The Astrophysical Journal Supplement Series, 283:35 (24pp), 2026 March https://doi.org/10.3847/1538-4365/ae3f24
© 2026. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Through Thick and Thin: The Cosmic Evolution of Disk Scale Height

Si-Yue Yu[1,2], Luis C. Ho[3,4], Takafumi Tsukui[2], John D. Silverman[2,5,6,7], Marc Huertas-Company[8,9,10,11,12], Anton M. Koekemoer[13], Maximilien Franco[14], Richard Massey[15], Lilan Yang[16], Rafael C. Arango-Toro[17], Andreas L. Faisst[18], Ghassem Gozaliasl[19,20], Kartik Sheth[21], Jeyhan S. Kartaltepe[16], Can Xu[2,22,23], Aryana Haghjoo[24], Xuheng Ding[2,25], Zhaoxuan Liu[2,5,6], and Jacqueline E. McCleary[26]

[1] Department of Astronomy, Xiamen University, Xiamen, Fujian 361005, People's Republic of China; syu@xmu.edu.cn
[2] Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of Tokyo Institutes for Advanced Study, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[3] The Kavli Institute for Astronomy and Astrophysics, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, People's Republic of China
[4] Department of Astronomy, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, People's Republic of China
[5] Department of Astronomy, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan
[6] Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[7] Center for Astrophysical Sciences, Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[8] Instituto de Astrofísica de Canarias (IAC), La Laguna, E-38205, Spain
[9] Observatoire de Paris, LERMA, PSL University, 61 avenue de l'Observatoire, F-75014 Paris, France
[10] Université Paris-Cité, 5 Rue Thomas Mann, 75014 Paris, France
[11] Universidad de La Laguna, Avda. Astrofísico Fco. Sanchez, La Laguna, Tenerife, Spain
[12] Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA
[13] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[14] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
[15] Department of Physics, Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK
[16] Laboratory for Multiwavelength Astrophysics, School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA
[17] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[18] Caltech/IPAC, MS 314-6, 1200 E. California Blvd. Pasadena, CA 91125, USA
[19] Department of Computer Science, Aalto University, P.O. Box 15400, FI-00076 Espoo, Finland
[20] Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland
[21] Empowered Earth Alliance, Washington, DC 20003, USA
[22] School of Astronomy and Space Science, Nanjing University, Nanjing, Jiangsu 210093, People's Republic of China
[23] Key Laboratory of Modern Astronomy and Astrophysics, Nanjing University, Ministry of Education, Nanjing 210093, People's Republic of China
[24] Department of Physics and Astronomy, University of California, Riverside, 900 University Ave, Riverside, CA 92521, USA
[25] School of Physics and Technology, Wuhan University, Wuhan 430072, People's Republic of China
[26] Department of Physics, Northeastern University, 360 Huntington Ave, Boston, MA, USA
Received 2025 November 27; revised 2026 January 21; accepted 2026 January 27; published 2026 March 3

Abstract

To investigate the formation and evolution of vertical structures in disk galaxies, we measure global sech^2 scale heights, averaging thin and thick components when present, for 2631 edge-on disk galaxies with $M_* > 10^{10} M_\odot$ at $0 < z < 3.5$ from the JWST COSMOS-Web survey. We show that dust extinction systematically overestimates scale heights at shorter rest-frame wavelengths and therefore adopt a fixed rest-frame wavelength of 1 μm. After further correcting for projection-induced bias using a new accurate method, we find that the median disk scale height increases from 0.56 ± 0.03 kpc at $z = 3.25$ to 0.84 ± 0.04 kpc at $z = 1.25$, and subsequently decreases to 0.67 ± 0.06 kpc at $z = 0.25$. The bias-corrected disk scale length-to-height ratio remains constant at 2.7 ± 0.2 for $z > 1.5$ but rises to 4.0 ± 0.4 at $z = 0.25$. These results imply that the high-redshift progenitors of present-day thick disks were of intermediate thickness, neither thin nor thick, yet dynamically hot and dense. The observed radial variation of scale height is consistent with the artificial flaring expected from observational effects, disfavoring minor mergers as the primary mechanism of disk thickening. Instead, we suggest that the high-redshift intermediate-thickness disks were single-component systems that increased their vertical scale height through decreasing surface mass density and/or violent gravitational instabilities, eventually producing thick disks. Thin-disk growth begins at $z \approx 2$ and dominates at $z \lesssim 1$, yielding a vertically more compact system with decreasing scale heights from $z \approx 1$ to 0. The inferred thin-disk mass fraction increases from 0.1 ± 0.03 at $z = 1$ to 0.6 ± 0.1 at $z = 0$. Together, these findings reveal a continuous evolutionary link between high-redshift single-component disks and present-day disks comprising thick and thin components.

Unified Astronomy Thesaurus concepts: Galaxy evolution (594); Milky Way evolution (1052); Galaxy structure (622); Galaxy stellar disks (1594); Galaxy photometry (611); Galaxy kinematics (602); High-redshift galaxies (734)

1. Introduction

Thick disks were first identified in 1979 as diffuse stellar components in edge-on S0 galaxies, distinct from both bulge and thin disks (D. Burstein 1979), and were subsequently recognized in the Milky Way in 1983 through star counts

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

toward the south Galactic pole (G. Gilmore & N. Reid 1983). These discoveries established the now standard view that the Milky Way disk comprises two components: a thin and a thick disk (K. Freeman & J. Bland-Hawthorn 2002; T. Bensby et al. 2014; J. Bland-Hawthorn & O. Gerhard 2016).

The formation history of the Milky Way's disk appears to have proceeded in two nearly disjoint phases (P. E. Nissen et al. 2020; M. Xiang & H.-W. Rix 2022; M. Xiang et al. 2025). The thin disk, with an exponential scale height of ∼0.3–0.4 kpc (M. Jurić et al. 2008; J. Bland-Hawthorn & O. Gerhard 2016), is dominated by younger stars and was formed from the cold gaseous midplane during the phase of steady star formation over the past ∼8 Gyr (J. Forbes et al. 2012; M. Xiang & H.-W. Rix 2022; S. Yu et al. 2023). The thick disk, with an exponential scale height of ∼0.7–1.2 kpc, consists primarily of old, [α/Fe]-enhanced stars that originated during an earlier, bursty star formation phase within the first ∼5 Gyr (M. Jurić et al. 2008; T. Bensby et al. 2014; J. Bland-Hawthorn & O. Gerhard 2016; S. Yu et al. 2021, 2023). Thick disks are now known to be nearly ubiquitous among nearby disk galaxies (P. Yoachim & J. J. Dalcanton 2006; D. Bizyaev & S. Mitronova 2009; D. V. Bizyaev et al. 2014; S. Comerón et al. 2014, 2016, 2018; G. Kauffmann et al. 2025).

Despite their prevalence, the physical origin of thick disks remains debated. They may have been born as inherently thick structures under turbulent, gas-rich conditions in the early Universe (F. Bournaud et al. 2009; S. Yu et al. 2023). Alternatively, they may have grown from initially thin disks through heating by minor mergers or internal secular processes (P. J. Quinn et al. 1993; Á. Villalobos & A. Helmi 2008; Y. Qu et al. 2011; J. C. Bird et al. 2021).

Studying galaxy disks across cosmic time provides crucial insight into the mechanisms that construct their vertical structure. In the Hubble Space Telescope (HST) era, the limited spatial resolution made it difficult to distinguish thin and thick disks, restricting studies to global scale-height measurements. Tracing the rest-frame optical to UV wavelength, B. G. Elmegreen et al. (2017) reported a median global scale height of 0.63 kpc, with a scatter of 0.24 kpc among individual measurements, at $z \sim 2$ (see also B. G. Elmegreen & D. A. Hunter 2006) and detected vertical color gradients that suggest the coexistence of thin and thick components at early epochs. A later analysis by K. A. Hamilton-Campos et al. (2023) yielded a median scale height of 0.74 kpc, with a scatter of 0.35 kpc, after correcting for the overestimation caused by the projection effect of deviations from perfect edge-on inclination; however, such corrections depend essentially on the ratio of scale height to scale length, a dependence explored in this study.

Progress in characterizing disk structure beyond $z \sim 1$ has recently accelerated thanks to the unprecedented sensitivity and resolution of JWST's NIRCam. JWST observations have revealed a substantial population of regular disk galaxies already in place within the first few Gyr, suggesting that disk assembly and evolution began earlier than previously thought (C. Cheng et al. 2022, 2023; L. Ferreira et al. 2023; C. Jacobs et al. 2023; J. S. Kartaltepe et al. 2023; E. J. Nelson et al. 2023; B. E. Robertson et al. 2023; Z. A. Le Conte et al. 2024; D. Xu & S.-Y. Yu 2024; M. Huertas-Company et al. 2025; S.-Y. Yu et al. 2025). Using F115W imaging that traces rest-frame near-infrared (NIR) at $z \sim 0$ but UV at $z > 3$ for a sample with a median stellar mass of $10^9 M_\odot$, J. Lian & L. Luo (2024) measured a median global scale height of 0.38 kpc, with a scatter of 0.13 kpc, at $z = 0.2$–5, lower than the previous HST results. T. Tsukui et al. (2025),[27] using the F277W, F356W, and F444W filters to trace the rest-frame 1–2 μm wavelength at $z \leqslant 3$, obtained a median global scale height of 0.66 ± 0.07 kpc (where all quoted values in this work follow the $\pm$ notation for uncertainties) for a stellar mass of $10^{9.2} M_\odot$, consistent with earlier HST results. Moreover, they performed the first thick-thin disk decomposition at $z > 1$, finding median scale heights of 0.33 ± 0.04 and 0.94 ± 0.07 kpc for the thin and thick components, respectively.

Still, as these pioneering JWST studies do not adopt a fixed rest-frame wavelength for the measurements, the derived scale heights may be biased by extinction from midplane dust lanes or by spatial variations in stellar populations. Furthermore, their analyses were based on relatively small samples (fewer than 200 galaxies). Expanding to larger, statistically robust samples is essential to establish the redshift evolution of disk scale heights. Equally important is correcting for the bias introduced by variation in the rest-frame wavelength and developing a more accurate method to correct for the bias introduced by deviations from a perfectly edge-on inclination. Motivated by these needs, we present a comprehensive analysis of edge-on disk galaxies observed by the COSMOS-Web survey (C. M. Casey et al. 2023). In this work, we focus on the global disk scale height, averaging thin and thick components when present, while future studies will address thin-thick disk decompositions.

This paper is organized as follows. Section 2 describes the sample selection and its properties. Section 3 details the image analysis and scale-height measurements. Section 4 presents the main results. Section 5 discusses the implications for the formation and evolution of galaxy disks. Finally, Section 6 summarizes the main findings. Throughout this paper, we adopt AB magnitudes, a G. Chabrier (2003) initial mass function, and assume a flat ΛCDM cosmology with $(\Omega_m, \Omega_\Lambda) = (0.27, 0.73)$ and a Hubble constant of $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2. Sample and Data

The COSMOS-Web survey is the largest galaxy survey conducted in JWST Cycle 1 (C. M. Casey et al. 2023). Its NIRCam observations in four filters (F115W, F150W, F277W, and F444W) provide a deep and high-resolution dataset for our structural and photometric analysis. For this study, we use the first data release of COSMOS-Web NIRCam mosaics[28] (M. Franco et al. 2025). The mosaics have a pixel scale of 0.″03 at all filters. M. Shuntov et al. (2025) constructed the COSMOS-Web photometric catalog by performing source extraction across 33 photometric bands spanning 0.3–8 μm and measured photometry across available bands using a multiband model-fitting approach with `SourceX-Tractor++` (E. Bertin et al. 2020; M. Kümmel et al. 2022). Using these measurements, M. Shuntov et al. (2025) derived physical properties for each source through template spectral energy distribution fitting using `LePhare` (S. Arnouts et al. 2002;

[27] Due to different definitions of scale height (their Equation (1)), the values reported by T. Tsukui et al. (2025) should be multiplied by a factor of 2 to match those presented here.

[28] https://cosmos2025.iap.fr

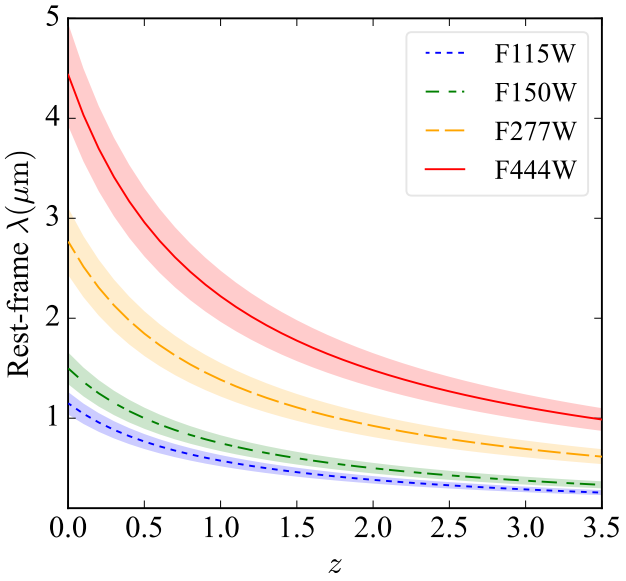


Figure 1. Observed rest-frame wavelengths of the NIRCam filters used in COSMOS-Web as a function of redshift. The shaded region marks the bandwidth of the filter. The upper redshift limit is chosen to ensure that the pivot wavelength of the F444W filter probes rest-frame wavelengths of at least 1 μm. We obtain the scale heights at a fixed rest-frame 1 μm by interpolating the measurements across the available filters (see Section 4.1).

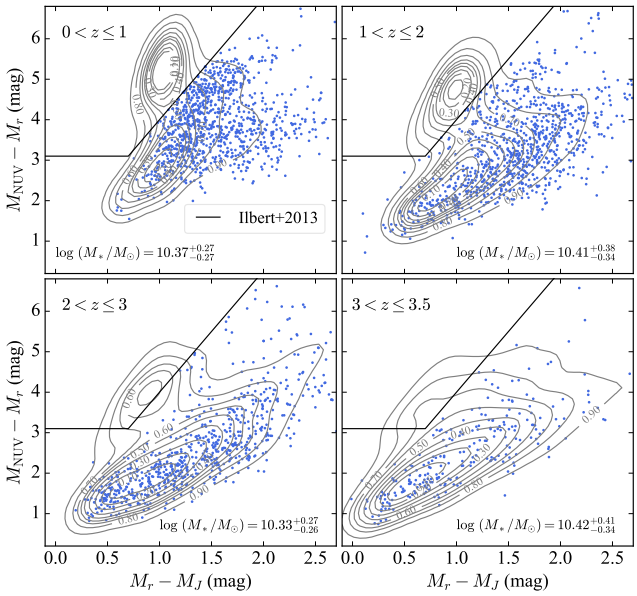


Figure 2. Diagram of $M_r - M_J$ versus $M_{\rm NUV} - M_r$ for our edge-on disk galaxy sample (blue dots). Contours indicate regions enclosing a given fraction of galaxies in the parent sample within each redshift range. Quiescent galaxies, which mostly lie above the dividing line proposed by O. Ilbert et al. (2013), are excluded. Median stellar mass and scatter for our sample for each redshift bin are shown at the bottom.

O. Ilbert et al. 2006). From the COSMOS-Web catalog (v1.1), we obtain photometric redshifts (z), stellar masses (M_*), and absolute magnitudes in the near-ultraviolet (NUV), r, and J bands ($M_{\rm NUV}$, M_r, and M_J). Figure 1 illustrates the rest-frame wavelength probed by each NIRCam filter as a function of redshift. In particular, the F444W filter observes the rest-frame 1 μm at $z = 3.5$, above which the rest-frame pivot wavelength drops below 1 μm. To minimize the impact of dust extinction on our morphological measurements, we require that the reddest filter probes at least 1 μm. We therefore limit our analysis to galaxies within $0 < z < 3.5$. The choice of the upper redshift limit also aligns with the epoch of the Milky Way's thick-disk formation (C. Conroy et al. 2022).

We further apply the following selection criteria to construct our parent sample. To ensure a clean and reliable galaxy sample, we require the following flags provided in the COSMOS-Web catalog: *type = 0* (select galaxies instead of stars), *flag_star_hsc = 0* (exclude sources severely contaminated by nearby stars), and *warn_flag = 0* (remove spurious detections such as hot pixels, sources detected in only a single filter, or other imaging artifacts). We restrict our analysis to galaxies with $M_* \geqslant 10^{10} M_\odot$, as recent studies suggest that lower-mass galaxies at $z > 1$ may be prolate rather than disklike systems (V. Pandya et al. 2024; J. Vega-Ferrero et al. 2024; C. Klein et al. 2025). Applying these criteria yields a parent sample of 19,199 galaxies.

We visually inspected each galaxy in the parent sample and excluded 1445 objects that were under interaction/merger or were too close to other galaxies or bright stars for reliable morphology measurements. This step resulted in a cleaned sample of 17,754 galaxies. For each object, we used the Python package `SEP` (E. Bertin & S. Arnouts 1996; K. Barbary 2016) on the F444W images to generate a target galaxy segmentation map and generate a mask file for removing contaminating sources. Mask generation was performed with both cold and hot detection modes, with parameters adjusted manually. These masks were used for subsequent model fitting.

Figure 2 shows $M_{\rm NUV} - M_r$ versus $M_r - M_J$ color–color diagrams for four redshift bins. The background contours indicate the distribution of the parent sample, with contours enclosing 10%, 20%, ..., 90% of galaxies. Quenched galaxies were identified using the division line (black line in Figure 2) proposed by O. Ilbert et al. (2013) and were excluded, leaving 13,453 star-forming galaxies for analysis. We first perform single Sérsic fitting using the code `IMFIT` (P. Erwin 2015) and remove unresolved sources with an effective radius in the F444W band less than the F444W point-spread function (PSF) FWHM. We then perform bulge-disk decomposition, modeling each galaxy with an exponential disk and a round Sérsic bulge with the Sérsic index fixed at 4, in all available filters. The generation of PSFs used in the fitting is described in Section 3.1. Catastrophic fitting results are excluded. From this decomposition, we calculate the bulge-to-total light ratio (B/T) and exclude potential elliptical galaxies, either introduced by photometric measurement uncertainties or representing star-forming ellipticals, by adopting a threshold of $B/T > 0.5$. Using a higher threshold (e.g., $B/T = 0.7$) does not affect our main results (see Section 4.3). We measure the axis ratio of the disk component ($q = b/a$, where a and b are the semimajor and semiminor axes, respectively) in each filter. Using interpolation, we derive the disk axis ratio at a consistent rest-frame wavelength of 1 μm for all galaxies, which we define as the final disk axis ratio (q_d). Edge-on disks appear as flattened systems and can be efficiently selected using q_d. This criterion is objective and reproducible, in contrast to visual classification. We therefore select our nearly edge-on disk sample by requiring $q_d < 0.4$, following the criterion adopted by K. A. Hamilton-Campos et al. (2023), which is slightly more permissive than the $q_d < 0.3$ threshold used by B. G. Elmegreen et al. (2017). As demonstrated in Section 4.2, our method for correcting biases introduced by

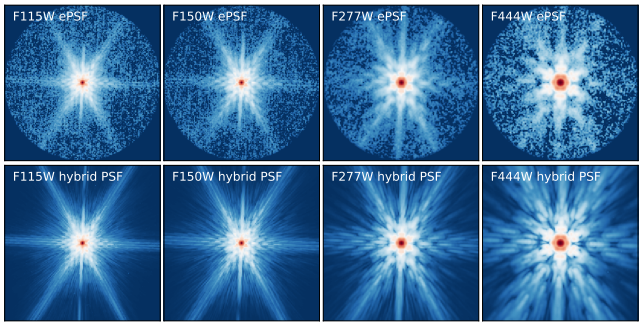


Figure 3. Comparison of the (top) ePSF and (bottom) our hybrid PSF for each JWST NIRCam filter of COSMOS-Web. The ePSF is constructed using PSFEX, while the hybrid PSF is constructed by replacing the low-S/N pixels in the ePSF with the values from simulated theoretical PSFs using a smooth top-hat function. These hybrid PSFs are adopted for all model fittings throughout this work.

projection effects is effective, such that adopting either $q_d < 0.3$ or $q_d < 0.4$ yields nearly identical bias-corrected scale heights. The above criteria reduce the 13,453 galaxies to 2631 edge-on disks, the sample studied in this work. This sample is 15–30 times larger than in previous studies.

In contrast to our PSF-corrected measurements, applying an axis ratio cut based on values derived without PSF correction (such as those from segmentation analysis), as done in some previous studies, can bias the sample toward intrinsically flatter systems at higher redshifts due to the smaller apparent sizes of galaxies. Our edge-on disk sample is shown in Figure 2 in the color–color diagram (blue dots), with median stellar mass values ($\sim 2.5 \times 10^{10} M_{\odot}$) and scatter indicated at the bottom. The stellar masses of our sample lie well above the completeness limits at each redshift (M. Shuntov et al. 2025).

We note that there is apparent clustering of galaxies at some specific redshifts, which is a known artifact of the spectral template fitting process for determining the photometric redshift. This can occur when prominent spectral features, such as the Balmer or 4000 Å break, transition between photometric bands. We have verified that the clustering is observed for both edge-on and face-on galaxies, indicating that the selection of edge-on disks does not introduce any unusual biases caused by the redshift determination. The overall quality of the photometric redshifts in our sample is high, with a catastrophic failure rate of only 1.44% and a normalized median absolute deviation of 0.011 compared to spectroscopic redshifts (M. Shuntov et al. 2025).

3. Measurements

3.1. Hybrid PSFs

Accurate PSFs are essential for robust morphological fitting of galaxies. Theoretical PSFs simulated with `WebbPSF` (M. D. Perrin et al. 2014) tend to be narrower than those derived from real stars in drizzled mosaics (Y. Ono et al. 2023; M.-Y. Zhuang & Y. Shen 2024), as `WebbPSF` is designed to simulate the PSF on a single exposure and should not be expected to agree with the PSF from drizzled mosaics, which include additional pixel-level convolutions. Since our analysis is based on drizzled mosaics, we construct empirical PSFs (ePSFs) from field-star images in each NIRCam filter using `PSFEx` (E. Bertin 2011), as shown in the top row of Figure 3. Still, these ePSFs suffer from relatively low signal-to-noise ratios (S/Ns) in their outer regions. A common strategy to improve PSF quality is to combine the high-S/N core of an ePSF with the noise-free wings of a theoretical one, forming a hybrid PSF (e.g., A. van der Wel et al. 2012; C.-H. Chen et al. 2025). Following this approach, we generate theoretical PSFs for each NIRCam filter using `WebbPSF`[29] and rotate them to match the orientation angle of the corresponding ePSFs by minimizing their difference. We then replace the noisy pixels (approximately the faintest 10% in flux) of the ePSF with values from the theoretical PSF. A smoothed step function and reversed step function are applied during this combination to ensure a seamless transition. Finally, we normalize the combined image to produce the hybrid PSF, as illustrated in the bottom row of Figure 3. These hybrid PSFs, which preserve the realistic core structure of observed stars while maintaining accurate outer profiles, are adopted for all model fittings throughout this work.

3.2. Scale-height Measurement

The scale height provides a quantitative measure of disk thickness, with larger values corresponding to thicker disks. In a coordinate system aligned with the edge-on inclination, we define the radial distance along the disk plane (parallel to the major axis) as R and the vertical distance from the plane (parallel to the minor axis) as h. The galaxy surface brightness at (R, h) is denoted as $I(R, h)$. Assuming an isothermal disk, the light distribution of an edge-on disk can be described by (P. C. van der Kruit & L. Searle 1981)

$$I(R, h) = I_0 \cdot (R/R_d) \cdot K_1(R/R_d) \cdot \mathrm{sech}^{2/m}(mh/h_0), \quad (1)$$

where I_0 is the central intensity, K_1 is the modified Bessel function of the second kind of order one, R_d is the exponential scale length in the disk plane, h_0 is the scale height, and m controls the overall shape of the vertical profile. Because m and h_0 are degenerate, m is usually fixed to 1 or $+\infty$ when determining the scale height, and Equation (1) then simplifies to the following two common forms: a sech^2 function,

$$I \propto \mathrm{sech}^2(h/h_0), \quad (2)$$

and an exponential function,

$$I \propto \exp(-2h/h_0). \quad (3)$$

Due to the above functional forms, it is generally approximated that the sech^2 scale height is twice the exponential scale height (e.g., P. Yoachim & J. J. Dalcanton 2006). However, this approximation is not exact, as the vertical profile of a galaxy is independent of the functional form adopted. J. Bland-Hawthorn & O. Gerhard (2016) suggest that the sech^2 scale height is 1.8 times the exponential scale height, whereas our analysis (see discussion later in this section) indicates that a factor of 1.37 provides a more accurate conversion for our sample.

Two-dimensional (2D) model fitting is an efficient approach for measuring scale heights (e.g., P. Yoachim & J. J. Dalcanton 2006; N. Ranaivoharimina et al. 2024; T. Tsukui et al. 2025). We use `IMFIT` (P. Erwin 2015) to perform the 2D fittings, incorporating an edge-on disk model and a bulge model. The hybrid PSFs generated in Section 3.1 are used. The bulge is modeled with a round Sérsic function

[29] https://stpsf.readthedocs.io/en/latest/

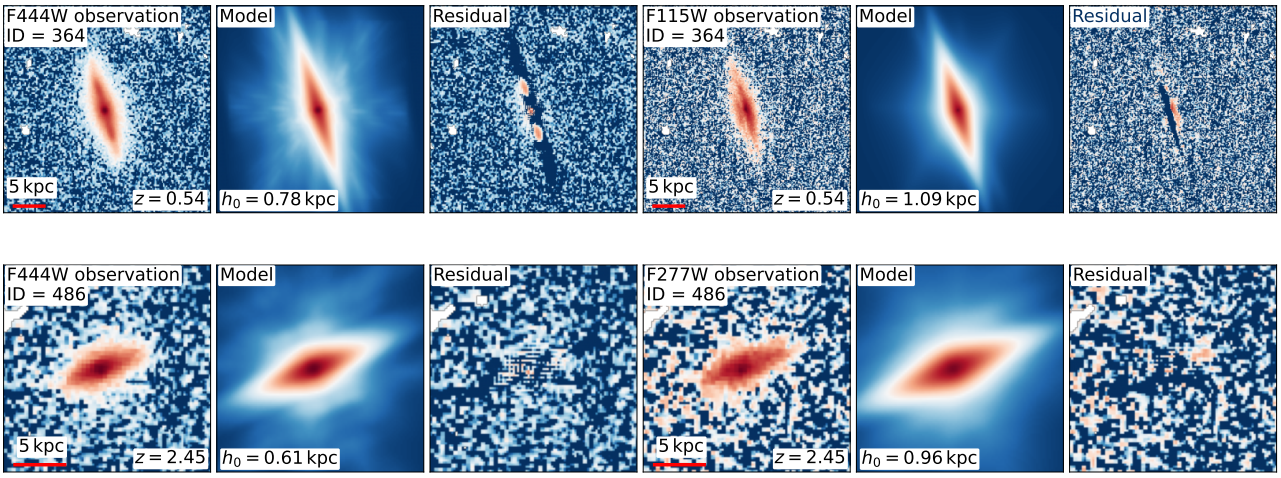


Figure 4. Example of the Sérsic-sech^2 model fitting. The top row shows a galaxy (ID = 364) at $z = 0.54$, with the left three panels and right three panels displaying the results for the F444W and F115W filters, respectively. The bottom row shows a galaxy (ID = 486) at $z = 2.45$, with results for the F277W and F115W filters. Scale bars are shown in the bottom-left corners. The derived sech^2 vertical scale height (h_0) is indicated in each case.

(J. L. Sersic 1968) with $n = 4$:

$$I(R) = I_e \exp\{-b_n[(R/R_e)^{1/n} - 1]\}, \quad (4)$$

where R_e is the half-light radius, I_e is the intensity at R_e, and b_n is a function of n determined by the incomplete gamma function.

We adopt the sech^2 vertical profile and the exponential radial profile for the edge-on disk model:[30]

$$I(R, h) = I_0 \cdot (R/R_d) \cdot K_1(R/R_d) \cdot \mathrm{sech}^2(h/h_0), \quad (5)$$

where h_0 is the sech^2 scale height. The Sérsic-sech^2 decomposition is carried out for each galaxy in all available filters. The galaxy center is a free parameter to fit, as this can mitigate the effects of dust lanes (see Section 4.1). The uncertainties of the best-fit parameters are provided by `IMFIT` as the statistical uncertainties obtained from evaluating the structure of the χ^2 minimum. The F115W and F150W images generally have lower S/Ns than the F277W and F444W images, as the stellar emission of galaxies typically peaks at the NIR. At high redshifts, some galaxies observed in F115W or F150W are indistinguishable from, or only marginally above, the background noise. To ensure robust measurements, we exclude any filter whose mean flux within the galaxy segmentation falls below the background noise level and instead adopt filters with sufficiently high S/Ns. The scale height derived from the Sérsic-sech^2 fitting represents the scale height averaging thin and thick components when present. Figure 4 presents representative fitting results for galaxies at $z = 0.54$ and $z = 2.45$. We further interpolate the measurements to a fixed rest-frame wavelength of 1 μm to mitigate the effects of dust extinction and spatial variation in stellar populations, as detailed in Section 4.1. The correction for projection biases arising from deviations from perfect edge-on orientation is described in Section 4.2.

For comparison, we also measure the exponential scale height ($h_{0,\mathrm{exp}}$). To minimize degeneracy between components, we fix the bulge parameters to those obtained from the Sérsic-sech^2 fit and then refit the image using an exponential vertical profile:

$$I(R, h) = I_0 \cdot (R/R_d) \cdot K_1(R/R_d) \cdot \exp(-h/h_{0,\mathrm{exp}}). \quad (6)$$

We use $h_{0,\mathrm{exp}}$ instead of $h_{0,\mathrm{exp}}/2$ in Equation (3), as $h_{0,\mathrm{exp}}$ is more commonly adopted in the literature. Figure 5 presents the correlation between the two measurements, which is well described by a tight linear relation given by

$$\mathrm{sech}^2\ h_0 = 1.37 \times \mathrm{exponential}\ h_{0,\mathrm{exp}}. \quad (7)$$

The Pearson correlation coefficient between sech^2 h_0 and the exponential $h_{0,\mathrm{exp}}$ is $r_\mathrm{p} = 0.99$, indicating a statistically significant correlation (p-value < 0.01). Therefore, our results remain unaffected by whether the vertical profile is described using a sech^2 or an exponential function. In the rest of this work, the term "scale height" denotes the sech^2 scale height unless the exponential scale height is specified.

In addition to the 2D method, one-dimensional (1D) fitting has also been used to measure scale heights (B. G. Elmegreen et al. 2017; K. A. Hamilton-Campos et al. 2023; N. Ranaivoharimina et al. 2024). The 1D method is able to obtain scale height as a function of radius. We evaluate the consistency between h_0 derived using 1D and 2D approaches. For the 1D method, we first subtract the best-fit bulge component (if present) and rotate the galaxy so that its major axis aligns with the image x-axis. We then extract vertical intensity profiles along the galaxy major axis. Profiles with peak intensities below 3 times the background noise are excluded. Each vertical profile is fitted with a 1D sech^2 function, convolved with the line-spread function (LSF), from which we derive the radial profile of the scale height. The LSF is constructed by convolving a one-pixel line with a PSF following the strategy in B. G. Elmegreen et al. (2017). The profiles are stacked to create a composite profile. A final sech^2 fit to the stacked profile yields the average h_0 from the 1D method, with weighting similar to that of the 2D fitting.

A comparison of the 1D and 2D measurements is shown in Figure 6. The median difference between the two is 0.04 kpc, with a standard deviation of 0.05 kpc. These small values indicate that the scale heights derived from the 1D and 2D

[30] Our measured scale height from Equation (5) is twice that derived from the default sech^2 function in `IMFIT`.

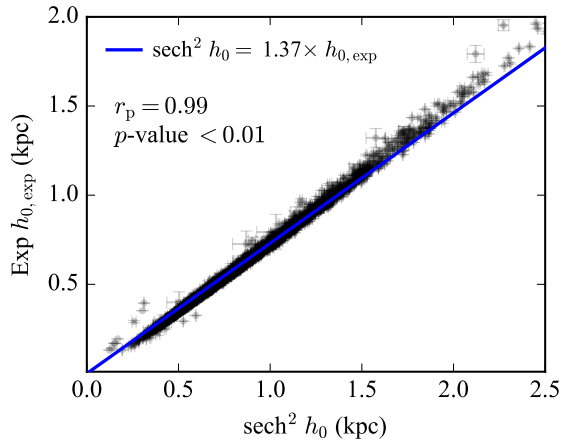


Figure 5. Correlation between the scale heights derived from the sech^2 function (h_0) and the exponential function ($h_{0,\rm exp}$). The two measurements are tightly correlated, with h_0 being on average 1.37 times as large as $h_{0,\rm exp}$.

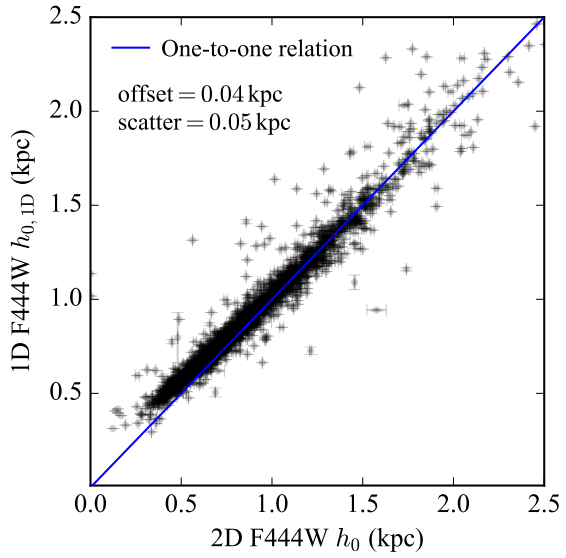


Figure 6. Comparison of sech^2 scale heights derived from 1D ($h_{0,\rm 1D}$) and 2D (h_0) fitting methods. The median difference and corresponding standard deviation are displayed at the top.

methods are consistent when similar weighting schemes are applied. The radial variation of the scale height will be discussed further in Section 4.

4. Results

4.1. Wavelength Dependence of Scale Height and the Correction

Since younger stellar populations dominate the thin disks of galaxies, observations at shorter rest-frame wavelengths without dust extinction are expected to reveal thinner disks with smaller scale heights. However, for edge-on systems, particularly the actively star-forming disks in our sample, dust lanes along the midplane substantially attenuate the flux, especially at shorter wavelengths. This obscuration impacts the apparent vertical light profile and can overestimate the measured scale height (D. Bizyaev & S. Mitronova 2009). As described by Equation (5), the scale height is defined as the height at which the intensity falls to sech^2(1) of the central value, rather than as an absolute geometric thickness. When the midplane flux is suppressed by a dust lane, the observed profile appears artificially broadened, leading to an overestimation of the intrinsic scale height.

In Figure 7, we compare the vertical scale heights (h_0) measured in the F444W band with those measured in F277W, F150W, and F115W. The scale heights derived from the bluer filters are systematically larger than those from redder filters. Quantitatively, the h_0 measured in F277W, F150W, and F115W exceeds those in F444W by 0.12 ± 0.11, 0.29 ± 0.28, and 0.30 ± 0.29 kpc, respectively.

To visualize how dust attenuation overestimates the measured scale height, we plot in Figure 8 examples of the 1D vertical profile fitting (see Section 3.2 for the method) for a galaxy observed in F444W (top) and F115W (bottom). The first column shows the bulge-subtracted images, where the prominent dust lane in F115W is clearly visible. The second column presents the radial profile of h_0, showing systematically larger values in F115W than in F444W, consistent with the 2D fitting results for the same galaxy shown in Figure 4. The third column shows the corresponding LSFs, while the final column compares the vertical light profiles. The F444W vertical light profile is symmetric and smooth, whereas the F115W profile exhibits a flux depression caused by the dust lane near the center. This depression shifts the apparent center and broadens the observed distribution, producing an overestimated scale height. When all fitting parameters are allowed to vary freely, we obtain $h_0 = 0.71$ kpc in F444W and $h_0 = 0.99$ kpc in F115W. If the galaxy center in F115W is fixed to that of F444W, the fitted value in F115W increases further to 1.37 kpc. Setting the center free for the fitting can mitigate the effect caused by the dust lane, which motivates our choice of leaving the center unconstrained in the two-component decomposition into a bulge and an edge-on disk (Section 3). In addition to dust extinction, spatial variations in the stellar population, such as younger stars concentrated near the midplane and older stars located at larger vertical distances, can in principle also affect the scale height measured at different wavelengths, although this effect is less significant than dust extinction.

Figure 9 presents the dependence of the measured h_0 on the rest-frame wavelength, with results shown separately for $0 < z \leqslant 1$ (top) and $1 < z < 3.5$ (bottom). The median values and their uncertainties reveal a clear trend: scale heights systematically decrease toward longer rest-frame wavelengths, where dust extinction becomes weaker. This dependence on rest-frame wavelength should have also incorporated the effects of spatial variations in the stellar population. The measured h_0 at the rest-frame 5000 Å wavelength is overestimated by roughly 30–40%, whereas at the rest-frame 1 μm wavelength, the impact of extinction is small.

This wavelength dependence therefore acts as a confounding factor when a single filter is used to study the redshift evolution of scale height, if the filter traces rest-frame optical to UV wavelengths at a high redshift but NIR wavelengths at a low redshift. To isolate the true redshift evolution, we determine the h_0 for each galaxy at a fixed rest-frame wavelength of 1 μm ($h_{0,1\ \mu\rm m}$) by fitting a power law to the h_0 versus rest-frame wavelength, using the uncertainties from 2D fitting as weights. A similar wavelength dependence affects the disk scale length, R_d, owing to the inside-out growth of

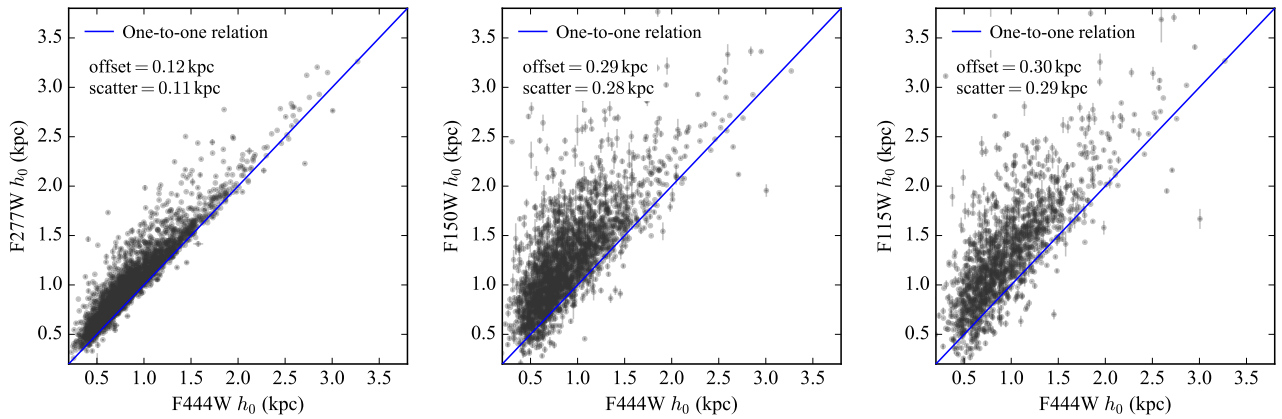


Figure 7. Correlation between scale heights measured in F444W and those in F277W, F150W, and F115W. The blue line marks the one-to-one relation. The median offset relative to F444W and scatter are indicated.

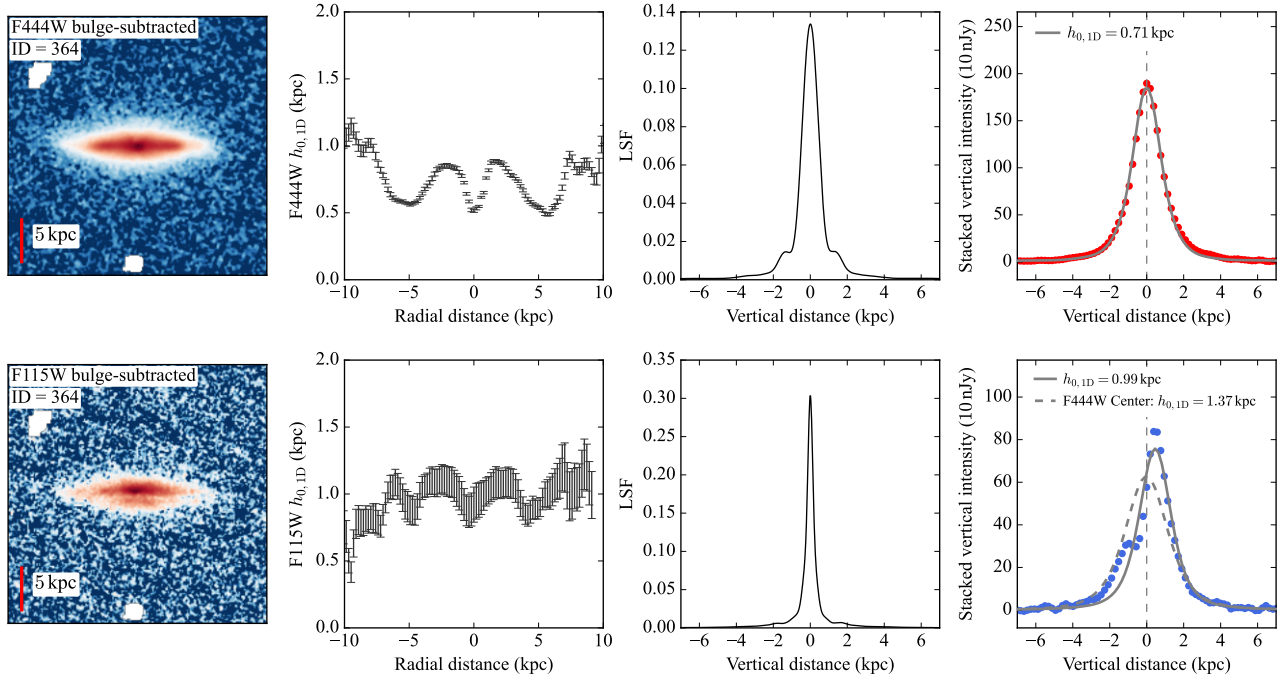


Figure 8. Illustration explaining why the scale height derived from the bluer filter appears thicker than that from the redder filter. From left to right: the bulge-subtracted image, the radial profile of scale height, the LSF constructed from the PSF, and the 1D sech^2 fit to the stacked vertical light profile. The top and bottom rows correspond to the F444W and F115W filters, respectively. In the final column, solid curves mark the best fits with all parameters free, and the dashed curve marks the fit with the center fixed to that in F444W.

galaxies (A. van der Wel et al. 2014; S. J. Lilly & C. M. Carollo 2016). To ensure consistency, we likewise derive R_d at 1 μm, $R_{d,1\ \mu\mathrm{m}}$. The effective radius of the disk component at 1 μm ($R_{e,\mathrm{disk},1\ \mu\mathrm{m}}$) is converted from $R_{d,1\ \mu\mathrm{m}}$.

We adopt 1 μm as the reference wavelength as extinction effects are largely mitigated at this wavelength, and it remains accessible at high redshifts with the F444W filter. At longer rest-frame wavelengths ($\gtrsim$1.5 μm), h_0 gradually converges to values $\sim$10% lower than $h_{0,1\ \mu\mathrm{m}}$ at both low and high redshifts. Nevertheless, by anchoring both h_0 and R_d to the same rest-frame wavelength, our analysis mitigates wavelength-dependent effects, including those caused by dust extinction and stellar population effects, and enables a clean and uniform comparison of disk structure across redshift. The measurements of scale heights and scale lengths can be accessed at Zenodo (doi:10.5281/zenodo.18059559 ; S.-Y. Yu 2026).

4.2. Projection Effects of Deviations from a Perfect Edge-on Inclination and the Bias Correction

A randomly oriented three-dimensional (3D) galaxy disk is unlikely to be perfectly edge on. A deviation from perfect edge-on inclination ($i = 90°$) amplifies the scale height measured in a 2D projection image, which is done by assuming a perfect edge-on inclination. This constitutes a projection effect overestimating the scale height. Consequently, correcting for this effect is essential to derive the intrinsic scale heights and their intrinsic cosmic evolution, particularly for our sample, which is selected based on a cut in the observed disk axis ratio $q_d < 0.4$.

A constant correction factor of 0.78 has been proposed to address this bias (K. A. Hamilton-Campos et al. 2023). However, this correction is not accurate enough, because the influence of the deviation from a perfect edge-on inclination

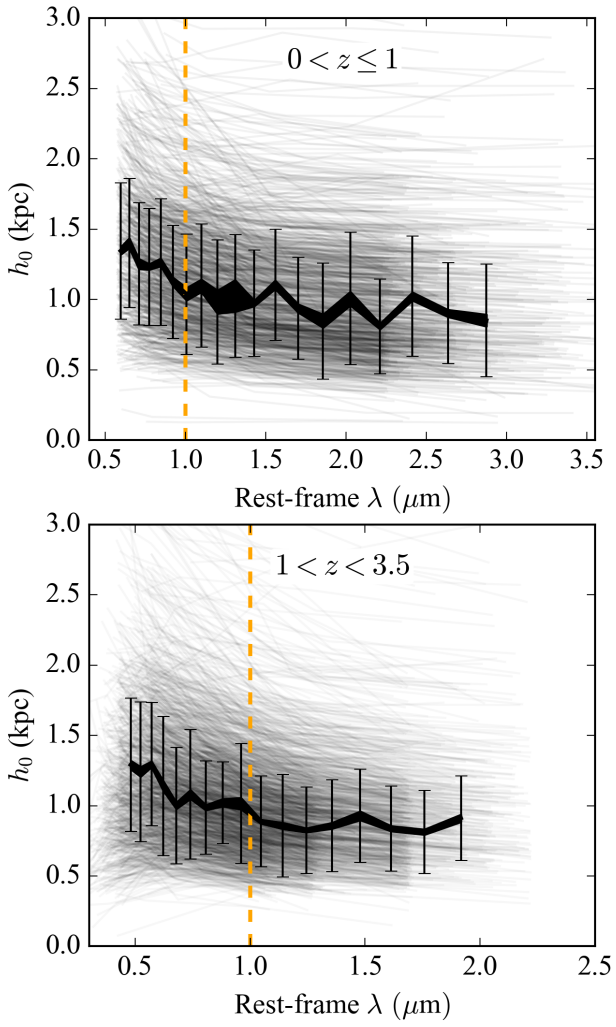


Figure 9. Measured scale heights (h_0) as a function of rest-frame wavelength. Results for individual galaxies are shown as background gray curves. The data are grouped into 20 bins in the rest-frame wavelength. The center and width of the black region (curve) indicate the median values and their uncertainties, while the error bars represent the scatter among individual measurements. The vertical orange dashed line marks the rest-frame 1 μm. The top and bottom panels show results for $0 < z \leqslant 1$ and $1 < z < 3.5$, respectively. Dust extinction produces a central depression in the vertical light profile, especially at a short rest-frame wavelength, which causes the fitted scale height to be overestimated.

on a scale-height measurement actually depends on the relative scale height of the disk, demonstrated below. The relative scale height (or relative thickness) is defined as

$$\text{relative scale height} = h_0/R_d. \tag{8}$$

We generate simulated projected 2D images of disks at specific inclination by using IMFIT (P. Erwin 2015), which performs line-of-sight integration of 3D disk models with exponential radial and sech^2 vertical profiles. For illustration purposes, we consider two disks with $h_0/R_d = 0.4$ and 0.2, respectively, and rotate them at $i = 90°$, 85°, 80°, and 75°. The projected 2D images with intrinsic and observed values of the disk parameters are shown in Figure 10. For the model with $h_0/R_d = 0.4$, when inclination is set from $i = 90°$ to $i = 75°$, h_0 is amplified from 0.5 to 0.8 kpc, and q_d is amplified from 0.246 to 0.389. But for $h_0/R_d = 0.2$, h_0 is amplified from 0.5 to 1.25 kpc, and q_d is amplified from 0.123 to 0.307. We can learn two key points from this test: (1) the scale height of a disk with a lower relative scale height (e.g., lower h_0 at fixed R_d, or larger R_d at fixed h_0) suffers more severely from the projection effect; (2) a disk with a lower relative scale height is less likely to meet the cut in q_d when selecting a sample.

Next, we quantify the impact of this projection effect and derive a new bias correction method. We use IMFIT to construct a suite of disk models spanning a range of h_0/R_d and inclination angles. For each model, we perform fits to measure h_0, R_d, and q_d. The measurement of R_d remains robust, but h_0 is overestimated. To derive the correction for measured h_0, we apply our sample selection criterion by excluding models with $q_d \geqslant 0.4$ and compute the median inclination angle and median observed h_0, weighted by the inclination probability distribution $P(i) \propto \sin(i)$.

The results are presented in Figure 11. The top-left panel illustrates that the median observed value of h_0 of disks with larger intrinsic h_0/R_d are less overestimated. The top-middle panel demonstrates that disks with larger intrinsic h_0/R_d ratios can have smaller maximum deviations from $i = 90°$ before exceeding the q_d threshold, simply because a relatively thinner disk has a smaller axis ratio. The top-right panel presents the median inclination angle as a function of intrinsic h_0/R_d. The bottom-left panel shows the conversion between intrinsic and median observed h_0/R_d. The bottom-middle and bottom-right panels show the derived statistical correction factor $f_{\rm corr}$, defined as

$$f_{\rm corr} = \frac{\text{median intrinsic } h_0}{\text{median observed } h_0}, \tag{9}$$

plotted as a function of intrinsic and median observed h_0/R_d, respectively. Disks with larger h_0/R_d ratios exhibit correction factors closer to 1, because (1) their h_0 are less affected for a given inclination compared to those with lower intrinsic h_0/R_d (see Figure 10) and (2) their maximum deviation from a perfect edge-on inclination before meeting the cut in q_d is smaller (top-middle panel). The variation in $f_{\rm corr}$ highlights the importance of the newly developed correction method, as a redshift evolution of h_0/R_d indeed exists, as shown in the next section. Table 1 shows the parameters for deriving the correction factors in samples selected using the axis ratio cut $q_d < 0.4$ as in this work. Table 2 provides values for samples selected using $q_d < 0.3$. If a smaller critical axis ratio is used, the $f_{\rm corr}$ is closer to unity.

The application of our bias correction is straightforward: we first compute the median observed h_0/R_d and determine the corresponding $f_{\rm corr}$, and then obtain the corrected median h_0 by multiplying the observed value by $f_{\rm corr}$. As described above, the correction curve is derived for a given intrinsic h_0/R_d. In real observations, however, disks exhibit a distribution of h_0/R_d rather than a single value. To assess the effectiveness of our method under real observational conditions, we create four groups of models, each containing 10,000 models with random inclinations. The distributions of R_d, h_0, and h_0/R_d for these four groups are shown in Figure 12. A truncation is observed at $h_0/R_d = 0.65$ in the h_0/R_d distribution, as models with h_0/R_d above this threshold exhibit an axis ratio $q_d \geqslant 0.4$ even in a perfectly edge-on inclination. For each inclined model, the observed h_0 and h_0/R_d are measured using 2D sech^2 fitting.

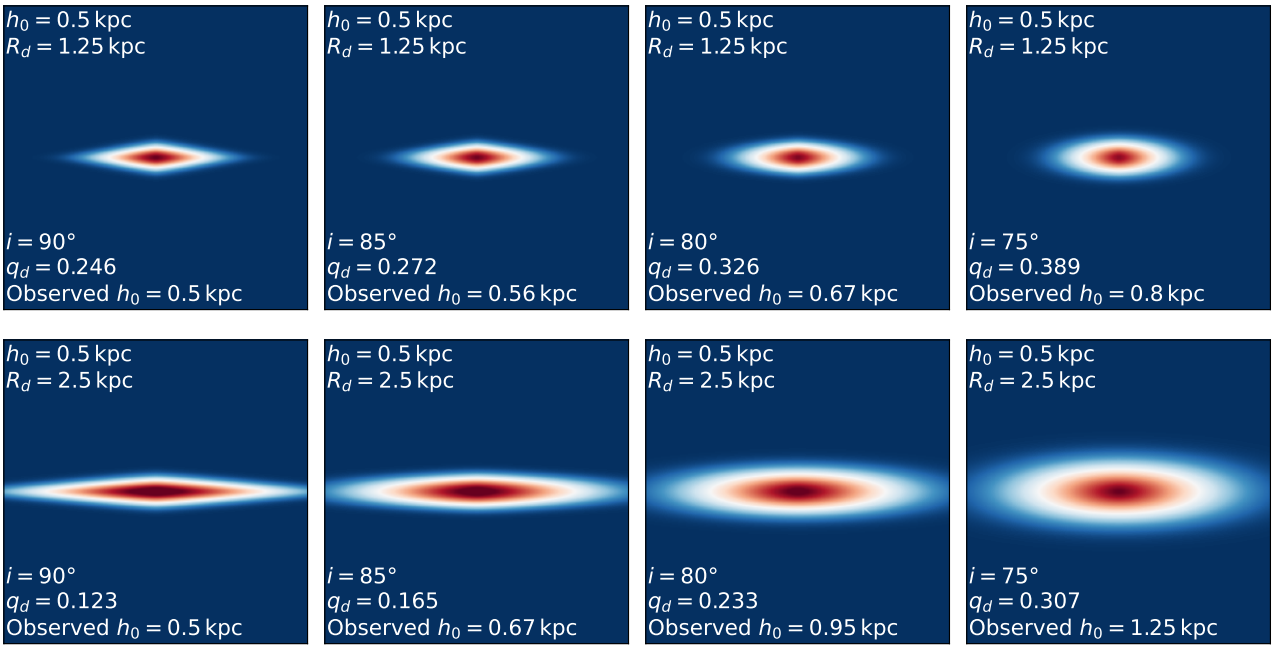


Figure 10. Illustration of the projection effect caused by a deviation from a perfect edge-on inclination on scale-height h_0 measurements. The top and bottom rows present the analysis for models with intrinsic relative scale height $h_0/R_d = 0.4$ and $h_0/R_d = 0.2$, respectively. Columns 1 to 4 correspond to inclination angles of $i = 90°$, 85°, 80°, and 75°, respectively. The intrinsic properties are listed at the top in each panel, while the observed properties are listed at the bottom. The observed h_0 is more strongly overestimated when the disk deviates more from $i = 90$ and has a lower relative scale height.

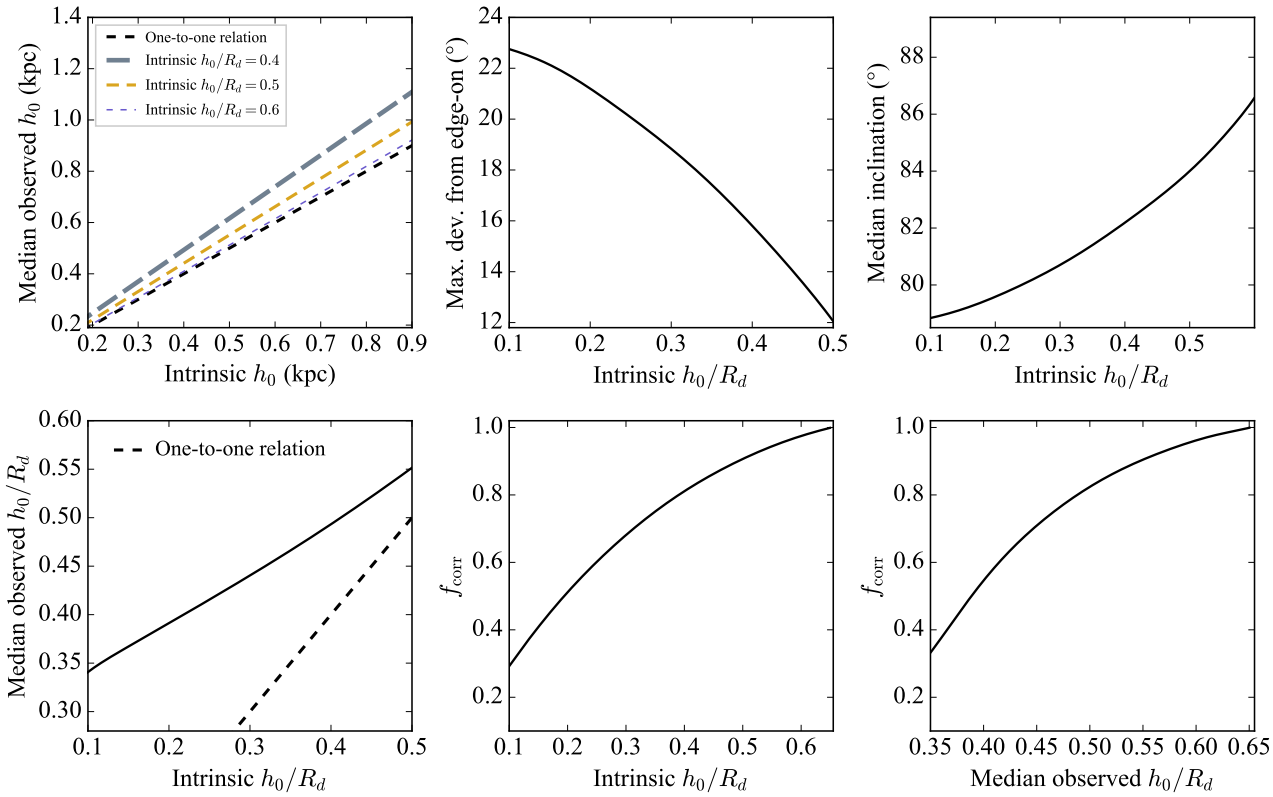


Figure 11. Projection effects of the deviation from a perfect edge-on inclination on scale-height measurement and the correction curve in a disk galaxy sample with $q_d < 0.4$. The top row shows the relation between the median observed and intrinsic h_0 (left), the dependence of the maximum deviation from a perfect edge-on inclination required for a disk to satisfy $q_d < 0.4$ (middle), and the median inclination angle (right) on intrinsic h_0/R_d. The bottom row shows the median observed h_0/R_d (left) and the correction factor $f_{\rm corr}$ (middle) as functions of intrinsic h_0/R_d and $f_{\rm corr}$ as a function of median observed h_0/R_d (right).

For each group, we calculate the median observed h_0 and h_0/R_d, and apply our new correction method to obtain the bias-corrected h_0.

Figure 13 presents a comparison of the median measured h_0, with and without correction, against the intrinsic values. The results without correction are shown as crosses. As expected,

Table 1
Parameters for Deriving the Correction Factors for Removing the Projection Effect Caused by Deviations from Edge-on Orientation in Samples Selected with an Axis Ratio Cut $q_d < 0.4$

Observed h_0/R_d	Intrinsic h_0/R_d	Max. Deviation i from Edge On (deg)	Median i (deg)	$f_{\rm corr}$
(1)	(2)	(3)	(4)	(5)
0.350	0.116	22.6	78.9	0.332
0.375	0.166	21.9	79.3	0.442
0.400	0.218	20.8	79.8	0.546
0.425	0.270	19.6	80.3	0.635
0.450	0.319	18.3	81.0	0.709
0.475	0.366	16.9	81.7	0.771
0.500	0.412	15.4	82.4	0.824
0.525	0.456	13.8	83.1	0.868
0.550	0.498	12.2	83.9	0.905
0.575	0.538	10.3	84.9	0.935
0.600	0.577	8.3	85.9	0.962
0.625	0.614	5.9	87.1	0.982
0.650	0.649	1.8	89.2	0.999

Note. Column (1): median observed relative scale heights. Column (2): intrinsic relative scale heights. Column (3): maximum deviation from a perfect edge-on inclination before reaching $q_d \geqslant 0.4$. Column (4): median inclination. Column (5): correction factor for the median observed scale height accounting for bias due to deviation from a perfect edge-on orientation.

Table 2
Parameters for Deriving the Correction Factors for Removing the Projection Effect Caused by Deviations from Edge-on Orientation in Samples Selected with an Axis Ratio Cut $q_d < 0.3$

Observed h_0/R_d	Intrinsic h_0/R_d	Max. Deviation i from Edge On (deg)	Median i (deg)	$f_{\rm corr}$
(1)	(2)	(3)	(4)	(5)
0.26	0.090	16.6	81.8	0.348
0.28	0.126	16.1	82.0	0.450
0.30	0.166	15.3	82.4	0.553
0.32	0.206	14.3	82.9	0.645
0.34	0.246	13.3	83.4	0.722
0.36	0.283	12.1	84.0	0.787
0.38	0.319	10.9	84.6	0.839
0.40	0.353	9.7	85.2	0.883
0.42	0.387	8.4	85.8	0.920
0.44	0.418	6.8	86.6	0.951
0.46	0.448	5.1	87.5	0.975
0.48	0.477	2.7	88.7	0.993

Note. Column (1): median observed relative scale heights. Column (2): intrinsic relative scale heights. Column (3): maximum deviation from a perfect edge-on inclination before reaching $q_d \geqslant 0.3$. Column (4): median inclination. Column (5): correction factor for the median observed scale height accounting for bias due to deviation from a perfect edge-on orientation.

the observed h_0 values are systematically overestimated compared to their intrinsic values when no correction is applied. Furthermore, the degree of overestimation varies across different model groups, suggesting there is no constant factor that can entirely correct for the effects. If a constant correction factor of 0.78 is applied, as suggested by K. A. Hamilton-Campos et al. (2023), the corrected h_0 values (shown as diamonds) do not reproduce the intrinsic values. In contrast, our correction method, represented by solid points, yields h_0 values that align closely with the one-to-one relation, demonstrating the effectiveness of our approach in accounting for the projection effect caused by deviations from a perfect edge-on inclination. We emphasize that our method is statistical in nature and can be applied only to a sample, not to an individual source.

4.3. Redshift Evolution of Bias-corrected Scale Heights

Before examining the bias-corrected scale heights, we first probe the uncorrected measurements derived from the F444W, F277W, F150W, and F115W filters as a function of redshift (Figure 14). Only a weak trend is seen in the F444W filter, where the scale height increases slightly toward lower redshifts and then remains roughly constant. At high redshifts, the number of galaxies included in the F115W and F150W measurements decreases because in these filters the galaxies become too faint in surface brightness to yield reliable measurements and are therefore excluded (see Section 3).

To remove the bias introduced by variations in the rest-frame wavelength, we have computed the scale height ($h_{0,1\ \mu\rm m}$), scale length ($R_{d,1\ \mu\rm m}$), and relative scale height ($h_{0,1\ \mu\rm m}/R_{d,1\ \mu\rm m}$) at a fixed rest-frame wavelength of 1 μm (see Section 4.1). Their redshift dependence is shown in Figure 15. The relation between $h_{0,1\ \mu\rm m}$ and redshift is more clearly defined than that derived from single-band measurements, as biases caused by variations in the rest-frame wavelength, such as dust extinction and stellar population effects, have been removed. The $R_{d,1\ \mu\rm m}$ exhibits a general increase toward lower redshifts, consistent with the well-established growth of galaxy sizes over cosmic time (A. van der Wel et al. 2014; N. Allen et al. 2025; L. Yang et al. 2025). The observed $h_{0,1\ \mu\rm m}/R_{d,1\ \mu\rm m}$ decreases with lower redshifts, resulting in smaller values of $f_{\rm corr}$ at low redshifts. This indicates that the measured scale heights are more overestimated at lower redshifts than at higher redshifts.

To remove the bias introduced by the projection effect, we divide the sample into seven redshift bins and compute the median observed value of $h_{0,1\ \mu\rm m}/R_{d,1\ \mu\rm m}$ in each bin. These values are used as inputs to determine the $f_{\rm corr}$ (see Figure 11 or Table 1), which ranges from 0.6 to 0.8. $f_{\rm corr}$ is then applied to $h_{0,1\ \mu\rm m}$ to obtain the bias-corrected scale height. The bias-corrected $h_{0,1\ \mu\rm m}$ as a function of redshift are presented in Figure 15 (bottom right) and are listed in Table 3. Therefore, after correcting for the effects of variation in the rest-frame wavelength and the projection effect, the median and scatter of the scale heights are 0.56 ± 0.03, 0.61 ± 0.04, 0.72 ± 0.03, 0.78 ± 0.03, 0.84 ± 0.04, 0.82 ± 0.03, and 0.67 ± 0.06 kpc at $z = 3.25$, 2.75, 2.25, 1.75, 1.25, 0.75, and 0.25, respectively. The scatters among individual measurements are 0.2–0.3 kpc. The overall trend shows an increase in scale height from $z \sim 3$ to $z \sim 1$, followed by a decline toward lower redshifts. The bias-corrected ratios of $h_{\rm vert,1\ \mu\rm m}$ to $R_{d,1\ \mu\rm m}$ and their inverse ratios are also presented in Table 3.

Furthermore, to examine whether, during sample selection, the adopted cut in B/T used to remove potential elliptical galaxies, which may fall within the star-forming region of the color–color diagram due to measurement uncertainties or are intrinsically star-forming, affects our results, we plot the redshift evolution of the bias-corrected $h_{0,1\ \mu\rm m}$ using a higher B/T threshold of 0.7, shown as the open circles in Figure 15. The result does not change significantly, suggesting that our conclusions are insensitive to adopting a higher B/T threshold.

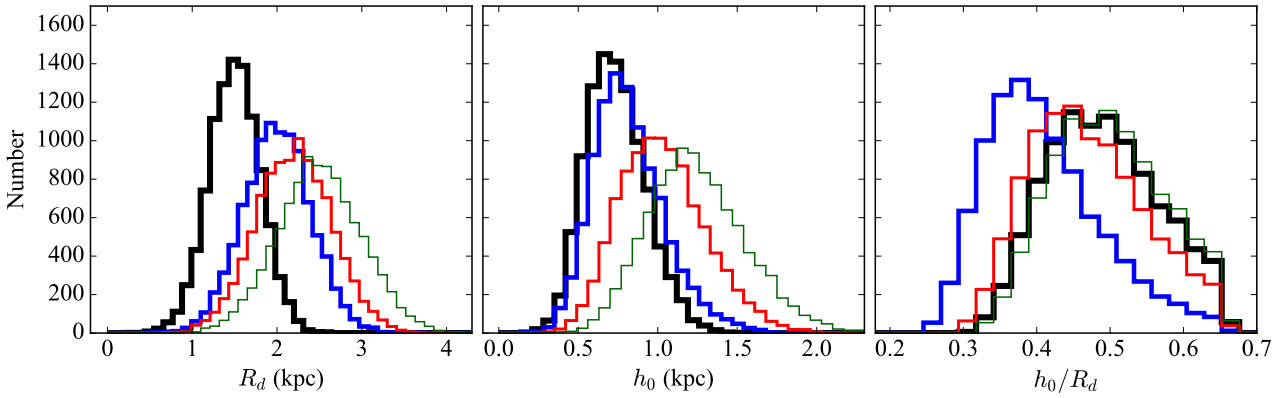


Figure 12. Distributions of R_d, h_0, and h_0/R_d of the four groups of disk models used to verify the effectiveness of our correction method. The four groups are marked in black, blue, green, and red. Each group consists of 10,000 models. A truncation at $h_0/R_d = 0.65$ is applied, as models with values above this threshold exhibit an axis ratio $q_d > 0.4$ even for a perfect edge-on inclination.

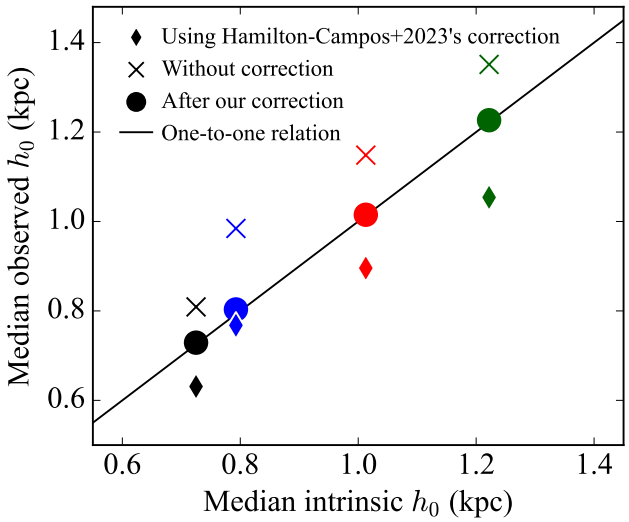


Figure 13. Comparison of the median measured scale height (h_0), with and without correction, against the intrinsic values. Symbol colors correspond to the models shown in Figure 12. Crosses represent results without correction, diamonds indicate results using the correction from K. A. Hamilton-Campos et al. (2023), and points show results from our new correction method.

4.4. Artificial Radial Increase of Measured Scale Heights

The projection effect of a deviation from a perfectly edge-on inclination not only causes a systematic overestimation of the global vertical scale height but also produces an artificial radial increase of the measured height toward larger radii, as previously noted by D. V. Bizyaev et al. (2014). In addition, the LSF cannot fully account for the light from the bright central region scattered into the outskirts of the disk, leaving a residual PSF effect that introduces further artificial flaring. Figure 16 illustrates these effects using 3D model disks with an intrinsic constant scale height. The measured radial profiles are obtained by fitting 1D vertical light profiles, convolved with an LSF if needed, at each radius in the 2D projected image. For a perfectly edge-on disk without PSF convolution, the measured profile normalized by the central scale height remains flat at 1. When a PSF is convolved, however, the measured scale height becomes artificially larger even though an LSF is applied when doing the 1D fitting (see Section 3). For a disk at $i = 70°$ without PSF convolution, the measured scale height artificially increases with radius due to projection effects; when the PSF is convolved, this artificial increase becomes slightly stronger owing to the combined impact of projection and residual PSF effects.

To examine whether the disk galaxies in COSMOS-Web exhibit genuine or artificial radial gradients, we measured the vertical scale height at each radius using the 1D method described in Section 3. We refrain from interpolating to rest-frame 1 μm because the 1D profiles at individual radii, extracted from 2D flux maps, have relatively lower S/Ns than global measurements, which would propagate larger uncertainties through interpolation. Instead, we adopted results from the filter, probing rest-frame wavelengths just above 1 μm, where the effects of dust extinction are small. The galaxies were grouped into two redshift bins: $0 < z \leqslant 1$ and $1 < z \leqslant 3.5$. For each galaxy, the radial profile of the scale height was normalized by its central value within one FWHM of the PSF, and the radius was normalized by the disk scale length R_d. The resulting normalized profiles are shown in Figure 17, where large points and error bars indicate the median and scatter, respectively. In both redshift ranges, the measured scale height systematically increases with radius, giving the appearance of disk flaring.

However, as shown above, such a trend can arise purely from the projection effect and the residual PSF effect. To understand whether this observed increase is physical or artificial, we estimated the median observed $h_{0,1\ \mu\mathrm{m}}/R_{d,1\ \mu\mathrm{m}}$ for each redshift bin and derived the corresponding intrinsic $h_{0,1\ \mu\mathrm{m}}/R_{d,1\ \mu\mathrm{m}}$ and median inclination. Using these parameters, we generated PSF-convolved 2D projected models from 3D model disks with a constant intrinsic scale height. The 1D fitting method yields the predicted artificial radial trends shown as blue solid curves in Figure 17. The predicted trends closely reproduce the observed radial profiles at both low and high redshifts. This agreement indicates that the apparent flaring is primarily a projection artifact, with a minor contribution from residual PSF effects, rather than a genuine structural feature. The implications of this result for the origin of disk scale heights are discussed in Section 5.

4.5. Comparison with Scale-height Measurements in Previous Works

There are several studies on the redshift evolution of disk scale height based on HST or JWST images (B. G. Elmegreen & D. M. Elmegreen 2006; B. G. Elmegreen et al. 2017;

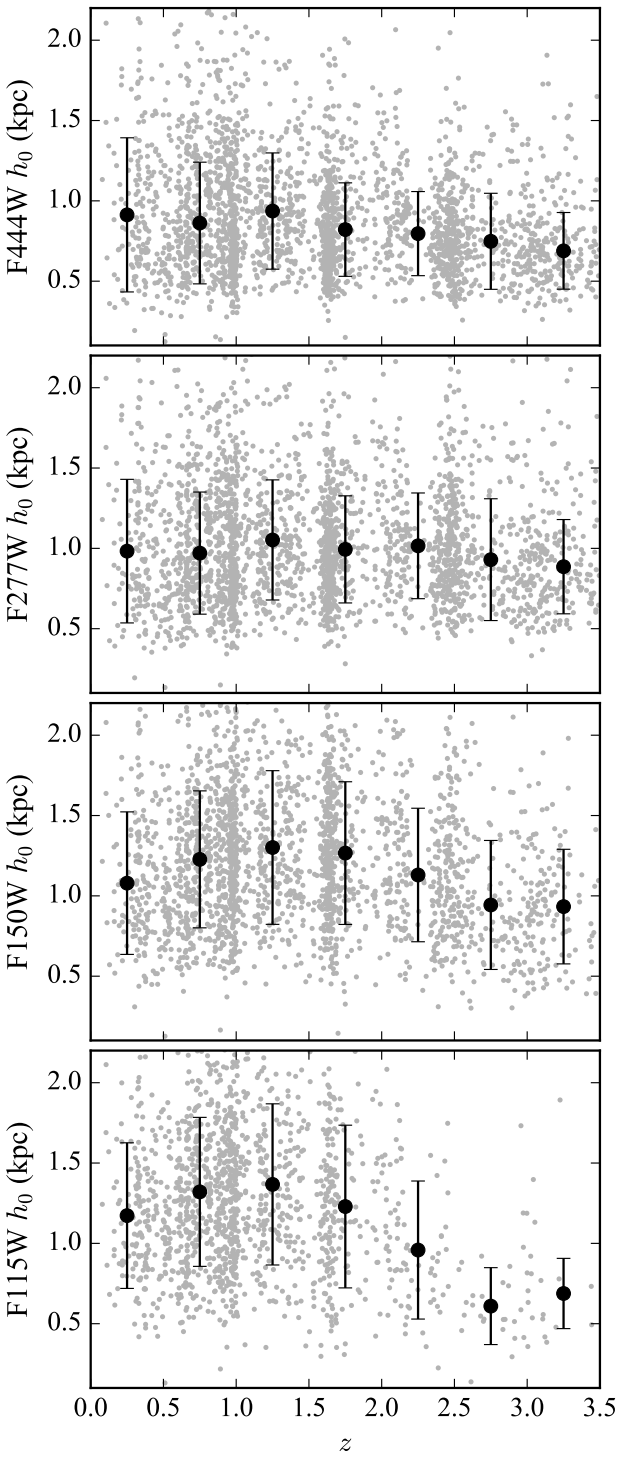


Figure 14. Scale heights derived from the F444W, F277W, F150W, and F115W bands as a function of redshift. The black points and associated error bars mark the median values and the scatter among individual measurements. At high redshifts, the number of galaxies included in the F115W and F150W measurements decreases because in these filters the galaxies become too faint in surface brightness to yield reliable measurements (see Section 3).

K. A. Hamilton-Campos et al. 2023; T. Tsukui et al. 2025). We compare our results with their measurements in Figure 18. Our measurements are plotted as black points, with the black curve showing the best-fit relation and the gray curves indicating its uncertainties. The fitting method is described later in Section 5.

At the solar circle of the Milky Way, M. Jurić et al. (2008) reported exponential scale heights of 0.9 and 0.3 kpc for the thick and thin disks, respectively, with uncertainties of about 20%. If these values are multiplied by the commonly used, although inappropriate, factor of 2 to convert to a sech^2 scale height, one obtains 1.8 and 0.6 kpc. Under this assumption, the Milky Way's thick disk would appear substantially thicker than the extragalactic values with a median of less than ~ 1 kpc, whereas its thin disk would be comparable to the global scale height at $z = 0$, suggesting that the Milky Way might be an unusually thick system. However, as shown in Equation (7), the appropriate conversion factor is 1.37 rather than 2. Using this conversion, the sech^2 scale heights of the Milky Way's thick and thin disks become 1.23 and 0.41 kpc, respectively, as indicated by the arrows in Figure 18. These values are a few hundred parsecs higher and lower, respectively, than the global scale heights at $z \approx 0$ in this work and in most of the previous studies. This is expected, because the global measurement represents a flux-weighted average of the thin and thick components.

Based on HST/Advanced Camera for Surveys (ACS) imaging, B. G. Elmegreen et al. (2017) measured the scale heights of 107 visually selected edge-on disks with an average stellar mass of $10^{9.5-10.5} M_\odot$ at $0.5 \leqslant z \leqslant 3.5$. They reported a mean scale height of 0.63 kpc with a scatter of 0.24 kpc in the F814W band, which probes rest-frame wavelengths of 0.20–0.54 μm and is therefore potentially affected by dust extinction. We regrouped their sample into three redshift bins and plot the results as light-gray diamonds in Figure 18. The data reveal an increase in scale height from $\sim$0.32 kpc at $z \approx 3$ to $\sim$1.03 kpc at $z \approx 1$.

K. A. Hamilton-Campos et al. (2023) measured the vertical scale heights of 491 disk galaxies with stellar masses of $M_* = 10^{9-11} M_\odot$ (the majority below $10^{10} M_\odot$) over the redshift range $0.4 \leqslant z \leqslant 2.5$ using HST/ACS F850LP, HST/WFC3 F125W, and F160W imaging. The images trace the rest-frame optical wavelength range (0.46–0.66 μm). The galaxies were selected with an axis ratio cut of $q < 0.4$, and the scale heights were derived from 1D vertical profile fitting. A constant correction factor of 0.78 was applied to account for the projection effect, yielding a median scale height of 0.74 kpc with a scatter of 0.35 kpc. The median values of the three redshift bins are shown as gray squares in Figure 18. They found little to no redshift evolution in the scale height. Nevertheless, the relatively shallow HST observations and their lower spatial resolution may introduce larger measurement uncertainties.

Using F115W imaging, which traces the rest-frame NIR at $z \sim 0$ but shifts to the rest-frame UV at $z > 3$, J. Lian & L. Luo (2024) measured sech^2 scale heights for 191 disk galaxies with a median stellar mass of $\sim 10^9 M_\odot$ over $0.2 \leqslant z \leqslant 5$, reporting a global median of 0.38 kpc with a scatter of 0.13 kpc. The red crosses in Figure 18 show their median values in each redshift bin (plotted only for $z < 3.5$). Their measured global scale height exhibits only mild evolution at $z > 1.5$ but decreases toward lower redshifts at $z < 1.5$, reaching a surprisingly small value of $\sim$0.18 kpc at $z = 0.35$. Perhaps due to uncertainties caused by small sample statistics, this value is roughly half of the global scale heights reported by T. Tsukui et al. (2025) for JWST galaxies of similar stellar masses at similar redshifts. J. Lian & L. Luo (2024) further argued that scale-height measurements are consistent across NIRCam filters from

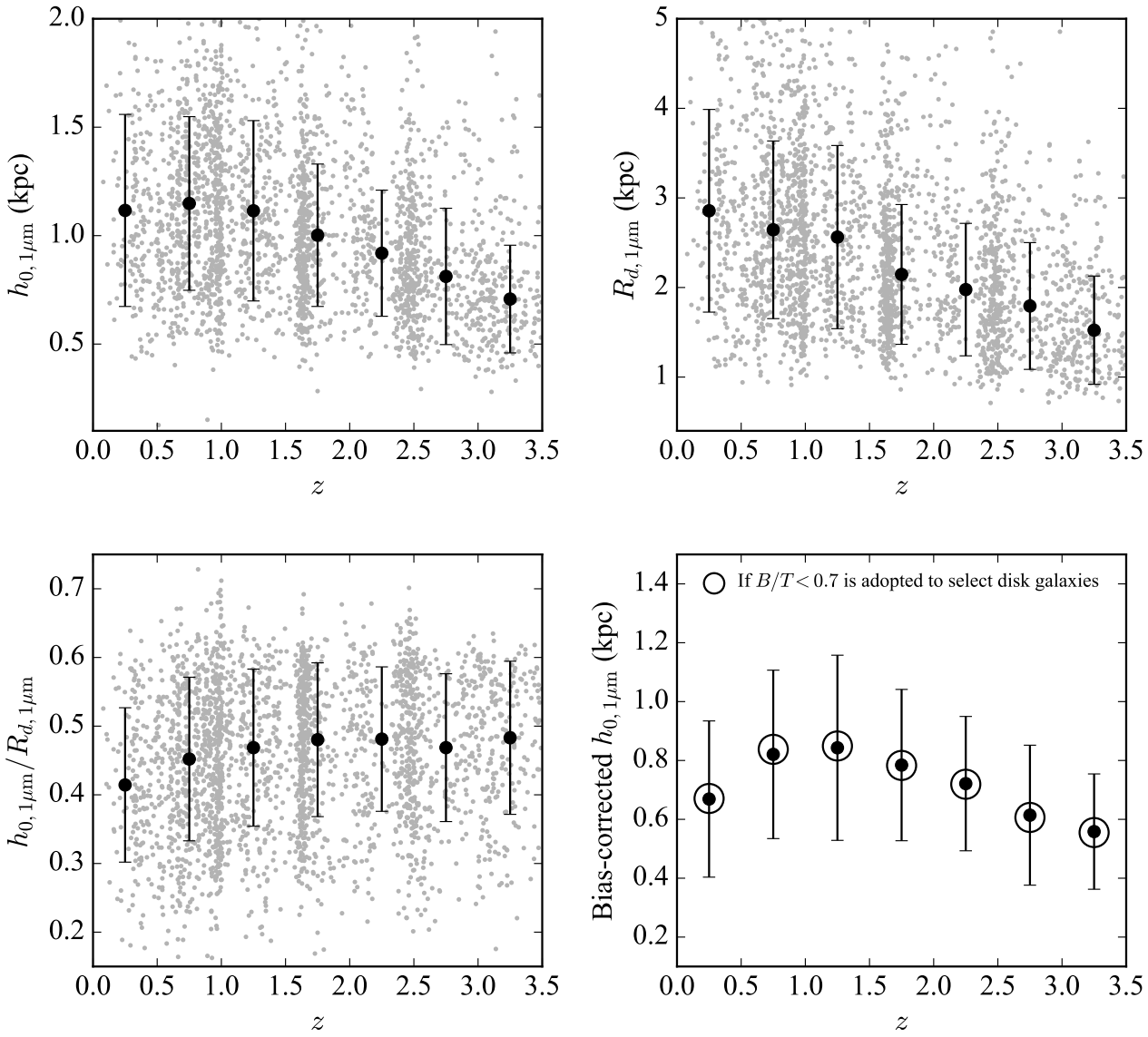


Figure 15. Measured at a fixed rest-frame wavelength of 1 μm, the scale height ($h_{0,1\ \mu m}$), scale length ($R_{d,1\ \mu m}$), relative scale height ($h_{0,1\ \mu m}/R_{d,1\ \mu m}$), and the bias-corrected $h_{0,1\ \mu m}$ (accounting for projection effects) are plotted as functions of redshift. Black points with error bars indicate the median values and the scatter among individual measurements across seven redshift bins.

F444W to F115W. However, as demonstrated in Section 4.1 and in D. Bizyaev & S. Mitronova (2009), measurements at shorter rest-frame wavelengths are systematically biased toward larger values due to dust extinction across the midplane.

T. Tsukui et al. (2025) studied 111 nearly edge-on disk galaxies at $z \leqslant 3$ using F277W, F356W, and F444W images, tracing rest-frame wavelengths of 1–2 μm, where the dust extinction effect is small. They identified 67 galaxies better fitted with a single-disk component and 44 galaxies requiring two disk components (thin and thick disks), with the earliest two-disk system found at $z = 1.96$. After correcting for mass dependence, the median scale heights at $M_* = 10^{9.2} M_\odot$ for single disks, thin disks, and thick disks are shown in Figure 18. Consistent with our results, the scale height of single-disk galaxies increases toward lower redshifts, with a median value of about 0.66 $\pm$ 0.07 kpc, slightly smaller than our measurements due to their lower stellar masses. In contrast, the thin and thick components of two-disk galaxies show nearly constant, or perhaps slightly decreasing, scale heights with decreasing redshift, with median values of 0.33 $\pm$ 0.04 and 0.94 $\pm$ 0.07 kpc, respectively.

5. Implications for Galaxy Evolution

It is important to emphasize that the redshift evolution of the disk scale height reported in this work refers to galaxies with a fixed stellar mass of approximately $2.5 \times 10^{10} M_\odot$. In reality, galaxies increase their stellar mass over cosmic time; therefore, our analysis does not trace the exact progenitor–descendant relationship of individual systems across redshifts. Nevertheless, the results offer valuable clues to the overall cosmic evolution of disk galaxies.

The formation mechanism of thin disks is relatively well understood. Thin disks are composed of stars born from a dynamically cold, rotationally supported gaseous disk. The newborn stars inherit the kinematics of the gas, exhibiting low ratios of velocity dispersion to rotation velocity, and

Table 3
Measured Scale Heights in Each Redshift Bin

z	$h_{0.1\ \mu m}$ (kpc)	$h_{0.1\ \mu m}/R_d$	f_{corr}	Bias-corrected $h_{0.1\ \mu m}$ (kpc)	Bias-corrected $h_{0.1\ \mu m}/R_d$	Bias-corrected $R_d/h_{0.1\ \mu m}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.25	1.12 ± 0.05 (0.44)	0.41 ± 0.02 (0.11)	0.599 ± 0.043	0.67 ± 0.06 (0.27)	0.25 ± 0.03 (0.07)	4.0 ± 0.4 (1.1)
0.75	1.15 ± 0.02 (0.40)	0.45 ± 0.01 (0.12)	0.715 ± 0.018	0.82 ± 0.03 (0.29)	0.32 ± 0.02 (0.09)	3.1 ± 0.2 (0.9)
1.25	1.11 ± 0.03 (0.42)	0.47 ± 0.01(0.11)	0.756 ± 0.025	0.84 ± 0.04 (0.31)	0.35 ± 0.02 (0.09)	2.8 ± 0.2 (0.8)
1.75	1.00 ± 0.02 (0.33)	0.48 ± 0.01(0.11)	0.783 ± 0.014	0.78 ± 0.03 (0.26)	0.38 ± 0.02 (0.09)	2.7 ± 0.1 (0.7)
2.25	0.92 ± 0.02 (0.29)	0.48 ± 0.01(0.10)	0.785 ± 0.016	0.72 ± 0.03 (0.23)	0.38 ± 0.02 (0.08)	2.7 ± 0.1 (0.6)
2.75	0.81 ± 0.03 (0.31)	0.47 ± 0.02(0.11)	0.756 ± 0.032	0.61 ± 0.04 (0.24)	0.35 ± 0.03 (0.08)	2.8 ± 0.2 (0.7)
3.25	0.71 ± 0.02 (0.25)	0.48 ± 0.02(0.11)	0.789 ± 0.026	0.56 ± 0.03 (0.20)	0.38 ± 0.03 (0.09)	2.6 ± 0.2 (0.7)

Note. Column (1): redshift. Columns (2)–(3): median, error of the median, and scatter among individual measurements for the measured $h_{0.1\ \mu m}$ and $h_{0.1\ \mu m}/R_d$, respectively; the scatter is shown in parentheses. Column (4): factor correcting for bias caused by the projection effect. Columns (5)–(7): median, error of the median, and scatter of the distribution for the bias-corrected $h_{0.1\ \mu m}$, $h_{0.1\ \mu m}/R_d$, and $R_d/h_{0.1\ \mu m}$, respectively.

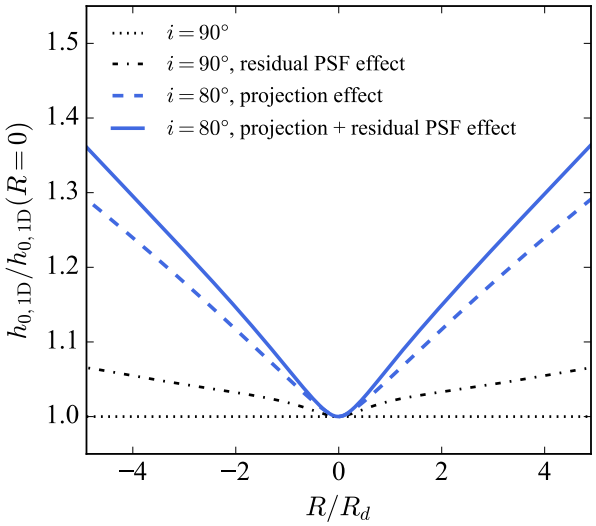


Figure 16. Artificial increases of scale height as a function of normalized radius (R/R_d). The increase is primarily driven by deviations from a perfectly edge-on orientation, with a minor contribution from residual PSF effects. The black dotted and dotted–dashed curves represent the results for a perfectly edge-on disk ($i = 90°$) without and with residual PSF effects, respectively, while the blue dashed and solid curves show the corresponding results for a disk inclined at $i = 80°$.

consequently form geometrically thin stellar structures (J. Forbes et al. 2012; S. Yu et al. 2023). However, open questions remain regarding when thin disks first emerged in the Universe and how they evolved over time.

In contrast, the dominant formation mechanism of thick disks remains debated despite decades of research. This uncertainty partly stems from observational limitations, particularly the difficulty in resolving the vertical structure of distant galaxies prior to the JWST era. Broadly speaking, proposed thick-disk formation mechanisms fall into two categories:

(1) Thick disks are born thick. In this picture, stars form with large scale heights of ∼1 kpc either at birth or shortly thereafter. This may occur through rapid stellar scattering driven by violent gravitational instabilities in gas-rich, clumpy high-redshift disks (F. Bournaud et al. 2009; L. Beraldo e Silva et al. 2020, 2021), combined with the dissolution of massive clumps (B. G. Elmegreen & D. A. Hunter 2006; B. G. Elmegreen et al. 2017). It may also arise from star formation within turbulent high-redshift gaseous disks (A. B. Romeo et al. 2010; A. B. Romeo & O. Agertz 2014), where giant molecular clouds (GMCs) have sizes comparable to clumps, typically ∼300 pc (B. G. Elmegreen et al. 2013; M. Dessauges-Zavadsky et al. 2017; A. Cava et al. 2018; M. Messa et al. 2022; U. Meštrić et al. 2022; A. Claeyssens et al. 2023). Together with the effect that the GMCs may not be coplanar, these conditions may directly produce stars with high velocity dispersion, naturally giving rise to thick-disk populations and providing a plausible explanation for the thick disks found in hydrodynamical simulations (S. Yu et al. 2021, 2023). Furthermore, the sloshing effect, where perturbation from nonaxisymmetric structures and the stochastic star formation in the disk induce bulk motion in the disk, increases the velocity dispersion of the stars (J. Bland-Hawthorn et al. 2024, 2025).

(2) Thick disks were born thin but later thickened. The potential mechanisms include dynamical heating by minor mergers (P. J. Quinn et al. 1993; Y. Qu et al. 2011), secular evolution by stellar structures (J. A. Sellwood 2014; L. A. Martinez-Medina et al. 2015), and a decrease in surface mass density due to galaxy size growth.

5.1. Thick Disks Begin as Intermediate-scale-height, Dynamically Hot, Dense Structures

The detection of relatively small disk scale heights at high redshifts was first reported by B. G. Elmegreen et al. (2017), who found that the median scale height of disks at $z = 3$ is ∼0.3 kpc, which is based on rest-frame ∼0.2 μm observations with HST. This result was recently reinforced by T. Tsukui et al. (2025), who analyzed JWST NIR rest-frame images of 111 edge-on disks. After correcting their measurements to a median stellar mass of $10^{9.2} M_\odot$, they obtained a median global disk scale height of 0.51 kpc at $z = 2.6$, compared to 0.77 kpc at $z = 0.26$. Consistently, using a JWST sample that is 15–30 times larger than in previous studies, we find a median scale height of 0.56 kpc at $z = 3.25$. Together, these findings suggest that high-redshift disks are neither thin nor thick, but instead occupy an intermediate-scale-height regime between typical thin disks (0.41 kpc for the Milky Way) and thick disks (1.23 kpc for the Milky Way), and experienced vertical growth over Gyr timescales. Consequently, the category (1) mechanisms may involve physical conditions in their galaxy models that do not fully match real observations, and may therefore require further improvement.

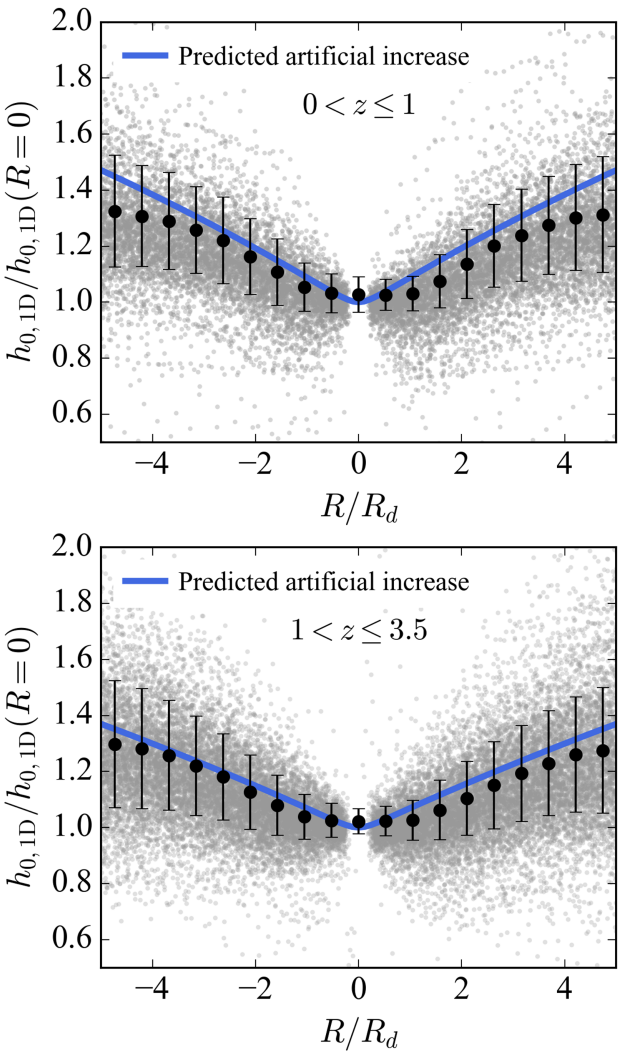


Figure 17. Normalized measured scale-height profiles for galaxies at $0 < z \leqslant 1$ (top) and $1 < z \leqslant 3.5$ (bottom). Black points with error bars show the median and scatter. The solid blue curves indicate the predicted artificial increase due to the median residual PSF effects and the median deviation from a perfectly edge-on inclination (predicted median $i = 80.^{\circ}9$ and $81.^{\circ}7$, respectively). Note that $5R_d \approx 3R_e$.

This, however, does not exclude the role of gravitational instability-driven scattering (F. Bournaud et al. 2009) or the sloshing effect (J. Bland-Hawthorn et al. 2025). On the contrary, such processes likely play a crucial role in producing dynamically hot stellar populations within thick disks. In the same vein, the "born thick" disk stars formed from large-size GMCs or clumps from bursty star formation seen in the simulations of S. Yu et al. (2023) also contribute to form hot stars. Our results are not necessarily inconsistent with these simulations, because the bridge connecting simulations and observations—namely, realistic mock observations derived from simulations—is still lacking.

Although disks at $z \approx 3.5$ are of geometrically intermediate thickness rather than thin or thick, they are not scaled-up versions of local cold thin disks. Instead, they more closely resemble thick disks, as they are dynamically hot and exhibit high stellar mass surface densities, as discussed below. In Figure 19, we show that the stellar surface mass density, $\Sigma_{*,\rm disk} = M_{*,\rm disk}/(2\pi R^2_{e,\rm disk,1\mu m})$, increases toward higher redshifts: the median value is $\Sigma_{*,\rm disk} = 10^{8.13} M_{\odot}\,{\rm kpc}^{-2}$ at $z = 0.25$ and $\Sigma_{*,\rm disk} = 10^{8.67} M_{\odot}\,{\rm kpc}^{-2}$ at $z = 3.25$. A best-fit relation to the median values in the seven redshift bins yields

$$\Sigma_{*,\rm disk}(z) = 10^{8.07\pm0.02}\exp((0.34 \pm 0.02)\, z)\, M_{\odot}\,{\rm kpc}^{-2}. \quad (10)$$

Assuming an isothermal disk (P. C. van der Kruit & L. Searle 1981), the vertical stellar velocity dispersion is related to h_0 and $\Sigma_{*,\rm disk}$ through (also see the derivation in Appendix A)

$$h_0 = \frac{\sigma^2_{*,\rm vert}}{\pi\, G\, \Sigma_{*,\rm disk}}, \quad (11)$$

or, equivalently,

$$\sigma_{*,\rm vert} = \sqrt{\pi\, G\, \Sigma_{*,\rm disk}\, h_0}. \quad (12)$$

Given $\Sigma_{*,\rm disk} = 10^{8.67} M_{\odot}\,{\rm kpc}^{-2}$, and a bias-corrected $h_{0,1\ \mu\rm m} = 0.56$ kpc at $z = 3.25$, we obtain $\sigma_{*,\rm vert} = 59\,{\rm km\,s^{-1}}$. This value is consistent with $\sigma_{*,\rm vert} = 50 \pm 5\,{\rm km\,s^{-1}}$ for stars in the solar neighborhood in the Milky Way's thick disk (J. Bland-Hawthorn & O. Gerhard 2016). If including the contribution of the gaseous component, assumed to account for 40% of the total mass (D. Narayanan et al. 2012), the inferred velocity dispersion increases to $\sim$78 km s^{-1}. Therefore, although the disks at $z = 3.25$ appear geometrically thin, their vertical velocity dispersions of 59–78 km s^{-1} indicate that they are dynamically hot. These hot stellar populations are likely produced by star scattering (F. Bournaud et al. 2009), the sloshing effect (J. Bland-Hawthorn et al. 2025), or are formed inherently hot (S. Yu et al. 2023). The strong gravitational potential of the dense midplane confines stars with high vertical velocity dispersions within a geometrically thin structure.

We may gain further insights from gas kinematics. Observations show that the velocity dispersion of the gaseous component, including the ionized, atomic, and molecular phases, increases systematically with redshift (K. Glazebrook 2013). By compiling and analyzing measurements from the literature, H. Übler et al. (2019) derived a redshift-dependent relation for the gas velocity dispersion: $\sigma_{\rm g}/{\rm km\,s^{-1}} = 10.9 + 11.0\, z$. Applying this relation yields $\sigma_{\rm g,vert} \sim 50\,{\rm km\,s^{-1}}$ at $z = 3.25$, consistent with the inferred $\sigma_{*,\rm vert}$ above. Such turbulent, gas-rich disks facilitate the formation of stars with high velocity dispersion (S. Yu et al. 2023). Consequently, even though the stellar disks at $z \approx 3.5$ are of geometrically intermediate thickness rather than thick, with a median $h_{0,1\ \mu\rm m} = 0.56$ kpc, they are dynamically hot. Furthermore, these $z \approx 3.5$ disks exhibit a single-component vertical structure and occupy a structural parameter space comparable to that of thick disks identified at lower redshifts, while a distinct thin component only begins to emerge at $z < 2$ (T. Tsukui et al. 2025). This may suggest that most galaxies initially form a thick disk, observed as a dense, dynamically hot, intermediate-thickness single component, followed by the subsequent development of a thin disk.

Because our sample is selected to have an approximately constant median stellar mass across redshifts, it does not trace the evolutionary paths of individual galaxies. However, T. Tsukui et al. (2025) similarly find relatively small scale heights ($\sim$0.51 kpc) in high-redshift disks at a lower median mass of $10^{9.2} M_{\odot}$. Combined, these findings point toward a picture in which present-day thick disks originate from dense, dynamically hot, intermediate-thickness disks formed in the turbulent, gas-rich conditions of early cosmic epochs.

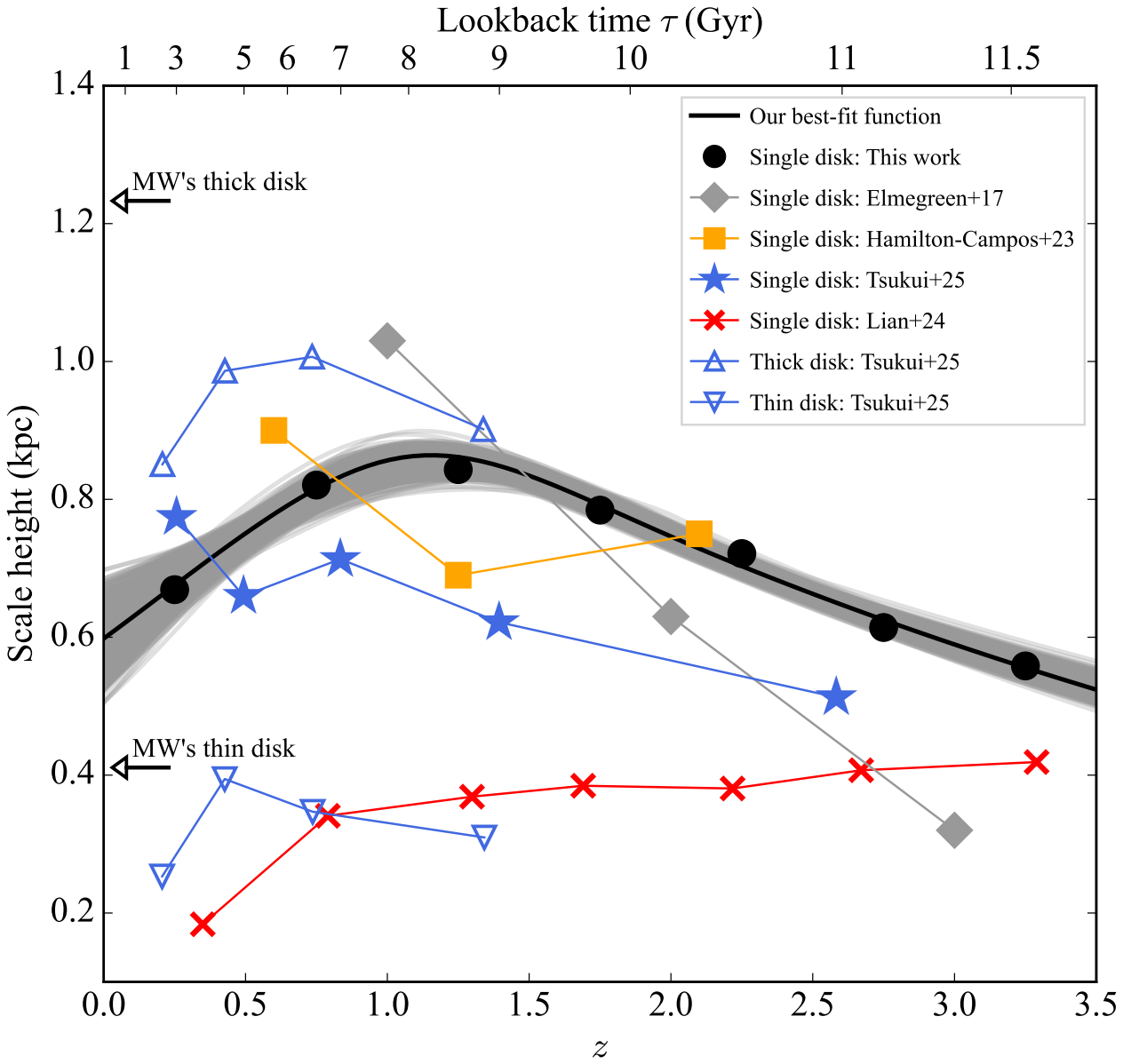


Figure 18. Redshift evolution of our bias-corrected scale heights and comparison with previous studies. Error bars are omitted for clarity. The global disk scale heights are from this work (points; median stellar mass $M_* \approx 10^{10.4} M_\odot$), B. G. Elmegreen et al. (2017; diamonds; $10^{10} M_\odot$), K. A. Hamilton-Campos et al. (2023; squares; $10^{9.3} M_\odot$), T. Tsukui et al. (2025; stars; $10^{9.2} M_\odot$), and J. Lian & L. Luo (2024; crosses; $10^{9} M_\odot$). Results for thick and thin disks are taken from T. Tsukui et al. (2025; open upward and downward triangles, respectively; $10^{9.2} M_\odot$). The black curve shows the best-fit function to our measurements, while the gray curves, indicating the fitting uncertainty, are the 1000 best-fit functions from 1000 resamplings of the data. The arrows indicate the sech^2 scale heights (1.23 and 0.41 kpc) of the Milky Way's thick and thin disks, derived by converting the exponential scale heights (0.9 and 0.3 kpc) reported by M. Jurić et al. (2008) using the conversion in Equation (7).

5.2. Minor Mergers and Secular Evolution Do Not Drive Present-day Thick Disks

Dynamical heating by minor mergers (P. J. Quinn et al. 1993; Y. Qu et al. 2011) has long been proposed as a possible mechanism for thick-disk formation, especially at early cosmic times when merger rates were significantly higher (K. Duncan et al. 2019; J. A. O'Leary et al. 2021). This process, however, is predicted to induce strong vertical flaring of stellar disks (F. Bournaud et al. 2009; B. P. Moster et al. 2010), i.e., an intrinsic increase of scale height with radius. Yet observations suggest otherwise. F. Bournaud et al. (2009) showed that high-redshift disks do not exhibit pronounced flaring, at least based on visual inspection of edge-on images. B. G. Elmegreen & D. A. Hunter (2006) did not find a signature of minor mergers in their sample. Moreover, D. V. Bizyaev et al. (2014) pointed out that even a slight deviation from a perfectly edge-on orientation can produce artificial flaring, which may mimic the expected intrinsic trend.

As shown in Figure 17, the observed radial gradient of the measured scale height in our sample is consistent with the amount of artificial flaring expected from the projection effect and the residual PSF effect (see Section 4.4 for a detailed calculation). This agreement implies that the apparent flaring in our sample is likely due to observational effects rather than intrinsic structural variations. Therefore, we find no evidence for intrinsic vertical flaring across our edge-on disk sample

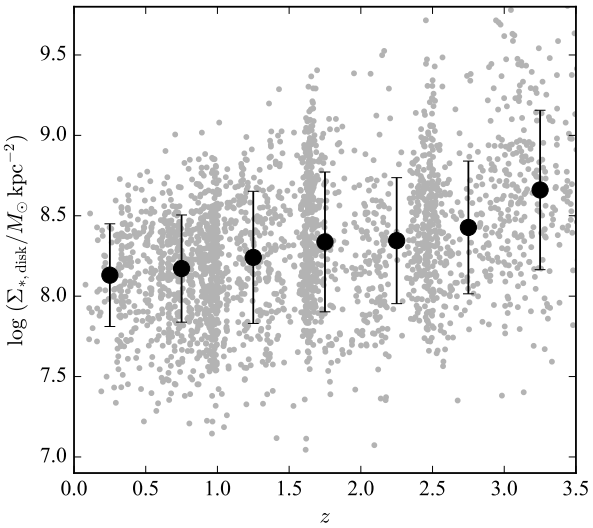


Figure 19. Stellar mass surface density of the disk component as a function of redshift. Gray dots represent individual galaxies, while black points with error bars show the median values and the scatter among individual measurements in each redshift bin.

over the redshift range $0 < z < 3.5$, effectively ruling out minor mergers as the dominant mechanism responsible for the formation of the thick-disk component.

Nevertheless, mergers likely played an important role in specific cases such as the Milky Way, where they contributed to disk heating and the formation of in situ halo stars with chemical abundances similar to those of thick-disk stars (V. Belokurov et al. 2020; A. Bonaca et al. 2020; C. Conroy et al. 2022).

Our results also do not support scenarios in which thick disks formed primarily through slow heating of a preexisting thin disk by bar buckling (J. A. Sellwood 2014) or spiral scattering (L. A. Martinez-Medina et al. 2015). These mechanisms are potentially important, as bars and spirals are common in nearby galaxies (S.-Y. Yu & L. C. Ho 2018, 2019, 2020; S.-Y. Yu et al. 2018) and have been shown to play a key role in driving galaxy secular evolution (S.-Y. Yu et al. 2021, 2022a, 2022b). If bar buckling or spiral scattering played a dominant role in disk thickening, one would expect disk thickness to increase toward lower redshifts at $z < 1$, when bars and spiral arms are well developed and dynamically active. In contrast, we observe a clear decrease in disk thickness over this redshift range (Figure 18). Furthermore, these secular mechanisms are also disfavored by the observed high-[α/Fe] stars in the thick disk of the Milky Way. If they were responsible for building the thick disk over the last 8 Gyr of the Universe, the gas reservoir would inevitably produce a substantial number of low-[α/Fe] stars that would be heated into the thick disk, which is contrary to observations. Consistently, hydrodynamical simulations also show no significant thick-disk formation at $z \lesssim 1$, after the end of the bursty star formation phase (S. Yu et al. 2023).

5.3. Scale Heights from z = 3 to 1

From $z = 3.25$ to $z = 1.25$, the disk scale height increases from 0.56 to 0.84 kpc, and R_d grows from 1.52 to 2.56 kpc, while $\Sigma_{*,\rm disk}$ declines from $10^{8.67}$ to $10^{8.24} M_\odot$ kpc^{-2}. The errors of these values are approximately 0.015 kpc, 0.05 kpc, and $10^{7.3} M_\odot$ kpc^{-2}, respectively. For comparison, $\Sigma_{*,\rm disk} = 10^{8.07} M_\odot$ kpc^{-2} at $z = 0$ in our sample. As discussed below, these trends reflect a combination of effects, including the decreasing of $\Sigma_{*,\rm disk}$ due to size growth and the generation of dynamically hot stars formed through bursty star formation, stellar scattering, and the sloshing effect.

We first examine the impact of size growth. It is well established that, at a fixed stellar mass, galaxy size increases toward lower redshifts (A. van der Wel et al. 2014; N. Allen et al. 2025; L. Yang et al. 2025), likely as a result of dark-matter halo growth and radial stellar migration. Because our galaxies have a nearly constant stellar mass of $10^{10.4} M_\odot$ across redshifts, size growth naturally leads to a lower $\Sigma_{*,\rm disk}$, which plays a key role in determining disk scale height. To test whether the decreasing of $\Sigma_{*,\rm disk}$ due to size growth alone can account for the observed increase in h_0, we assume that there are no exchanges of mass, energy, or angular momentum between the system and its environment to isolate the pure effect of size growth. The size growth is therefore assumed to be adiabatic. The scale height varies with $\Sigma_{*,\rm disk}$, as follows (see the Appendix A for the formula derivation):

$$h_0 \propto \Sigma_{*,\rm disk}^{-0.2}. \tag{13}$$

This relation predicts an increase in scale height by a factor of 1.16 from $z = 3.25$ to 1.75, smaller than the observed factor of 1.4. Hence, additional energy sources are required to raise the stellar velocity dispersion.

Both the observations and simulations suggest that thick disks were established rapidly during a turbulent phase rather than through a gradual buildup at early cosmic times ($z \gtrsim 1$) (P. E. Nissen et al. 2020; S. Yu et al. 2021, 2023; C. Conroy et al. 2022; M. Xiang & H.-W. Rix 2022; M. Xiang et al. 2025). During this phase, fueled by intense gas inflows along cosmological cold streams (D. Ceverino et al. 2010), high-redshift disks exhibited elevated star formation rates (S. Wuyts et al. 2011; J. S. Speagle et al. 2014; S. Gillman et al. 2024; T. Morishita et al. 2024) and clumpy, irregular morphologies (C. J. Conselice et al. 2008; A. Mortlock et al. 2013; L. Ferreira et al. 2023; J. S. Kartaltepe et al. 2023; A. L. Faisst et al. 2025), driven by violent gravitational instabilities (A. B. Romeo et al. 2010; A. B. Romeo & O. Agertz 2014). In such physical conditions, dynamically hot stars that form the thick disk may form, via star scattering (F. Bournaud et al. 2009) or the sloshing effect, whereby perturbation from nonaxisymmetric structures and stochastic star formation in the disk induce bulk motion in the disk with respect to the halo potential, increasing the velocity dispersion (J. Bland-Hawthorn et al. 2024, 2025).

If we consider an extreme case where bursty star formation and gravitational instabilities are strong enough to keep $\sigma_{*,\rm vert}$ unchanged during size growth (isothermal thickening), the scale height follows (see Equation (11))

$$h_0 \propto \Sigma_{*,\rm disk}^{-1}. \tag{14}$$

This predicts an increase in scale height by a factor of 2.1 from $z = 3.25$ to 1.75, larger than observed. Thus, the actual evolution lies between the adiabatic and isothermal limits. The observed rise in scale height at high redshifts therefore reflects the combined effects of size growth, bursty star formation, and violent disk instabilities. These disks are single-component systems that serve as progenitors of present-day thick disks, followed by the subsequent development of a thin disk.

After $z \approx 1$–2, the penetration of cold streams into massive halos becomes inefficient as halos develop shock-heated gas atmospheres, leading to a decline in gas accretion rates (A. Dekel & Y. Birnboim 2006; A. Dekel et al. 2009; D. Ceverino et al. 2010). Consequently, galaxies become less clumpy, reducing stellar scattering and the rate of thickening. Simulations have demonstrated that, at this epoch, some stars with low vertical velocity dispersions begin to form (S. Yu et al. 2023), facilitating the formation of old thin disks.

Observations have also confirmed the presence of old thin disks (S. Comerón et al. 2016; B. G. Elmegreen et al. 2017; T. Tsukui et al. 2025). S. Comerón et al. (2016) found old stellar populations in thin disks of S0 galaxies, and B. G. Elmegreen et al. (2017) inferred thin disks at $z \sim 2$ from vertical color gradients. Using JWST imaging, T. Tsukui et al. (2025) decomposed disks at $z > 1$ and identified the first thin disk at $z = 1.96$, finding that the thin-disk fraction increases toward lower redshifts. The detection of spiral arms and bars in high-redshift galaxies (L. Ferreira et al. 2023; J. S. Kartaltepe et al. 2023; Z. A. Le Conte et al. 2024, 2025; X. Liang et al. 2024; D. Xu & S.-Y. Yu 2024; T. Géron et al. 2025; Y. Guo et al. 2025; M. Huertas-Company et al. 2025; S.-Y. Yu et al. 2025) further supports the presence of dynamically cold, settled disks (K. Kraljic et al. 2012; D. M. Elmegreen & B. G. Elmegreen 2014). T. Tsukui et al. (2025) show that the thin disk forms in a downsizing way: thin disks appear to form earlier in more massive galaxies, and on average thin disks appear to dominate over the past 8–9 Gyr for the stellar mass bin $2.5 \times 10^{10} M_{\odot}$, corresponding to the bend in Figure 18.

5.4. Scale Heights from z = 1 to 0

At $z \approx 1$, the virialization of the inner circumgalactic medium begins to regulate cosmological cold streams, reducing gas inflow to the disk and marking the transition from the bursty phase to a steady star-forming phase that enables the formation of dynamically cold stellar disks (D. Ceverino et al. 2010; S. Yu et al. 2023). Although minor mergers may still occur during this later phase at $z \lesssim 1$, simulations suggest that they do not account for the majority of thick-disk stars (S. Yu et al. 2023), consistent with our result that no intrinsic flaring in the disks is observed. Instead, minor mergers mildly heat some thin-disk stars and contribute to the young-star tail of the thick-disk population (S. Yu et al. 2023). Fueled by cold gas accretion into the midplane and leftover gas from the previous phase, the thin disk gradually builds up through steady star formation.

Toward $z = 0$, galaxy sizes at a fixed stellar mass continue to increase (A. van der Wel et al. 2014), reducing $\Sigma_{*,\mathrm{disk}}$ and the midplane gravitational potential, thereby tending to increase the scale height of the disk. This effect competes with thin-disk formation, which drives a reduction in the global scale height through two mechanisms. First, the deepening of the midplane gravitational potential by cold gas accretion contracts the thick disk vertically (B. G. Elmegreen & D. M. Elmegreen 2006). These effects become more pronounced with the progressive growth of the thin disk over cosmic time, driving the overall decline in global thickness. Second, the superposition of thin and thick components introduces a geometric effect that reduces the measured global scale height in the single-disk sech^2 fitting of the combined vertical profile. During this period, Type Ia supernovae enrich the interstellar medium with Fe, lowering the $[\alpha/\mathrm{Fe}]$ ratios

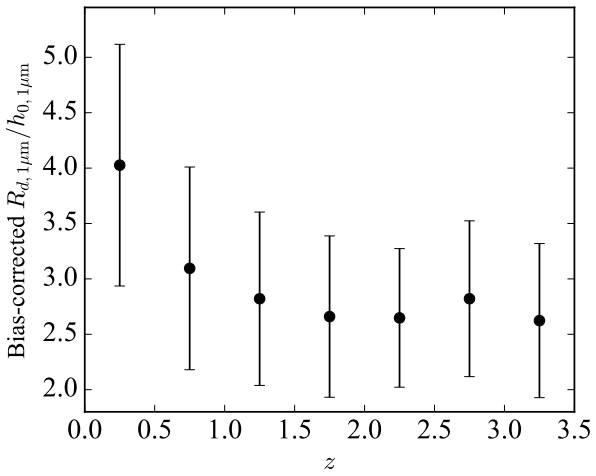


Figure 20. Bias-corrected ratio of disk scale length to scale height (the inverse of the relative scale height) as a function of redshift. Black points indicate the median values, and error bars represent the scatter among individual galaxies. The errors on the median are given in Table 3.

characteristic of thin-disk stars (J. Bland-Hawthorn & O. Gerhard 2016).

The ratio of scale length to scale height ($R_{d,1\ \mu\mathrm{m}}/h_{0,1\ \mu\mathrm{m}}$), i.e., the inverse of the relative scale height, serves as a useful indicator of the galaxy's dynamical state, roughly tracing the v/σ of the system (J. Kormendy 1982). The bias-corrected values are summarized in Table 3 and shown in Figure 20. The ratio remains approximately constant at 2.7 ± 0.2 for $z > 1.5$, but increases to 4.0 ± 0.4 at $z = 0.25$, indicating that galaxies become increasingly rotationally supported toward lower redshifts, which is consistent with the picture in which thin-disk growth dominates galaxy evolution in this epoch.

5.5. Simplified Model for Evolution and Its Fitting

Based on the discussion above, we construct a simplified model for the redshift evolution of scale height, aiming to gain further insight on galaxy disk evolution (Figure 21). Since the midplane mass density is the key factor determining disk thickness, the model is formulated as a function of $\Sigma_{*,\mathrm{disk}}$, which evolves with redshift and is connected back to the redshift according to the best-fit relation in Equation (10). We begin by assuming the dense, dynamically hot, intermediate-thickness disk observed at $z \approx 3.5$ represents the progenitors, or "seed galaxies," of present-day thick disks. An example of its vertical density profile is shown in the bottom-right inset panel of Figure 21. As previously demonstrated, the scale-height increase, due to the decrease in $\Sigma_{*,\mathrm{disk}}$, violent gravitational instabilities, and bursty star formation, between $z = 3.5$ and $z = 1$, is an intermediate process between the adiabatic process and isothermal regimes; we parameterize the relation between scale height and $\Sigma_{*,\mathrm{disk}}$ as

$$h_0 = D\ \Sigma_{*,\mathrm{disk}}^{k}, \tag{15}$$

where D is a normalization constant, and k lies between -0.2 and -1. Here, we adopt kiloparsecs and $M_{\odot}\ \mathrm{kpc}^{-2}$ as the units of scale height and stellar mass surface density, respectively. For simplicity, the increase in the mass of the gas component is absorbed into the factor D, given that both the stellar surface density $\Sigma_{*,\mathrm{disk}}$ and the gas fraction increase with redshift.

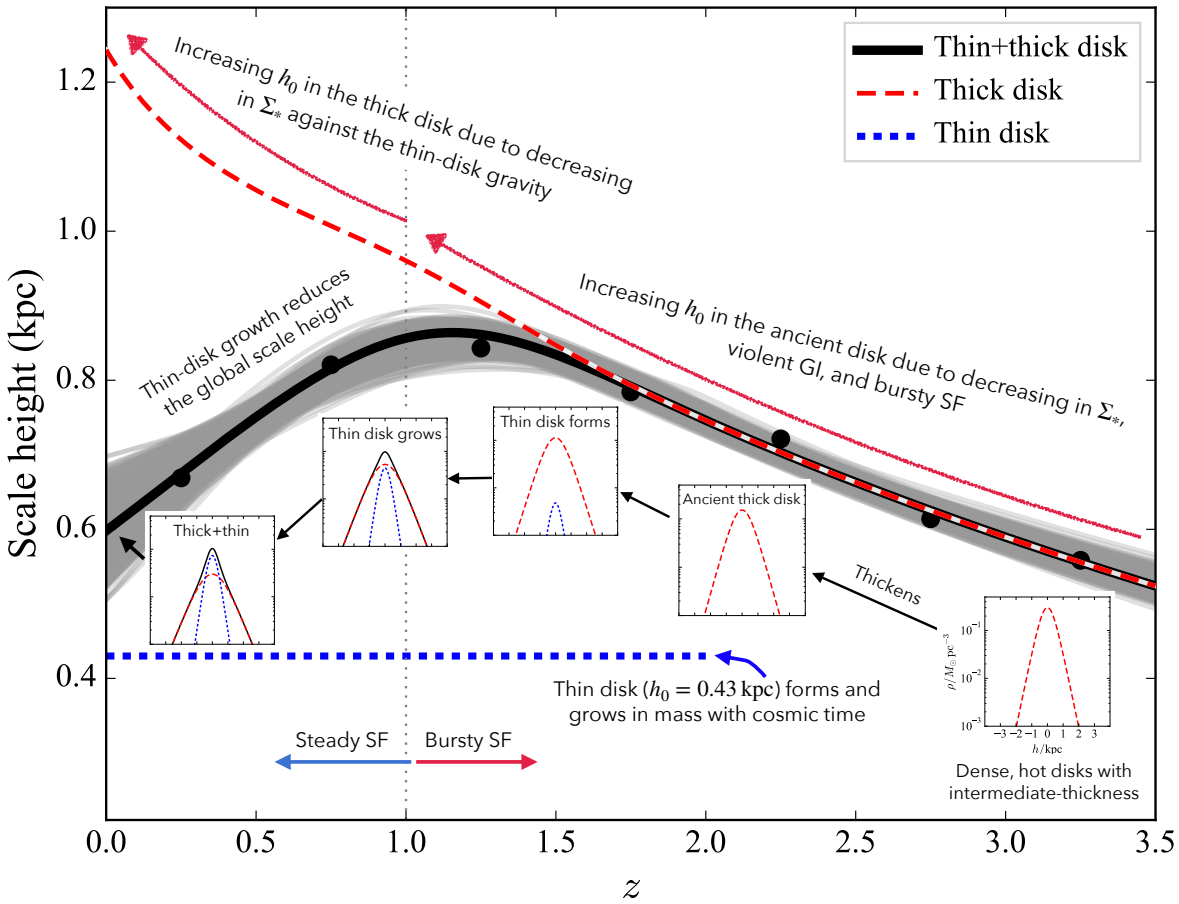


Figure 21. Simplified model for the redshift evolution of global disk scale height, averaging the thin and thick disk, at a fixed stellar mass of $M_* = 2.5 \times 10^{10} M_\odot$. The solid curve shows the best-fit relation, and the gray curves (shaded region) represent 1000 realizations from resampled datasets. The red dashed curve indicates the best-fit evolution of the thick-disk scale height, while the thin-disk scale height is fixed at 0.43 kpc. The inset panels indicate how the vertical light profiles of the single-disk-component galaxies—and subsequently thin, thick, and combined disks—evolve with redshift.

For $z \lesssim 1$, the formation of thick disks is expected to be significantly reduced. We therefore assume pure adiabatic thickening for the thick disk at $z < 1$, and use a step function ζ to connect $z > 1$ and $z < 1$ formulae (before the addition of the thin disk):

$$h^0_{\rm thick}(z) = D[\zeta(z)\,\Sigma^{k+0.2}_{*,\rm disk}(z=1)\,\Sigma^{-0.2}_{*,\rm disk}(z) + (1-\zeta(z))\,\Sigma^{k}_{*,\rm disk}(z)]. \quad (16)$$

where $\Sigma^{k+0.2}_{*,\rm disk}(z=1)$ is to ensure the continuity at $z = 1$. Motivated by the detection of a thin disk as early as $z \sim 1.96$ and the increasing number fraction of thin disks toward lower redshifts reported in T. Tsukui et al. (2025), we adopt the following functional form for the mass fraction of a thin disk:

$$f_{\rm thin} = \begin{cases} f_0\,(2-z)^t, & {\rm for}\, z \leqslant 2 \\ 0, & {\rm for}\, z > 2 \end{cases} \quad (17)$$

where f_0 and t characterize the amplitude and growth rate. We assume that the mass-to-light ratio for the thick and thin disks is 1.2/1 (S. Comerón et al. 2011). T. Tsukui et al. (2025) find that scale heights for thin disks are approximately unchanged with z, with a median value of ∼0.33 kpc. We therefore adopt a constant scale height for the thin-disk model:

$$h_{\rm thin}(z) = 0.43\ {\rm kpc}, \quad (18)$$

scaled upward by a factor of 1.3, the difference between their and our global scale heights, to account for stellar mass differences between their sample and ours. This value is almost the same with the sech^2 scale height of the Milky Way's thin disk (0.41 kpc). An example vertical light profile showing a thick disk overlapped with a newly formed thin disk is displayed in the inset at $z = 1.5$, while a later stage with a more developed thin disk is illustrated at $z = 1$ in Figure 21.

Because our sample has a constant stellar mass across redshifts, the growth of the thin disk must be accompanied by a reduction of the stellar mass in the thick component. To model this redistribution, we assume that the vertical velocity dispersion of thick-disk stars remains unchanged, with a modest increase caused by adiabatic contraction during thin-disk buildup. This is reasonable because the dominant heating mechanisms responsible for thick-disk formation operate primarily at $z \gtrsim 1$ and are largely ineffective at later times. The gradual buildup of the thin disk, with surface density $f_{\rm thin}\Sigma_{*,\rm disk}(z)$, is assumed to proceed adiabatically, contracting the vertical structure and slightly increasing the velocity dispersion of stars of the preexisting thick disk (B. G. Elmegreen & D. A. Hunter 2006; B. G. Elmegreen et al. 2017). We assume the approximation that the thin and thick components each maintain independent pressure equilibria, while their gravitational potentials are coupled (B. G. Elmegreen & D. A. Hunter 2006; C. J. Jog 2007). The connection between

the scale heights of the thin and thick components in a thick–thin system can then be approximated as (see the Appendix B and J. Forbes et al. 2012):

$$h_{\rm thick} = \frac{\sigma^2_{\rm thick}}{\pi G\left((1-f_{\rm thin})\Sigma_{*,\rm disk} + f_{\rm thin}\,\Sigma_{*,\rm disk}\frac{h_{\rm thick}}{h_{\rm thin}}\right)}, \quad (19)$$

where $\sigma_{\rm thick}$ is linked to $h^0_{\rm thick}$ and $\Sigma_{*,\rm disk}$ through the adiabatic invariant of the thick disk (Equation A10). The quantity $h_{\rm thick}(z)$ is then solved numerically.

The corresponding vertical light profiles are given by

$$I_{\rm thick}(h,\,z) = \frac{(1-f_{\rm thin})\,\Sigma_{*,\rm disk}(z)}{2\,h_{\rm thick}(z)} \times {\rm sech}^2\left(\frac{h}{h_{\rm thick}(z)}\right) \quad (20)$$

for the thick disk and

$$I_{\rm thin}(h,\,z) = \frac{f_{\rm thin}\,\Sigma_{*,\rm disk}(z)}{2\,h_{\rm thin}(z)}\,{\rm sech}^2\left(\frac{h}{h_{\rm thin}(z)}\right) \quad (21)$$

for the thin disk. An example of the combined thin+thick profile at $z = 0$, showing a dominant thin disk embedded within a thicker component, is presented in Figure 21. To obtain the disk scale height averaging the thin and thick disk, we fit a single sech^2 function to the combined profile, mimicking our observational procedure applied to edge-on disks in COSMOS-Web:

$$h_0(z) = \Gamma(I_{\rm thin}(h,\,z) + I_{\rm thick}(h,\,z)), \quad (22)$$

where Γ denotes the fitting operation.

We fit Equation (22) to the bias-corrected $h_{0,1\ \mu{\rm m}}(z)$ data using four free parameters: D, k, f_0, and t. The best-fit function is shown as the black curve in Figures 18 and 21. To estimate uncertainties, we perform 1000 bootstrap resamplings with replacement and refit the model each time; the resulting best-fit curves are shown as gray lines in Figures 18 and 21. The resulting parameters yield $D = 10^{6.0} \pm 10^{4.9}$, $k = -0.73 \pm 0.02$, $f_0 = 0.11 \pm 0.03$, and $t = 2.52 \pm 0.52$. The best-fit evolutionary curve for the thick-disk component is plotted as a red dashed line in Figure 21. The continued increase of the thick-disk scale height below $z = 1$ is primarily due to the decline in $\Sigma_{*,\rm disk}$ caused by size growth and mass reduction. The redshift corresponding to the peak of the best-fit curve is $z_{\rm peak} = 1.12 \pm 0.04$.

We assess the robustness of the fit by comparing the Bayesian information criterion (BIC) between Equation (22) and a simple linear model. The BIC for Equation (22) is lower by 20, indicating a strong statistical preference for our simplified model and confirming the presence of a peak in the evolution of the disk scale height.

The best-fit thin-disk mass fraction, $f_{\rm thin}$, is shown in Figure 22. It increases from 0 at $z = 2$ to 0.10 ± 0.03 at $z = 1$, and 0.60 ± 0.10 at $z = 0$. Our inferred $f_{\rm thin}$ values are systematically higher than those reported by T. Tsukui et al. (2025) for a sample with a median mass of $10^{9.2} M_\odot$, likely because galaxies with higher stellar masses tend to host a larger fraction of thin disks (S. Comerón et al. 2018, and see Figure 10 of T. Tsukui et al. 2025). Another plausible explanation is that the decomposition of thin- and thick-disk components may be systematically incomplete relative to the idealized model expectations. Detecting two different components requires sufficient contrast between their vertical profiles, so galaxies with comparable thin- and thick-disk masses are more likely identified as having a two-component structure. In contrast, when either component dominates the flux, the secondary component becomes difficult to detect observationally, and the galaxy is more likely classified as a single-component system. This may lead to the absence of very low thin-disk fractions in massive galaxies reported by S. Comerón et al. (2018) and T. Tsukui et al. (2025). Nevertheless, the inferred fraction at $z = 0$ agrees well with the measurements by S. Comerón et al. (2018), which gives $f_{\rm thin} = 0.66 \pm 0.09$ for nearby galaxies with $M_* \sim$; 10^{10}–$10^{10.8}\,M_\odot$. This agreement suggests that, despite its simplicity, our model captures the essential physical processes governing the coevolution of thin and thick disks.

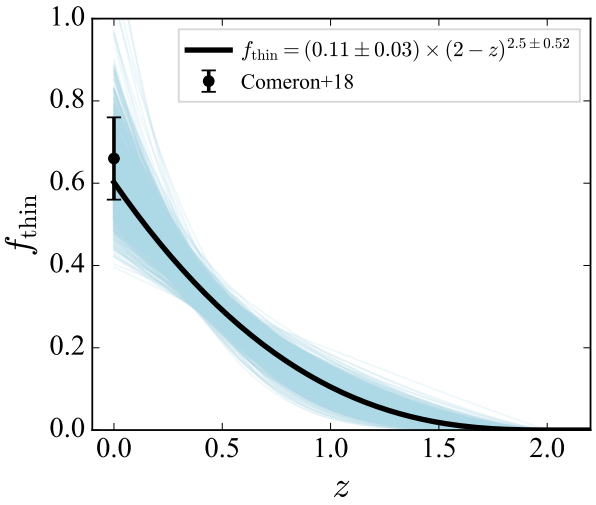


Figure 22. Best-fit mass fraction of a thin disk as a function of redshift. The solid black curve marks the best-fit function, while the shadowed light-blue curves, indicating the fitting uncertainty, are the 1000 fits of the 1000 resamplings of the data. The point and its error bar mark the results from S. Comerón et al. (2018) for nearby galaxies with stellar masses of 10^{10}–$10^{10.8}\,M_\odot$ in their sample.

In summary, the observed redshift evolution of disk scale height from COSMOS-Web can be explained by the interplay between thick- and thin-disk formation and evolution. The rise in h_0 from $z = 3.5$ to 1 results from size-driven decreasing of $\Sigma_{*,\rm disk}$, violent gravitational scattering, and bursty star formation. The subsequent decline from $z = 1$ to 0 reflects geometric overlap with thin disks, a continuing $\Sigma_{*,\rm disk}$ decrease, and contraction driven by thin-disk growth.

6. Summary and Conclusion

We conducted a comprehensive analysis of global disk scale heights for a sample of 2631 clean, nearly edge-on disk galaxies at $0 < z < 3.5$, observed by JWST/NIRCam in the COSMOS-Web survey with the F115W, F150W, F277W, and F444W filters. The sample has a median stellar mass of $2.5 \times 10^{10} M_\odot$, and a disk axis ratio cut of $q_d < 0.4$ was applied to ensure edge-on orientations. We measured the vertical scale height, h_0, by fitting a 2D model consisting of a bulge and an edge-on disk described by a single sech^2 vertical profile. The disk size is characterized by the radial exponential scale length R_d. The resulting h_0 represents an average scale height that combines the contributions of both thin and thick disks, where

such components coexist. For comparison, exponential vertical profiles were also tested. In addition, 1D vertical light profiles were fitted with an LSF, constructed using a PSF, to derive scale heights as a function of radius. The measurements of scale heights and scale lengths, as well as the correction curves used to account for projection effects, can be accessed at Zenodo (doi:10.5281/zenodo.18059559; S.-Y. Yu 2026). Our main findings are summarized as follows:

1. The h_0 derived from sech^2 vertical profiles are 1.37 times as large as those obtained from exponential profiles, with negligible scatter.
2. The h_0 measured at shorter rest-frame wavelengths are systematically overestimated relative to those measured at longer wavelengths, primarily due to dust extinction across the midplane. This effect is consistent with D. Bizyaev & S. Mitronova (2009). We adopt a fixed rest-frame 1 μm for our measurements to mitigate the dust extinction and spatial variation in the stellar population.
3. Deviations from a perfectly edge-on inclination systematically bias h_0 toward larger values, with this projection effect becoming progressively more severe in relatively thinner disks (i.e., those with smaller h_0/R_d). We developed a new statistical correction method for this inclination bias, which is more accurate than the constant correction factor proposed by K. A. Hamilton-Campos et al. (2023).
4. The observed radial variation of scale heights across our sample is consistent with artificial flaring primarily caused by projection effects, with a minor contribution from imperfect PSF treatment in the 1D fitting. This disfavors minor mergers as the dominant mechanism for the disk thickening and hence thick-disk formation, which are expected to produce genuine flaring.
5. After correcting for the biases caused by variations in the observed rest-frame wavelength and by the projection effect, we find an unambiguous trend in scale height as a function of redshift: it increases from 0.56 ± 0.03 kpc at $z = 3.25$ to 0.84 ± 0.04 kpc at $z = 1.25$, and then decreases to 0.67 ± 0.06 kpc at $z = 0.25$. The scatters among individual measurements are 0.2–0.3 kpc. The disks at $z \sim 3.5$ were of geometrically intermediate thickness, yet dynamically hot and highly dense, likely representing the progenitors of present-day thick disks.
6. The bias-corrected ratio of scale length R_d to h_0 (R_d/h_0) remains approximately constant at 2.7 ± 0.2 at $z > 1.5$, while it increases to 4.0 ± 0.4 at $z = 0.25$.
7. Using a simplified evolutionary model, we suggest that the early dense, dynamically hot, intermediate-thickness, single-disk galaxies at $z \approx 3.5$ thickened as a result of decreasing surface mass density (as galaxies grew in size), violent gravitational instabilities, and bursty star formation, leading to increasing scale height toward lower redshifts at $z > 1$. Below $z \sim 1$, thin-disk growth became dominant, producing a vertically more compact structure with smaller scale heights toward $z = 0$.
8. The model fitting further indicates that the thin-disk mass fraction $f_{\rm thin}$ increases from $f_{\rm thin} = 0$ at $z \approx 3.5$ to $f_{\rm thin} = 0.10 \pm 0.03$ at $z = 1$ and to $f_{\rm thin} = 0.60 \pm 0.10$ at $z = 0$. The best-fit value at $z = 0$ agrees well with the NIR observational constraints for nearby galaxies with a comparable stellar mass reported by S. Comerón et al. (2018).

A detailed investigation of disk thickness as a function of galaxy properties, such as stellar mass, will be pursued in future work. Future studies focusing on large-sample thick-thin disk decompositions, as well as direct comparisons between realistic mock images from high-resolution hydrodynamical simulations and observations, will be essential for disentangling the physical processes that govern the evolution of disk thickness and the transition from high-redshift dense hot disks with an intermediate scale height to present-day thin-thick disk systems.

Acknowledgments

We thank the referee for the insightful comments and suggestions. L.C.H. was supported by the China Manned Space Program (CMS-CSST-2025-A09) and the National Science Foundation of China (12233001). T.T. is supported by the JSPS Grant-in-Aid for Research Activity Start-up (25K23392) and the JSPS Core-to-Core Program (JPJSCCA20210003). S.-Y.U. acknowledges support from the UTokyo Global Activity Support Program for Young Researchers. We thank Jing Wang, Jianhui Lian, Raymond C. Simons, Qikang Feng, and Sijia Li for their discussions. Kavli IPMU is supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan. The JWST data presented in this article were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via doi:10.17909/ph8h-qf05.

Author Contributions

S.-Y.U. and L.C.H. developed the initial idea of the project. S.-Y.U. conducted the data reduction, developed the methods, and performed the analysis. S.-Y.U. wrote the manuscript, with revisions from other contributors.

Facility: JWST (NIRCam)

Software: astropy (Astropy Collaboration et al. 2013, 2018, 2022), IMFIT (P. Erwin 2015), SEP (E. Bertin & S. Arnouts 1996; K. Barbary 2016).

Appendix A
Derivation of Key Equations for a Self-gravitating Single-disk Model

In the main text, we use h to denote the vertical distance instead of the commonly used z, to avoid confusion with the symbol for the redshift. In Appendices A and B, however, we revert to z for vertical distance, as it is the convention most widely adopted in the literature.

For a stellar disk in vertical equilibrium, the balance between the vertical component of the gravitational force and the dynamical pressure ($P = \rho\,\sigma^2$) associated with the random stellar motions can be expressed as

$$\frac{1}{\rho}\frac{d(\rho\,\sigma_z^2)}{dz} = -\frac{d\Phi}{dz}, \tag{A1}$$

where $\rho(z)$ is the mass density, σ_z is the vertical velocity dispersion, and $\Phi(z)$ is the gravitational potential. This is the vertical component of the steady-state Jeans equation for a collisionless stellar system. If the system is isothermal in the

sense that σ_z is constant with height, the equation reduces to

$$\sigma_z^2 \frac{d\rho}{dz} = -\rho \frac{d\Phi}{dz}. \tag{A2}$$

The gravitational potential obeys the Poisson equation

$$\frac{d^2\Phi}{dz^2} = 4\pi G\rho. \tag{A3}$$

Combining these two equations and eliminating Φ gives

$$\frac{d^2}{dz^2}\ln\rho = -\frac{4\pi G}{\sigma_z^2}\rho, \tag{A4}$$

whose analytic solution is

$$\rho(z) = \rho_0 \,\mathrm{sech}^2\left(\frac{z}{z_0}\right), \tag{A5}$$

where ρ_0 is the midplane density, and z_0 is the vertical scale height. Substituting this form into Equation (A4) yields

$$z_0^2 = \frac{\sigma_z^2}{2\pi G\rho_0}. \tag{A6}$$

The total surface density of the disk is

$$\Sigma = \int_{-\infty}^{+\infty} \rho(z)\, dz = 2\rho_0 z_0. \tag{A7}$$

Substituting $\rho_0 = \Sigma/(2z_0)$ into Equation (A6) gives

$$z_0 = \frac{\sigma_z^2}{\pi G\Sigma}. \tag{A8}$$

If the disk evolves slowly in an adiabatic manner, the vertical pressure and density satisfy

$$P = K\rho^\gamma, \tag{A9}$$

where K is the adiabatic constant, and $\gamma = 5/3$ for a collisionless stellar system. With $P_0 = \rho_0\sigma_z^2$ and $\rho_0 = \Sigma/(2z_0)$, we obtain

$$\sigma_z^2 = K\left(\frac{\Sigma}{2z_0}\right)^{\gamma-1}. \tag{A10}$$

Combining this with Equation (A8) yields

$$\Sigma^{\,2-\gamma} {z_0}^{\gamma} = \text{const.} \tag{A11}$$

For $\gamma = 5/3$, the scale height and velocity dispersion vary as

$$z_0 \propto \Sigma^{-1/5}, \quad \sigma_z \propto \Sigma^{2/5}. \tag{A12}$$

Thus, during a slow adiabatic expansion, the disk becomes slightly thicker and dynamically cooler as its surface density decreases.

Appendix B
Derivation of Key Equations for a Self-gravitating Double-disk Model

To derive an approximate formula that connects the scale heights of the two disk components, we assume that the thin and thick components each maintain independent pressure equilibria, while their gravitational potentials are coupled. We emphasize that, within this approximation, the two resulting sech^2 profiles do not constitute a strictly self-consistent solution to the Poisson equation together with vertical hydrostatic balance. Even so, this approximation and the adoption of two sech^2 profiles offer a practical and decent representation of the observed scale heights. Each disk follows a sech^2 density profile:

$$\rho_i(z) = \rho_{0i}\,\mathrm{sech}^2\left(\frac{z}{z_{0i}}\right), \quad i = 1, 2, \tag{B1}$$

with corresponding surface densities

$$\Sigma_i = 2\rho_{0i}z_{0i}. \tag{B2}$$

The total gravitational potential satisfies the Poisson equation:

$$\frac{d^2\Phi}{dz^2} = 4\pi G[\rho_1(z) + \rho_2(z)]. \tag{B3}$$

Each disk is assumed to be isothermal in the vertical direction, so that vertical hydrostatic equilibrium for each component gives

$$\sigma_i^2 \frac{d\ln\rho_i}{dz} = -\frac{d\Phi}{dz}, \tag{B4}$$

where σ_i is the vertical velocity dispersion of disk i.

Taking the logarithmic derivative of the sech^2 profile, we obtain

$$\frac{d\ln\rho_i}{dz} = -\frac{2}{z_{0i}}\tanh\left(\frac{z}{z_{0i}}\right), \tag{B5}$$

so that hydrostatic equilibrium becomes

$$\frac{d\Phi}{dz} = \frac{2\sigma_i^2}{z_{0i}}\tanh\left(\frac{z}{z_{0i}}\right). \tag{B6}$$

Differentiating once more with respect to z gives

$$\frac{d^2\Phi}{dz^2} = \frac{2\sigma_i^2}{z_{0i}^2}\,\mathrm{sech}^2\left(\frac{z}{z_{0i}}\right). \tag{B7}$$

Evaluating at the midplane $z = 0$ where $\mathrm{sech}^2(0) = 1$ and comparing with the Poisson equation yields

$$\frac{2\sigma_1^2}{z_{01}^2} = \frac{2\sigma_2^2}{z_{02}^2} = 4\pi G(\rho_{01} + \rho_{02}). \tag{B8}$$

Substituting $\rho_{0i} = \Sigma_i/(2z_{0i})$ gives the coupled algebraic equations for the scale heights:

$$\begin{aligned}\sigma_1^2 &= \pi G\left(\Sigma_1 z_{01} + \Sigma_2\frac{z_{01}^2}{z_{02}}\right),\\ \sigma_2^2 &= \pi G\left(\Sigma_2 z_{02} + \Sigma_1\frac{z_{02}^2}{z_{01}}\right).\end{aligned} \tag{B9}$$

Dividing the two equations provides a simple relation between the scale heights:

$$\frac{z_{01}}{z_{02}} = \frac{\sigma_1}{\sigma_2}. \tag{B10}$$

The explicit expressions for z_{01} and z_{02} are then

$$z_{01} = \frac{\sigma_1^2}{\pi G\left(\Sigma_1 + \Sigma_2 \frac{z_{01}}{z_{02}}\right)},$$
$$z_{02} = \frac{\sigma_2^2}{\pi G\left(\Sigma_2 + \Sigma_1 \frac{z_{02}}{z_{01}}\right)}. \quad \text{(B11)}$$

ORCID iDs

Si-Yue Yu https://orcid.org/0000-0002-3462-4175
Luis C. Ho https://orcid.org/0000-0001-6947-5846
Takafumi Tsukui https://orcid.org/0000-0002-1499-6377
John D. Silverman https://orcid.org/0000-0002-0000-6977
Marc Huertas-Company https://orcid.org/0000-0002-1416-8483
Anton M. Koekemoer https://orcid.org/0000-0002-6610-2048
Maximilien Franco https://orcid.org/0000-0002-3560-8599
Richard Massey https://orcid.org/0000-0002-6085-3780
Lilan Yang https://orcid.org/0000-0002-8434-880X
Rafael C. Arango-Toro https://orcid.org/0000-0002-0569-5222
Andreas L. Faisst https://orcid.org/0000-0002-9382-9832
Ghassem Gozaliasl https://orcid.org/0000-0002-0236-919X
Kartik Sheth https://orcid.org/0000-0002-5496-4118
Jeyhan S. Kartaltepe https://orcid.org/0000-0001-9187-3605
Can Xu https://orcid.org/0000-0002-8437-6659
Aryana Haghjoo https://orcid.org/0009-0006-3071-7143
Xuheng Ding https://orcid.org/0000-0001-8917-2148
Zhaoxuan Liu https://orcid.org/0000-0002-9252-114X
Jacqueline E. McCleary https://orcid.org/0000-0002-9883-7460

References

Allen, N., Oesch, P. A., Toft, S., et al. 2025, A&A, 698, A30
Arnouts, S., Moscardini, L., Vanzella, E., et al. 2002, MNRAS, 329, 355
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Barbary, K. 2016, JOSS, 1, 58
Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, MNRAS, 494, 3880
Bensby, T., Feltzing, S., & Oey, M. S. 2014, A&A, 562, A71
Beraldo e Silva, L., Debattista, V. P., Khachaturyants, T., & Nidever, D. 2020, MNRAS, 492, 4716
Beraldo e Silva, L., Debattista, V. P., Nidever, D., Amarante, J. A. S., & Garver, B. 2021, MNRAS, 502, 260
Bertin, E. 2011, ASPC, 442, 435
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, E., Schefer, M., Apostolakos, N., et al. 2020, ASPC, 527, 461
Bird, J. C., Loebman, S. R., Weinberg, D. H., et al. 2021, MNRAS, 503, 1815
Bizyaev, D., & Mitronova, S. 2009, ApJ, 702, 1567
Bizyaev, D. V., Kautsch, S. J., Mosenkov, A. V., et al. 2014, ApJ, 787, 24
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Bland-Hawthorn, J., Tepper-Garcia, T., Agertz, O., & Federrath, C. 2024, ApJ, 968, 86
Bland-Hawthorn, J., Tepper-Garcia, T., Agertz, O., et al. 2025, ApJ, 994, 22
Bonaca, A., Conroy, C., Cargile, P. A., et al. 2020, ApJL, 897, L18
Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, ApJL, 707, L1
Burstein, D. 1979, ApJ, 234, 829
Casey, C. M., Kartaltepe, J. S., Drakos, N. E., et al. 2023, ApJ, 954, 31
Cava, A., Schaerer, D., Richard, J., et al. 2018, NatAs, 2, 76
Ceverino, D., Dekel, A., & Bournaud, F. 2010, MNRAS, 404, 2151
Chabrier, G. 2003, PASP, 115, 763
Chen, C.-H., Ho, L. C., Li, R., & Zhuang, M.-Y. 2025, ApJ, 983, 60
Cheng, C., Huang, J.-S., Smail, I., et al. 2023, ApJL, 942, L19
Cheng, C., Yan, H., Huang, J.-S., et al. 2022, ApJL, 936, L19
Claeyssens, A., Adamo, A., Richard, J., et al. 2023, MNRAS, 520, 2180
Comerón, S., Elmegreen, B. G., Knapen, J. H., et al. 2011, ApJ, 741, 28
Comerón, S., Elmegreen, B. G., Salo, H., et al. 2014, A&A, 571, A58
Comerón, S., Salo, H., & Knapen, J. H. 2018, A&A, 610, A5
Comerón, S., Salo, H., Peletier, R. F., & Mentz, J. 2016, A&A, 593, L6
Conroy, C., Weinberg, D. H., Naidu, R. P., et al. 2022, arXiv:2204.02989
Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS, 386, 909
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Natur, 457, 451
Dessauges-Zavadsky, M., Schaerer, D., Cava, A., Mayer, L., & Tamburello, V. 2017, ApJL, 836, L22
Duncan, K., Conselice, C. J., Mundy, C., et al. 2019, ApJ, 876, 110
Elmegreen, B. G., & Elmegreen, D. M. 2006, ApJ, 650, 644
Elmegreen, B. G., Elmegreen, D. M., Sánchez Almeida, J., et al. 2013, ApJ, 774, 86
Elmegreen, B. G., Elmegreen, D. M., Tompkins, B., & Jenks, L. G. 2017, ApJ, 847, 14
Elmegreen, B. G., & Hunter, D. A. 2006, ApJ, 636, 712
Elmegreen, D. M., & Elmegreen, B. G. 2014, ApJ, 781, 11
Erwin, P. 2015, ApJ, 799, 226
Faisst, A. L., Yang, L., Brinch, M., et al. 2025, ApJ, 980, 204
Ferreira, L., Conselice, C. J., Sazonova, E., et al. 2023, ApJ, 955, 94
Forbes, J., Krumholz, M., & Burkert, A. 2012, ApJ, 754, 48
Franco, M., Casey, C. M., Koekemoer, A. M., et al. 2025, arXiv:2506.03256
Freeman, K., & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Géron, T., Smethurst, R. J., Dickinson, H., et al. 2025, ApJ, 987, 74
Gillman, S., Smail, I., Gullberg, B., et al. 2024, A&A, 691, A299
Gilmore, G., & Reid, N. 1983, MNRAS, 202, 1025
Glazebrook, K. 2013, PASA, 30, e056
Guo, Y., Jogee, S., Wise, E., et al. 2025, ApJ, 985, 181
Hamilton-Campos, K. A., Simons, R. C., Peeples, M. S., Snyder, G. F., & Heckman, T. M. 2023, ApJ, 956, 147
Huertas-Company, M., Shuntov, M., Dong, Y., et al. 2025, A&A, 704, A94
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Jacobs, C., Glazebrook, K., Calabrò, A., et al. 2023, ApJL, 948, L13
Jog, C. J. 2007, ASSP, 3, 137
Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864
Kartaltepe, J. S., Rose, C., Vanderhoof, B. N., et al. 2023, ApJL, 946, L15
Kauffmann, G., D'Souza, R., Monachesi, A., et al. 2025, MNRAS, 542, 688
Klein, C., Bullock, J. S., Xia, L., et al. 2025, ApJ, 998, 125
Kormendy, J. 1982, SAAS, 12, 115
Kraljic, K., Bournaud, F., & Martig, M. 2012, ApJ, 757, 60
Kümmel, M., Álvarez-Ayllón, A., Bertin, E., et al. 2022, arXiv:2212.02428
Le Conte, Z. A., Gadotti, D. A., Ferreira, L., et al. 2024, MNRAS, 530, 1984
Le Conte, Z. A., Gadotti, D. A., Ferreira, L., et al. 2025, MNRAS, 545, staf2010
Lian, J., & Luo, L. 2024, ApJL, 960, L10
Liang, X., Yu, S.-Y., Fang, T., & Ho, L. C. 2024, A&A, 688, A158
Lilly, S. J., & Carollo, C. M. 2016, ApJ, 833, 1
Martinez-Medina, L. A., Pichardo, B., Pérez-Villegas, A., & Moreno, E. 2015, ApJ, 802, 109
Messa, M., Dessauges-Zavadsky, M., Richard, J., et al. 2022, MNRAS, 516, 2420
Meštrić, U., Vanzella, E., Zanella, A., et al. 2022, MNRAS, 516, 3532
Morishita, T., Stiavelli, M., Chary, R.-R., et al. 2024, ApJ, 963, 9
Mortlock, A., Conselice, C. J., Hartley, W. G., et al. 2013, MNRAS, 433, 1185
Moster, B. P., Macciò, A. V., Somerville, R. S., Johansson, P. H., & Naab, T. 2010, MNRAS, 403, 1009
Narayanan, D., Bothwell, M., & Davé, R. 2012, MNRAS, 426, 1178
Nelson, E. J., Suess, K. A., Bezanson, R., et al. 2023, ApJL, 948, L18
Nissen, P. E., Christensen-Dalsgaard, J., Mosumgaard, J. R., et al. 2020, A&A, 640, A81
O'Leary, J. A., Moster, B. P., Naab, T., & Somerville, R. S. 2021, MNRAS, 501, 3215
Ono, Y., Harikane, Y., Ouchi, M., et al. 2023, ApJ, 951, 72
Pandya, V., Zhang, H., Huertas-Company, M., et al. 2024, ApJ, 963, 54
Perrin, M. D., Sivaramakrishnan, A., Lajoie, C.-P., et al. 2014, SPIE, 9143, 91433X
Qu, Y., Di Matteo, P., Lehnert, M. D., & van Driel, W. 2011, A&A, 530, A10

Quinn, P. J., Hernquist, L., & Fullagar, D. P. 1993, ApJ, 403, 74
Ranaivoharimina, N., Randriamampandry, T., Wang, J., Menéndez-Delmestre, K., & Gonçalves, T. S. 2024, ApJ, 977, 66
Robertson, B. E., Tacchella, S., Johnson, B. D., et al. 2023, ApJL, 942, L42
Romeo, A. B., & Agertz, O. 2014, MNRAS, 442, 1230
Romeo, A. B., Burkert, A., & Agertz, O. 2010, MNRAS, 407, 1223
Sellwood, J. A. 2014, RvMP, 86, 1
Sersic, J. L. 1968, Atlas De Galaxias Australes (Observatorio Astronomico)
Shuntov, M., Ilbert, O., Toft, S., et al. 2025, A&A, 695, A20
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Tsukui, T., Wisnioski, E., Bland-Hawthorn, J., & Freeman, K. 2025, MNRAS, 540, 3493
Übler, H., Genzel, R., Wisnioski, E., et al. 2019, ApJ, 880, 48
van der Kruit, P. C., & Searle, L. 1981, A&A, 95, 105
van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Vega-Ferrero, J., Huertas-Company, M., Costantin, L., et al. 2024, ApJ, 961, 51
Villalobos, Á., & Helmi, A. 2008, MNRAS, 391, 1806
Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011, ApJ, 742, 96
Xiang, M., & Rix, H.-W. 2022, Natur, 603, 599
Xiang, M., Rix, H.-W., Yang, H., et al. 2025, NatAs, 9, 101
Xu, D., & Yu, S.-Y. 2024, A&A, 682, L17
Yang, L., Kartaltepe, J. S., Franco, M., et al. 2025, ApJS, 68, 281
Yoachim, P., & Dalcanton, J. J. 2006, AJ, 131, 226
Yu, S., Bullock, J. S., Gurvich, A. B., et al. 2023, MNRAS, 523, 6220
Yu, S., Bullock, J. S., Klein, C., et al. 2021, MNRAS, 505, 889
Yu, S.-Y. 2026, Data Dataset for "Through Thick and Thin: The Cosmic Evolution of Disk Scale Height", v1.0, Zenodo, doi:10.5281/zenodo.18059559
Yu, S.-Y., & Ho, L. C. 2018, ApJ, 869, 29
Yu, S.-Y., & Ho, L. C. 2019, ApJ, 871, 194
Yu, S.-Y., & Ho, L. C. 2020, ApJ, 900, 150
Yu, S.-Y., Ho, L. C., Barth, A. J., & Li, Z.-Y. 2018, ApJ, 862, 13
Yu, S.-Y., Ho, L. C., & Wang, J. 2021, ApJ, 917, 88
Yu, S.-Y., Kalinova, V., Colombo, D., et al. 2022b, A&A, 666, A175
Yu, S.-Y., Xu, D., Ho, L. C., Wang, J., & Kao, W.-B. 2022a, A&A, 661, A98
Yu, S.-Y., Xu, D., Kalita, B. S., et al. 2025, A&A, 693, L9
Zhuang, M.-Y., & Shen, Y. 2024, ApJ, 962, 139