



Improving PSF Reconstruction for CSST: A Combined Approach with Deep Learning Source Selection and Empirical Correction

You Wu¹, Nan Li^{1,2}, Huan-Yuan Shan³, Peng Wei¹ , Cheng-Liang Wei⁴ , Lin Nie³, Juan-Juan Ren¹ , Zhang Ban⁵,
Xiao-Bo Li⁵, Xun Yang⁵, Yu-Xi Jiang⁵, Hong-Cai Ma⁵, Wei Wang⁵, and Chao Liu^{1,2}

¹ Key Laboratory of Space Astronomy and Technology, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China;
wuyou@nao.cas.cn, nan.li@nao.cas.cn

² School of Astronomy and Space Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

³ Shanghai Astronomical Observatory (SHAO), Chinese Academy of Sciences, Shanghai 200030, China

⁴ Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China

⁵ Space Optics Department, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
Received 2025 May 7; revised 2025 August 26; accepted 2025 September 11; published 2025 October 15

Abstract

Accurate modeling of the point spread function (PSF) is essential for scientific measurements derived from observations. This study aims to present the issues related to PSF modeling for the Chinese Space Station Survey Telescope (CSST) and then propose a practical approach for PSF reconstruction specific to CSST. We utilize PSFEX and ePSF to reconstruct the PSFs from simulated images of CSST's PSFs and quantify the impact of image size and setups of the programs on PSF reconstruction by comparing the outputs with the ground truth. We also assessed the effect of the selection of stars on PSF reconstruction, such as unresolved binaries with varying binary fractions and separations in the same manner. Our analysis indicated that both PSFEX and ePSF programs tend to overestimate the radius and introduce biases in the ellipticity measurements of the PSFs, due to the undersampling of the CSST PSF. These biases remain consistent across different program setups and image sizes but can be corrected empirically. In addition, unresolved binaries have a significant influence on PSF modeling. However, our deep learning-based method can effectively identify suitable stars for PSF modeling to avoid this problem, achieving a classification accuracy of 89.1% and an AUC of 0.937. By integrating the source selection with empirical correction modules, the measured parameters of PSFs match the ground truth, as validated by Kolmogorov-Smirnov tests. This combined approach is expected to enhance the measurement of PSFs, leading to more accurate scientific results from CSST data.

Key words: gravitational lensing: weak – (stars:) binaries (including multiple): close – methods: analytical

1. Introduction

The point spread function (PSF) is the impulse response of the imaging system to a point source, and it describes how the system blurs the incoming light (Baat et al. 2008; Born & Wolf 2013). Due to factors such as the diffraction, atmospheric disturbances (for ground-based telescopes), aberrations in the telescope's optical system, and imperfections in the instrument detectors, the image of a point source does not appear as a single bright point. Instead, the light spreads out, forming a characteristic blurred pattern, which is referred to as the PSF (Beltramo-Martin et al. 2020). The observed image of any astronomical object is a convolution of its true shape with the PSF of the instrument. Essentially, the PSF reveals how the optical instrument distorts the light from astronomical objects (Schmitz 2019). Understanding the PSF is therefore crucial for accurately interpreting the shapes and structures of astronomical objects.

The PSF directly influences the distribution of light across the detector, thereby affecting the accuracy of photometric

measurements. Accurate PSF models help disentangle overlapping sources, allowing for more precise photometric measurements (Stetson 1987; Hildebrandt et al. 2012). PSF reconstruction plays a critical role in deconvolution techniques, which are particularly useful for resolving fine details in galaxy morphology, star clusters, or crowded stellar fields, thereby increasing the effective resolution of astronomical images (Hanisch et al. 1997; Bertero & Boccacci 2005). In addition to its impact on photometry and image resolution, PSF modeling is essential for astrometric measurements, such as parallax determinations and proper motion studies of stars (Mighell 2005; Lindegren et al. 2018). By affecting the centroiding of stars and galaxies, the PSF significantly determines the precision of positional measurements. Through accurate PSF reconstruction, potential biases introduced by instrumental effects are minimized, ensuring that the measured positions of astronomical objects remain accurate and reliable (Libralato et al. 2024). Among its various applications, PSF modeling is particularly critical in the context of weak

gravitational lensing. Weak lensing is a key observational tool in cosmology for probing the distribution of dark matter and constraining cosmological parameters (Mandelbaum 2018). Given the small lensing signals involved, even minor errors in PSF estimation can significantly affect the results. Accurate PSF modeling is crucial to distinguish the small distortions caused by gravitational lensing from those introduced by the instrument (Rhodes et al. 2000; Erben et al. 2001).

PSF modeling methods can be broadly categorized into parametric and non-parametric approaches (Liaudat et al. 2023). Parametric models use known instrument information for PSF modeling, typically involving forward modeling to refine the PSF, with star observations used for calibration. Such models, including the simple Gaussian and Moffat profiles, offer computational efficiency but often fall short in accurately capturing the complex variations in PSFs seen in real astronomical data. An example of a widely used parametric model is *Tiny Tim* (Krist 1993, 1995; Krist et al. 2011), which was designed for the Hubble Space Telescope (HST). *Tiny Tim* models the PSF by leveraging detailed knowledge of the telescope’s optics and simulates the PSF across different wavelengths and detector positions. In contrast, Anderson & King (2000) developed the effective PSF (ePSF) method, a non-parametric technique that iteratively models the PSF directly from observed stars. The ePSF describes how a star’s light is distributed across each pixel, offering a more data-driven and accurate representation of the PSF (Anderson 2016). Hoffmann & Anderson (2018) conducted a comparative study of PSF modeling techniques and found that the effective PSF method consistently outperforms *Tiny Tim*, revealing the limitations of physically based parametric approaches. Non-parametric approaches primarily rely on data-driven methods to model the PSF, rather than predefined functional forms. These methods directly measure the PSF from observed stars in the field and estimate it at other positions, making them more flexible and adaptable to real data. This category includes interpolation-based models like Kaiser–Squires–Broadhurst method (Kaiser et al. 1995), as well as techniques based on Principal Component Analysis (Lupton et al. 2001; Jee et al. 2007; Jee & Tyson 2011), Resolved Component Analysis (Ngolè et al. 2016), Multi-CCD PSF models (Liaudat et al. 2021), PSFs in the Full Field (Jarvis et al. 2021), and PSF Extractor (PSFEX) (Bertin 2011). Of these approaches, PSFEX stands out as the most typical and widely used tool, particularly in weak lensing surveys such as DES (Zuntz et al. 2018), HSC (Mandelbaum et al. 2018), and CFIS (Guinot et al. 2022). PSFEX models the PSF by extracting information from bright stars, making it highly accurate even in complex environments and capable of modeling spatially varying PSFs across wide fields of view. It has become a standard tool for large astronomical surveys due to its efficiency and adaptability.

The upcoming Chinese Space Station Survey Telescope (CSST; Zhan 2011; Cao et al. 2018) is a 2 m aperture space telescope designed to conduct photometric and slitless spectroscopic surveys across 17,500 square degrees of the sky. It will cover a wavelength range from 0.255 μm to 1.0 μm , including at least six bands (NUV, u , g , r , i , z) utilizing 18 filters, with each filter covering a single detector. CSST has a large field of view of 1 square degree, with a pixel size of approximately $0''.074$, and it boasts a spatial resolution of $\sim 0''.15$. Over its 10 yr mission, CSST is expected to collect photometric data for billions of stars and galaxies, as well as spectroscopic data for hundreds of millions of objects. For large-scale surveys like CSST, the accuracy of PSF modeling becomes a critical factor. It ensures precise galaxy shape measurements, which are essential for studying the distribution of dark matter and the large-scale structure of the Universe. Inaccurate PSF modeling can result in systematic errors in galaxy shape measurements, leading to biased estimates of the weak lensing shear and, consequently, inaccurate cosmological parameter determinations (Hirata & Seljak 2003).

Since the precision of data-driven PSF modeling heavily depends on the quality of the input data, selecting clean, isolated single stars is essential for achieving accurate results. However, distinguishing unresolved binaries⁶ from single stars remains a significant challenge. This issue is fundamentally linked to the spatial resolution of telescopes, which is determined by diffraction limits that restrict the ability to resolve two closely spaced objects. For example, the CSST has a photometric instrument with a pixel size of approximately $0''.074$ and a spatial resolution of about $0''.15$. As a result, stars separated by less than two physical pixels cannot be distinguished. The presence of unresolved binaries contaminates the sample used for PSF modeling, ultimately degrading the accuracy of the model. Furthermore, systematic errors may arise during the modeling process, compounding the impact on the precision of the PSF.

This study aims to optimize PSF modeling for the upcoming CSST mission, enhancing its capacity to achieve the precision required in large-scale surveys. We focus on eliminating the contamination caused by unresolved binaries in PSF modeling and addressing systematic errors. The analysis is conducted using simulated data from the 18 CCD detectors of CSST, providing valuable insights and preparations for future applications with real CSST observations.

The structure of this paper is as follows: In Section 2, we describe the generation of simulated CSST images and the methodologies used for PSF modeling and measurement. In Section 3, we analyze the accuracy of PSF modeling using

⁶ Unresolved binaries occur when two stars in a binary system are closely aligned along the same line of sight, with their angular separation smaller than the telescope’s spatial resolution limit. In this case, their light blends into a single point, causing them to overlap in the image and appear indistinguishable from a single star.

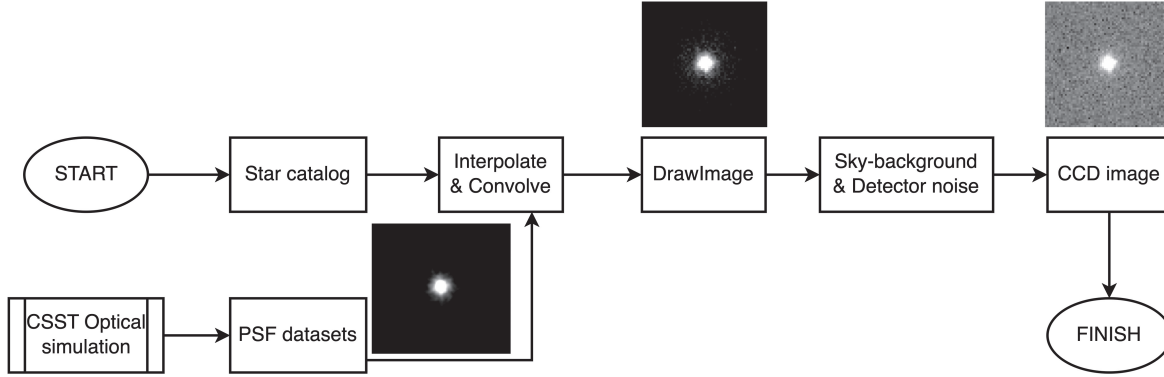


Figure 1. Imaging pipeline overview: An ideal image is first generated by convolving the flux of a given star from the catalog with an interpolated PSF from the CSST optical simulation data sets, which include various optical aberrations. Detector effects such as Poisson noise from photon counting, sky-background levels, CCD dark current, Gaussian read noise, as well as bias and gain factor corrections are subsequently applied, producing the final simulated CCD image.

simulated CSST images, conducting a detailed investigation into the effects of image size, sampling rate, and the contamination of modeling samples caused by unresolved binaries. Section 4 explores deep learning methods for source selection and applies empirical correction techniques to mitigate systematic errors. By combining both source selection and empirical correction, we evaluate the resulting improvements in PSF modeling precision. Finally, in Section 5, we present a summary of the results and discuss future directions.

2. Data Sets and Methodology

2.1. Mock PSF Images of CSST

To realistically simulate the CSST images, a high-fidelity data set of CSST PSFs is used to model the optical system’s impact on image quality. These PSFs are produced using an optical emulator with six modules to simulate the optical aberration, including mirror surface roughness, fabrication errors, CCD assembly errors, gravitational distortions, and thermal distortions. For each star in the catalog, we calculate its flux within the CSST filter system and generate its image by convolving the star with a PSF interpolated from the data set at a given position. Various types of noise are incorporated in our image simulator, including shot noise, sky background and detector effects. With the CSST throughputs we use *GalSim* (Rowe et al. 2015) to generate photons from a given object and add Poisson noise from both the sky background and the dark current of a CCD detector. The levels of sky background are (0.004, 0.021, 0.164, 0.207, 0.212, 0.123, 0.037) $e^-/\text{pixel}/s$ for the bands of (NUV, *u*, *g*, *r*, *i*, *z*, *y*), respectively, and the dark current is 0.02 $e^-/\text{pixel}/s$. Thus, for example, in *i*-band this results in an average of approximately 35 e^-/pixel in a 150 s exposure. Read noise is modeled as a Gaussian distribution with $\sigma = 5.0$ e^-/pixel . Bias and gain factor are then applied to simulate the production of mock images on the detector. These considerations ensure that our simulated

images closely resemble those obtained under real observational conditions. The resulting images have customizable sizes, with each pixel having a size of 0.074, corresponding to the pixel size of the photometric instrument in CSST. An overview of our imaging pipeline is shown in Figure 1.

2.2. PSF Modeling and Measurement

For the purpose of modeling the PSF, we utilize *PSFEx* to extract precise PSF models. *PSFEx* is a powerful tool designed to extract precise PSF models from images processed by *SExtractor* (Bertin & Arnouts 1996), facilitating model-fitting photometry and morphological analyses. Key features of *PSFEx* include modeling both non-parametric and parametric PSF models, reconstructing PSF models from undersampled images using super-resolution techniques, and modeling PSF variations as polynomial functions of position. The modeling process involves starting with input catalogs, pre-selecting vignettes, using built-in or user-supplied image bases, deriving PSF models, and computing homogenization kernels. PSF models can be controlled at different sampling rates and are represented as images with adjustable resolution, allowing for flexibility and interoperability with other software tools.

Accurate measurement of the PSF is essential for understanding and correcting distortions in astronomical images. We employed *GalSim* library to measure the PSF. The process begins by converting the PSF data into a *GalSim* image. Subsequently, the adaptive moments of the *GalSim* image are determined using the *FindAdaptiveMom* module. This method estimates the best-fit elliptical Gaussian to the object, computed iteratively by initially guessing a circular Gaussian as a weight function. Weighted moments are computed and then recomputed using the results of the previous step as the weight function. The iterative process continues until the measured moments converge with those used for the weight function, allowing for robust measurement of the PSF parameters even in the presence of noise or slight distortions.

(Bernstein & Jarvis 2002; Hirata & Seljak 2003). The adaptive moments of an image $I(x)$ are defined by the following equations:

$$M_{ii} = 2 \frac{\iint (x_i - x_{0,i})^2 \omega(x) I(x) d^2x}{\iint \omega(x) I(x) d^2x} \quad (1)$$

$$M_{jj} = 2 \frac{\iint (x_j - x_{0,j})^2 \omega(x) I(x) d^2x}{\iint \omega(x) I(x) d^2x} \quad (2)$$

$$M_{ij} = 2 \frac{\iint (x_i - x_{0,i})(x_j - x_{0,j}) \omega(x) I(x) d^2x}{\iint \omega(x) I(x) d^2x}. \quad (3)$$

In the above expressions, M_{ii} and M_{jj} represent the second moments along the i and j axes, respectively, while M_{ij} captures the cross-correlation between the i and j axes. These moments characterize the shape and orientation of the PSF in terms of its ellipticity and size. $\omega(x)$ denotes the weight function used in the calculation. The weighted centroid x_0 can be interpreted as a 2D vector $(x_{0,i}, x_{0,j})$, where the components are given by:

$$x_{0,i} = \frac{\iint x_i \omega(x) I(x) d^2x}{\iint \omega(x) I(x) d^2x} \quad (4)$$

$$x_{0,j} = \frac{\iint x_j \omega(x) I(x) d^2x}{\iint \omega(x) I(x) d^2x}. \quad (5)$$

Several key PSF parameters are derived from the adaptive moments, including the trace of the second moment matrix, T . The trace represents the total intensity distribution of the PSF and is calculated as the sum of the diagonal elements M_{ii} and M_{jj} , given by:

$$T = M_{ii} + M_{jj}. \quad (6)$$

The size of the PSF, denoted as R , is calculated as:

$$R = \sigma^* \sqrt{2}, \quad (7)$$

where σ , representing the width of the best-fit elliptical Gaussian, is defined as:

$$\sigma = \sqrt{\frac{T}{2}}. \quad (8)$$

Additionally, the ellipticity components e_1 and e_2 , as well as the total ellipticity e , are obtained from the observed shape of the PSF:

$$e_1 = \frac{M_{ii} - M_{jj}}{T} \quad (9)$$

$$e_2 = \frac{2M_{ij}}{T} \quad (10)$$

$$e = \sqrt{e_1^2 + e_2^2}. \quad (11)$$

The components e_1 and e_2 describe the shape distortion of the PSF, with e_1 measuring elongation along the x/y axes, and e_2 quantifying elongation along diagonal directions. These parameters are particularly important in weak lensing studies, where even small PSF distortions can introduce significant biases in the inferred shapes of galaxies. While ellipticity characterizes the PSF's shape, the radius parameter R measures its overall size and spread, affecting the resolution and the ability to accurately recover features in astronomical images. By combining R with the ellipticity components e_1 and e_2 , the PSF is comprehensively characterized in terms of both size and shape (Bernstein & Jarvis 2002; Mandelbaum et al. 2005). Precise measurements of these parameters allow us to apply effective corrections to observed galaxy shapes, minimizing biases introduced by the PSF in scientific analyses, which is essential for obtaining accurate and reliable astrophysical measurements.

3. Analysis of PSF Reconstruction

This section investigates several key factors that influence the accuracy of PSF reconstruction, including the input image size, the sampling rate used to model the PSF at sub-pixel resolution, and the contamination from unresolved binary stars in the PSF modeling sample.

3.1. Impact of Image Size on PSF Modeling Accuracy

To evaluate whether different image sizes affect PSF modeling, we generated simulated images with sizes of 16, 32, 64, and 128 pixels. Each set contained 500 images, with a pixel size of $0''.074$ and a total flux of 50,000. Then, PSFEX was employed to perform PSF modeling on each set of simulated images, using a sampling rate of 2 times the original pixel size. Following the modeling, we measured key PSF parameters, including the size R , and the ellipticity components e_1 , e_2 .

The results are summarized in Figure 2, which shows the distributions and cumulative probabilities of the PSF parameters (R , e_1 and e_2) for each image size. The density plot of e_1 shows that the 16 pixel image size exhibits a slight deviation from the others, with small variations between different image sizes. In contrast, for e_2 , there is virtually no variation across different image sizes. The density plot and cumulative probability plot of R indicate slight variations across different image sizes, but the overall spread is minimal. Overall, this analysis indicates that the image size has a minimal impact on PSF modeling, as the PSF parameters measured from different image sizes are consistent with each other, suggesting that variations in image sizes do not significantly affect the accuracy and reliability of the PSF modeling process.

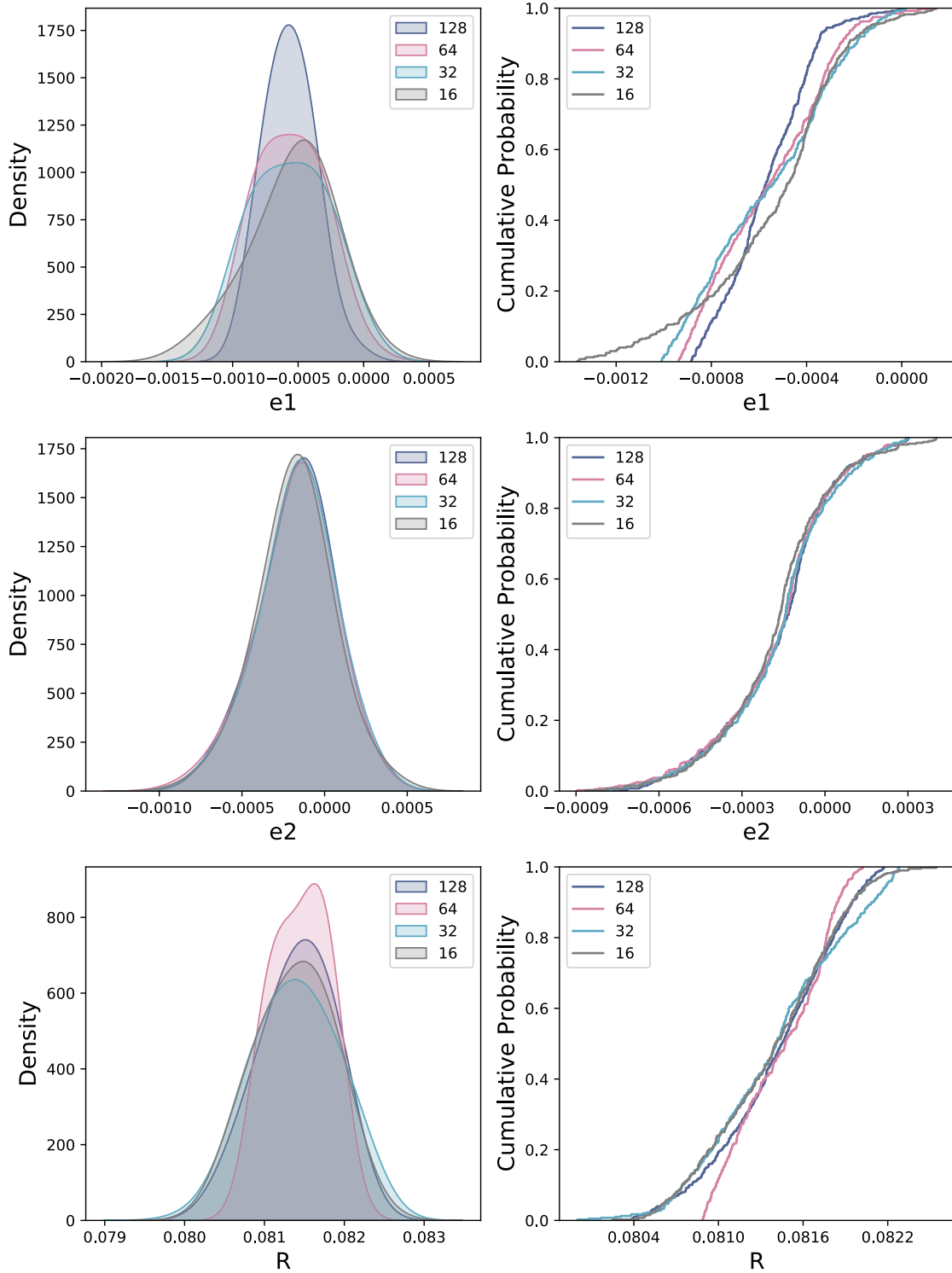


Figure 2. Impact of image sizes on PSF measurements of ellipticity components e_1 , e_2 , and radius parameters (R) for CCD12 of CSST. The left panels display the kernel density estimate (KDE) distributions, while the right panels show the cumulative distribution function (CDF) curves for each parameter. Different colors correspond to simulated images of varying sizes (16, 32, 64, and 128 pixels), reflecting the dimensions of the generated stellar images used in PSF reconstruction.

3.2. Impact of Sampling Rate on PSF Modeling Accuracy

Another crucial parameter in PSF modeling is the sampling rate, which determines the resolution at which the PSF is reconstructed relative to the original pixel scale. This parameter is particularly important when dealing with undersampled data, where the native pixel size is too large to adequately resolve the PSF structure. A higher sampling rate enables the recovery of finer PSF details by effectively increasing the resolution of the model, which is essential for accurate astrometric and photometric measurements. However, higher sampling rates also increase computational complexity and may introduce noise if not properly constrained. Conversely, lower sampling rates are less demanding computationally but may result in a loss of details and potential inaccuracies. Therefore, selecting an appropriate sampling rate is crucial for balancing detail resolution, computational efficiency, and noise control, ensuring robust PSF characterization.

To more precisely evaluate the impact of different sampling rates on PSF morphology, we conducted tests by performing PSF modeling using different sampling rates on simulated CSST images and simulated images with a Gaussian PSF function, separately, and then measured the PSF parameters R , $e1$ and $e2$. As a reference, in addition to `PSFEx`, we also employed the `ePSF` method for PSF reconstruction. This method provides a detailed empirical model of the PSF, which can be implemented via the `EPSFBuilder` module in the `Photutils` library (v1.2.0 Bradley et al. 2016). The `ePSF` is constructed by oversampling beyond the detector pixels, and its accuracy is iteratively refined. The model is fitted to star profiles, and adjustments are made based on the residuals, progressively improving its precision.

3.2.1. PSF Modeling from Mock CSST Images

We employed both `PSFEx` and `ePSF` methods to perform PSF modeling on mock images of CCD12. The sampling rates varied from 2 to 10. After constructing the PSF models using these methods, we measured the PSF parameters R , $e1$, and $e2$ for each model. To evaluate the reconstructed PSF models, we compared these measurements with the true values obtained by directly measuring the original PSF matrix of CCD12. These results are presented in Figure 3. The left panels show the results for CCD12 of CSST: the top panel corresponds to the ellipticity component $e1$, the middle panel to $e2$, and the bottom panel to the PSF size R . It can be seen that both `PSFEx` and `ePSF` systematically underestimate the $e1$ and $e2$ values, while the R values are slightly overestimated. Additionally, the variations across different sampling rates are minimal, suggesting that the choice of sampling rates has a limited impact on the accuracy of PSF parameter measurements. These measurement deviations could be attributed to intrinsic issues with the PSF models, characteristics of the simulated data, or systematic errors in the interpolation process used in both methods.

3.2.2. PSF Modeling from Gaussian PSF Images

To verify whether the deviations in measured values were due to the PSF itself, we replaced the CSST's PSF with a Gaussian function. Using the same methodology employed for generating simulated CSST data, we generated 500 simulated images with Gaussian PSF. These images had a size of 64×64 pixels, with each image having a total flux of 50,000, consistent with the previous tests. The Gaussian function was designed with an FWHM of $0''.15$ to match the FWHM of the CSST's PSF. We then employed both `PSFEx` and `ePSF` methods to perform PSF reconstruction on these simulated images at different sampling rates, ranging from 2 to 10. Following the reconstruction, we measured the PSF parameters for each model. This approach allowed us to evaluate the measurement deviations of the reconstructed Gaussian PSF under different sampling rates and compare the results with those obtained using the CSST PSF.

The right panels of Figure 3 illustrate the variations in the PSF parameters across different sampling rates for the Gaussian PSF model. From the top and middle panels, it is evident that the $e1$ and $e2$ values directly measured from the Gaussian model remain close to zero, as expected for a symmetric Gaussian PSF. Both the `PSFEx` and `ePSF` methods show deviations from the directly measured Gaussian $e1$ and $e2$ values, with the reconstructed parameters slightly overestimated. The deviations for the PSF reconstructed using the `PSFEx` method are relatively smaller. Furthermore, the variations in $e1$ and $e2$ remain consistent across different sampling rates, indicating that the sampling rate does not have a major impact on the reconstruction of the ellipticity components for the Gaussian PSF. The bottom panel shows that both methods slightly overestimate the R values. For the `ePSF` method, the degree of overestimation remains consistent across different sampling rates, while for the `PSFEx` method, the R value decreases when the sampling rate is 8.

Combining the analysis from Section 3.2.1, it is evident that although the `PSFEx` and `ePSF` methods can reconstruct PSFs from stellar images with reasonable accuracy, there are systematic biases leading to slight overestimation of the radius parameter and deviations in the ellipticity components. A potential cause of these biases is the interpolation process used to reconstruct the PSF, where small errors accumulate and systematically affect the estimated PSF shape. Additionally, the original input PSF matrix in the simulation is undersampled, introducing errors that propagate through the reconstruction process. Therefore, there may be inherent limitations in directly reconstructing PSF shapes from images, and these systematic biases must be carefully considered when interpreting measurement results.

3.3. Source Selection Effects

For accurate PSF reconstruction, it is essential to use a clean and isolated sample of single stars. However, the mistaken

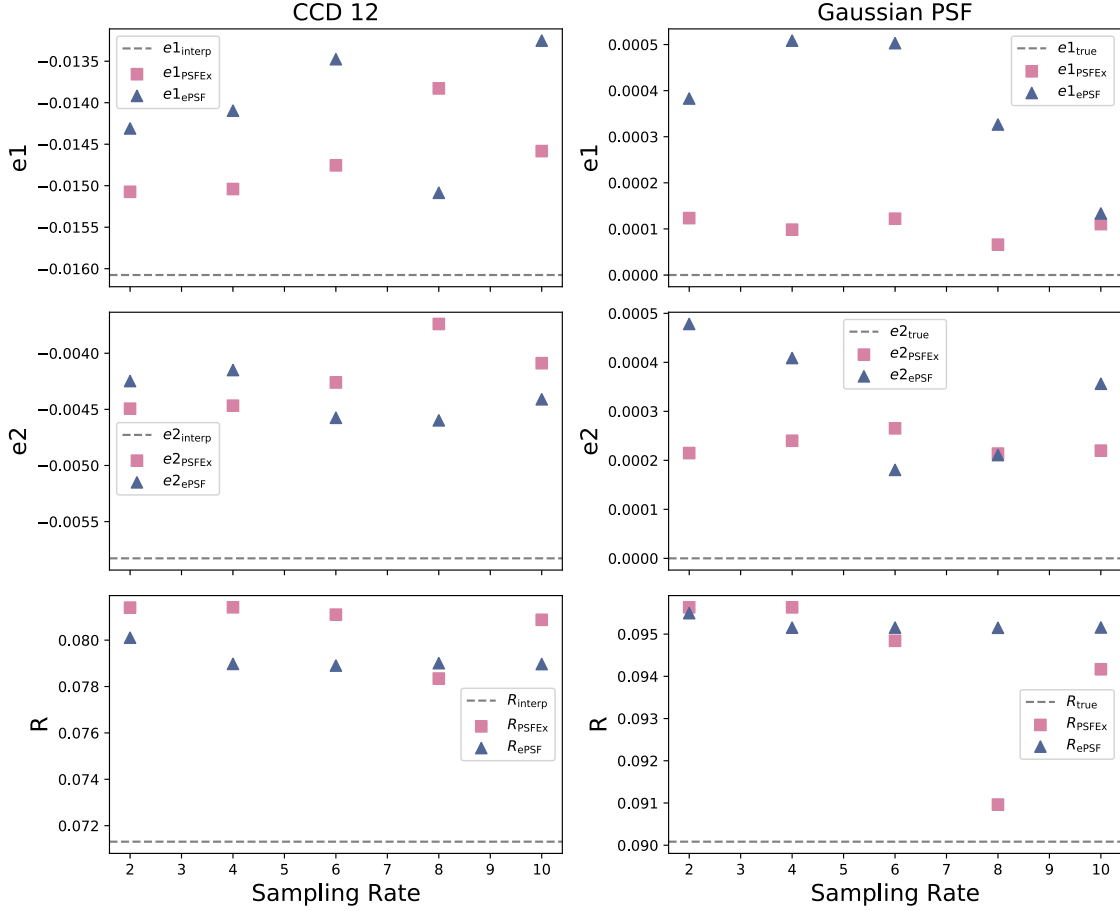


Figure 3. Variations in ellipticity components e_1 , e_2 , and radius parameter (R) measured using different PSF reconstruction methods (PSFEx and ePSF) across sampling rates (2–10). Pink squares represent parameter values measured after PSFEx reconstruction, navy triangles indicate measurements after ePSF reconstruction. Left panels: Results for CCD12 of CSST, where horizontal dashed lines indicate reference values obtained by interpolating the original CSST PSF matrix. Right panels: Results for the Gaussian PSF model, where horizontal dashed lines represent the true Gaussian PSF parameters.

inclusion of unresolved binaries in this sample can lead to significant errors in PSF modeling. In this section, we analyze the impact of unresolved binaries on PSF reconstruction.

3.3.1. Impact of Binary Fraction

To understand how different proportions of unresolved binaries in the sample affect the accuracy of PSF reconstruction, we established data sets with varying binary fractions. Following the method described in Section 2.1, we generated a series of simulated binary images with separations uniformly distributed from 0.1 to 2 pixels, ensuring a total flux of 50,000 for each pair. These binary star images were then combined with single star images to produce data sets with varying binary fractions. Next, we reconstructed the PSF models for these data sets using PSFEx and measured the resulting PSF parameters for each model.

Figure 4 illustrates the impact of varying binary fractions on the PSF models for CCD 12 of the CSST. In the top left panel, the mean eccentricity values exhibit a slight increasing trend as the binary fraction rises. However, the variations remain within the 1σ

uncertainty range of the true values (gray shaded region), indicating that the effect is statistically insignificant. We further analyzed the distributions of the ellipticity components e_1 and e_2 under varying binary fractions, as shown in the right panel of Figure 4. At a binary fraction of 0.1, the distributions of e_1 and e_2 are compact and concentrated near zero, indicating minimal distortion in the PSF models. As the binary fraction increases, the distribution of e_1 gradually expands and shifts, suggesting a systematic elongation of the PSF along the x/y axes. In contrast, e_2 primarily exhibits an increase in dispersion rather than a directional shift, indicating greater PSF shape variability along diagonal directions. At the highest binary fractions (e.g., 0.8–0.9), the e_2 distribution becomes significantly more elongated and dispersed compared to e_1 , further emphasizing the anisotropic nature of the PSF distortions. These differences indicate that increasing binary fraction introduces asymmetric distortions in the reconstructed PSF, with a dominant elongation along the diagonal axes rather than an isotropic expansion, which significantly affects the overall PSF shape. The left bottom panel of Figure 4 reveals a clear

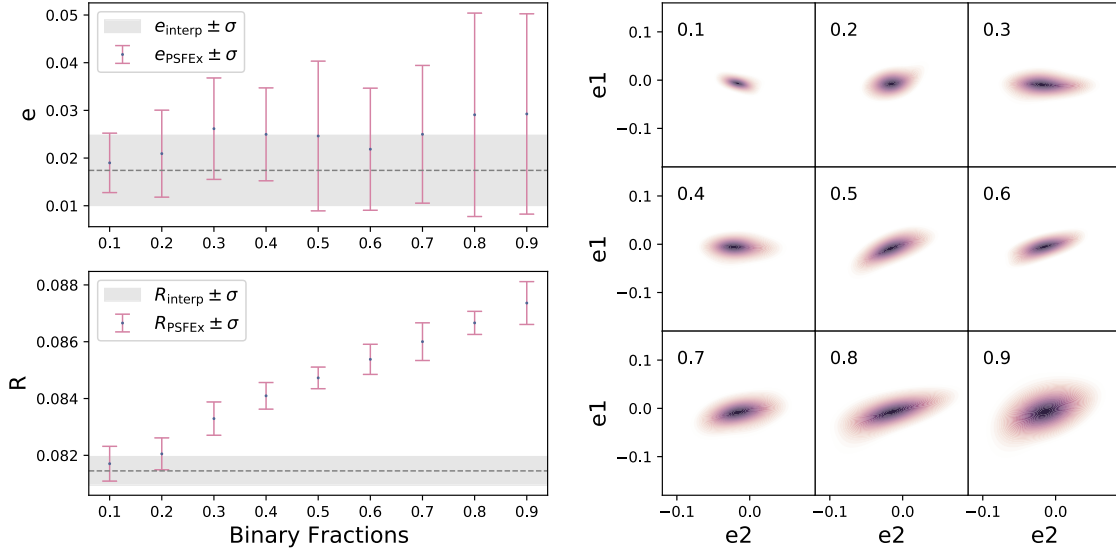


Figure 4. Effects of binary fractions on PSF reconstruction. Left panels: Impact of binary fractions on the eccentricity (e) and radius parameter (R) of PSF models. Each point represents the mean value for a given binary fraction, with error bars indicating the standard deviation. The gray dashed lines represent the mean parameter values obtained through direct interpolation, while the gray shaded areas indicate the 1σ uncertainty range for these values. Right panel: Distribution of ellipticity components (e_1 , e_2) across different binary fractions from 0.1 to 0.9 for PSF models.

increasing trend in the radius parameter R as the binary fraction increases, with values progressively exceeding the true mean and moving beyond the 1σ deviation range. These results indicate that the ellipticity components e_1 and e_2 as well as the radius parameter R of the PSF models are significantly affected by contamination from unresolved binaries in the modeling sample. As the binary fraction increases, these parameters show more significant variations, emphasizing the need to exclude unresolved binaries from the PSF reconstruction sample to maintain accurate measurements.

3.3.2. Impact of Binary Separation

For a more comprehensive understanding of the contamination caused by unresolved binaries, we examined how varying separations between them affect PSF reconstruction. Using CCD12 of the CSST as an example, we generated simulated binary star images with separations ranging from 0.1 to 2.0 pixels, in 0.1 pixel intervals, while keeping the total flux for each binary system constant at 50,000. These binary images were then analyzed to assess the changes in the ellipticity components and size parameters of the PSF models.

Figure 5 presents the results of this analysis. The top and middle panels illustrate the variations in e_1 and e_2 as binary separation increases. Both ellipticity components remain relatively stable for separations below 0.6 pixels but begin to exhibit significant deviations beyond this point, exceeding the standard deviation range of the true value. The bottom panel depicts the variation in the radius parameter R with increasing binary separation. For separations greater than 0.3 pixels, R begins to

show a noticeable linear increase, with values progressively exceeding the true value and falling outside the standard deviation range. These results demonstrate that unresolved binaries have a significant impact on the reconstructed PSF models. Larger binary separations introduce greater distortions, affecting the ellipticity measurements and systematically biasing the size measurements, leading to an overestimation of the PSF size. This highlights the importance of carefully selecting single-star sources to minimize contamination and ensure the accuracy and reliability of PSF reconstruction.

4. Toward Accurate PSF Parameter Calibration

4.1. Source Selection

We introduce a neural network-based approach for source selection, which aims to improve the precision of PSF modeling by ensuring that only clean, high signal-to-noise ratio (SNR) single-star sources are used, thus reducing distortions caused by unresolved binaries. It is important to emphasize that the current source selection model is designed specifically for identifying isolated point sources within 2 pixels ($\sim 0''.15$) in CSST images.⁷

⁷ Potential contamination from non-stellar sources (e.g., compact galaxies) can often be excluded during early-stage screening using morphological separation or spectroscopic identification. Instrumental artifacts such as cosmic rays, hot pixels, and detector defects are expected to be effectively flagged and masked by the CSST standard image processing pipeline before source catalogs are generated. Additionally, blended stellar sources involving unresolved triple or higher-order systems are intrinsically rare, and the probability of multiple stars falling within such a small angular separation ($\sim 0''.15$) is very low. Given these considerations, we assume a clean input catalog where most such contaminants have been removed or mitigated during early data reduction.

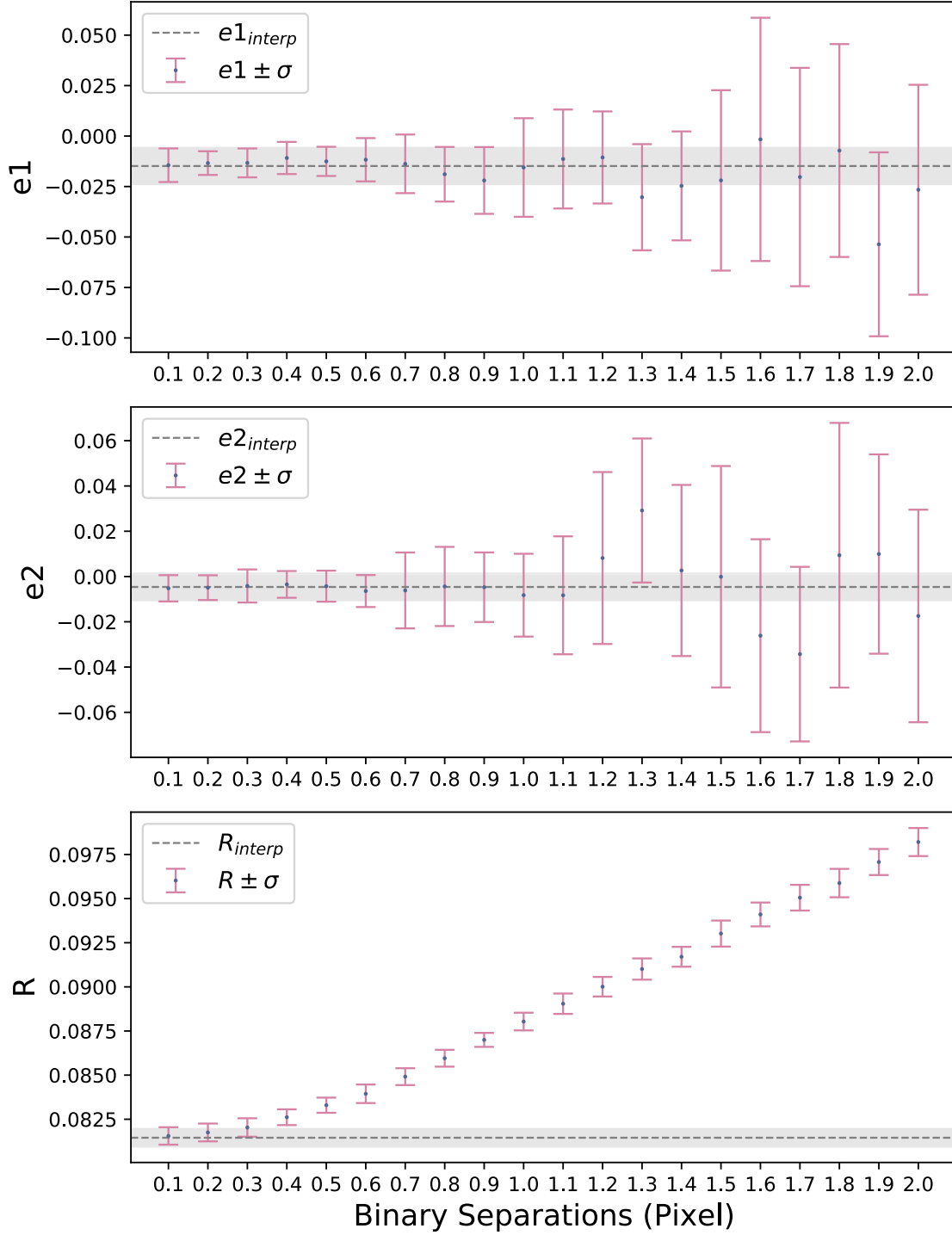


Figure 5. Impact of binary separations ranging from 0.1 to 2.0 physical pixels on the ellipticity components e_1 , e_2 , and radius parameters (R) of PSF models. All panels include error bars representing the standard deviation for each binary separation. The gray dashed line represents the mean parameter values obtained through direct interpolation, while the gray shaded areas indicate the 1σ uncertainty range for these values.

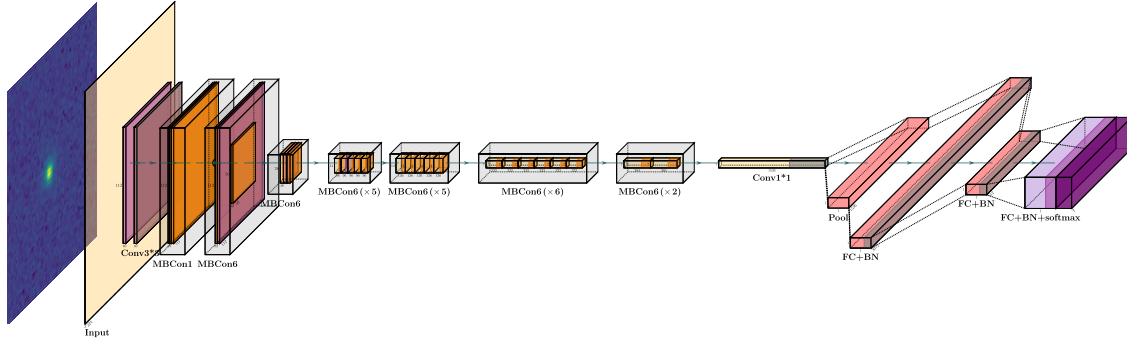


Figure 6. The architecture of the neural network model, featuring an *EfficientNet* backbone with multiple MBconv6 blocks followed by fully connected layers and a softmax output layer for classification. Each MBconv6 block consists of three main components: an expansion convolution layer to increase channel dimensions, a depth-wise separable convolution for spatial feature extraction, and a squeeze-and-excitation module for adaptive feature recalibration.

This is supported by Hu et al. (2011), who demonstrated through artificial-star tests that a separation of two pixels is the minimum at which stellar sources can be reliably resolved in photometric measurements.

4.1.1. Network Architecture

Building on the work of Wu et al. (2023), who developed a neural network method based on a modified *EfficientNet* architecture (Tan 2019) to distinguish between unresolved binaries and single stars, the model demonstrated high performance and accuracy on CSST simulated images and was further validated on real HST observations. Consequently, we employed the same approach for our analysis of CSST simulated images.

To create a more robust training and validation data set, we simulated stellar images using the PSFs from all 18 CCDs of CSST. For each CCD, we generated 3000 single-star images and 3000 unresolved binary images, with binary separations uniformly distributed between 0.1 and 2.0 pixels. This resulted in a total of 108,000 training images. To ensure consistency between the single-star and binary samples and avoid model bias, the total flux was uniformly distributed between 5000 and 50,000 for both types of stars. In addition, an independent validation data set containing 21,600 images was generated in the same way. The neural network was trained on the full training set and evaluated on this independent validation set, ensuring that the model's performance is tested against PSF variations across different CCDs.

The neural network architecture used in our work, as illustrated in Figure 6, is a modified version of *EfficientNet*, adapted to enhance its performance in distinguishing single stars from unresolved binaries. Specifically, the model consists of an initial convolutional layer for preliminary feature extraction, followed by multiple MBconv6 blocks. Each MBconv6 block incorporates an expansion convolution layer to increase feature channels, a depth-wise separable convolution for efficient spatial feature extraction, and a squeeze-and-excitation module for adaptive

recalibration of channel-wise features. After these convolutional layers, global pooling integrates spatial information, followed by fully connected layers with batch normalization and a softmax layer for binary classification. Further technical details of the model architecture and training procedure follow the methodology described in Wu et al. (2023).

4.1.2. Model Performance

After 25 epochs of training, the model's loss function converged to its minimum, achieving an accuracy of 89.1% on the independent validation set. The ROC curve shown in Figure 7 demonstrates that the model remains highly effective in distinguishing between unresolved binaries and single stars, with an AUC score of 0.937. The curve stays close to the upper-left corner, reflecting a high true positive rate and a low false positive rate across decision thresholds. The confusion matrix further demonstrates that the model correctly classified the vast majority of both binary and single-star images. The recall for binary systems reached 92.0%, reflecting strong sensitivity to unresolved binaries, while the recall for single stars was 86.6%, indicating reliable retention of clean single-star samples. Despite the additional complexity introduced by training on diverse PSFs across multiple CCDs, the neural network maintained high overall accuracy and robustness in source selection, ensuring that clean single-star inputs can be reliably identified for PSF reconstruction.

4.2. Empirical Correction of Measured PSF Parameters

In Section 3.2.1, we analyzed the potential systematic biases in the PSF reconstruction of CSST CCD12, which could lead to measurement errors in the radius parameter and ellipticity components. To address and correct these measurement deviations, we first measured the PSF reconstruction parameters for all CCDs of the CSST. This process involved generating simulated images for each CCD, using *PSFEX* to reconstruct the PSF models, and then measuring the $e1$, $e2$, and R parameters. These measurements were compared with the

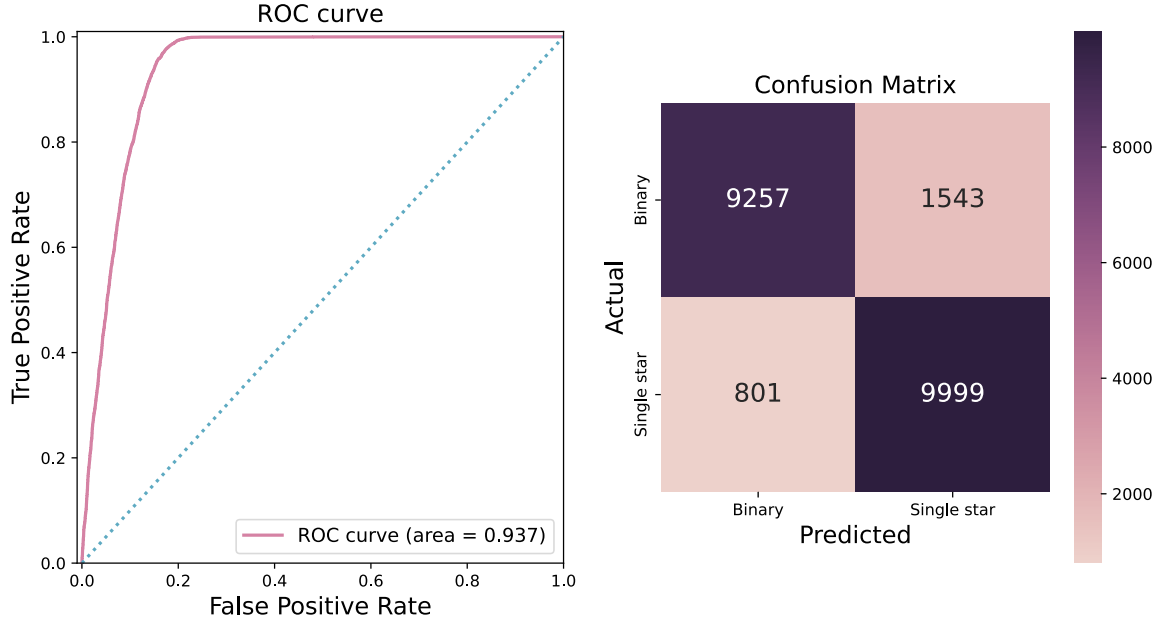


Figure 7. Left panel: ROC curve of the model. The navy dotted line represents the performance of a random classifier, while the AUC is calculated from the area under the ROC curve (pink line). Right panel: Confusion matrix of the model, displaying the number of true and predicted classifications for binary and single star images. The color intensity in each cell corresponds to the frequency of each classification result.

parameters obtained from directly interpolating the PSF models, and the results are presented in Figure 8. This figure illustrates the mean differences in the ellipticity components $e1$, $e2$, as well as the percent differences in the radius parameter R , for each CCD of CSST. The CCDs are arranged in increasing wavelength order from NUV to the y -band. The mean differences are calculated as the differences between the mean values derived using `PSFEX` and those obtained through direct interpolation for each CCD. The percent difference in R is the relative difference between the mean values, expressed as a percentage of the mean.

The top panel of Figure 8 shows the mean differences in $e1$, where most CCDs exhibit a negative difference between the measurements obtained through `PSFEX` reconstruction and those from direct interpolation of the PSF models. The mean differences are more pronounced in the shorter wavelengths. As the wavelength increases, both the mean differences and error bars show a decreasing trend, particularly in the longer wavelengths (such as the y -band). The middle panel of Figure 8 depicts the mean differences in $e2$. While the overall range of deviations in $e2$ is similar to that of $e1$, with both showing a decrease in deviation as the wavelength increases, the majority of CCDs show a positive difference. From the bottom panel of Figure 8, it can be seen that the `PSFEX` method consistently overestimates the R values compared to direct interpolation for all CCDs. The maximum deviation reaches approximately 15%, observed in the shorter wavelengths (such as NUV and the u band). As the wavelength

increases, starting from the r -band, the percentage deviation significantly decreases, showing a clear trend of error reduction. In longer wavelengths, particularly in the y -band, the deviation is approximately 8%, which is the smallest observed across all CCDs.

The analysis reveals that the reconstructed ellipticity components and radius parameter in CSST exhibit systematic biases across all CCDs, with varying degrees of deviation depending on the wavelength. A consistent trend is observed in which both systematic biases and uncertainties decrease as the wavelength increases, indicating that PSF reconstruction is more accurate at longer wavelengths. This can be attributed to the broadening of the PSF in longer wavelengths. As the PSF becomes wider, it is better sampled by the pixels, which allows for more accurate PSF reconstruction and reduces systematic errors. In contrast, at shorter wavelengths, the sharper PSF is more sensitive to undersampling and noise, leading to larger deviations and higher uncertainties in the reconstructed PSF. These results suggest that it is necessary to individually correct the systematic biases for each CCD to improve the PSF modeling accuracy, particularly for short-wavelength CCDs, where systematic errors have a greater impact on scientific measurements.

4.2.1. Correction of PSF Sizes

For each CCD, we generated simulated images at various SNRs, following the same simulation process as described in

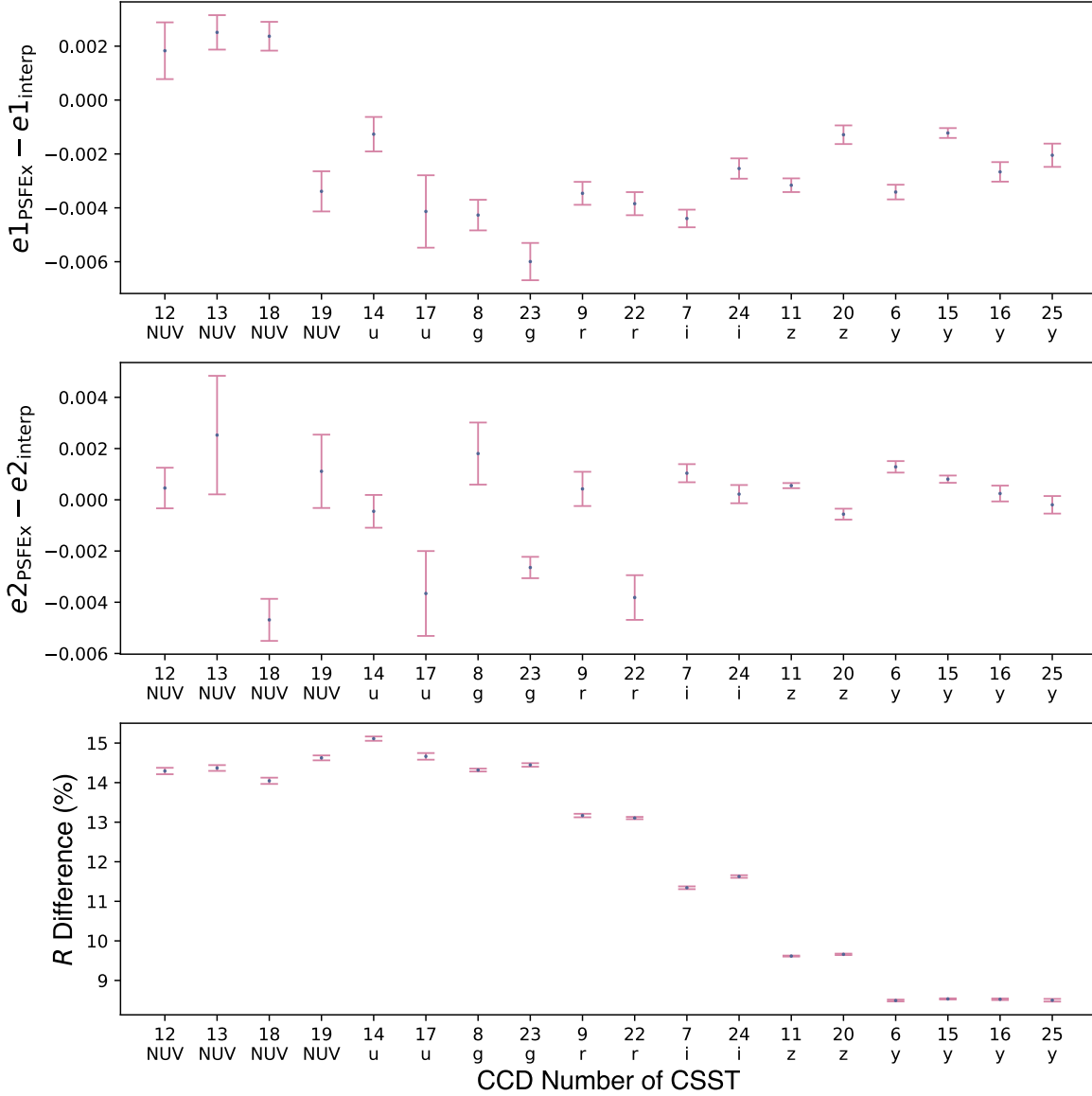


Figure 8. Mean differences in ellipticity components $e1$ and $e2$, and the percent difference in radius parameter R for each CCD of CSST. Error bars represent the 95% confidence intervals.

Section 3.2.1, but controlling the SNR through total flux. These images were then used to reconstruct the PSF models using PSFEX at a sampling rate of $2\times$. Subsequently, the ellipticity parameters ($e1$ and $e2$) and size (R) values of the PSF models were measured across the different SNR levels. We observed that the R values tend to decrease with increasing SNR. However, even at higher SNRs, the R values did not converge to those obtained from direct interpolation of the PSF models. Figure 9 illustrates examples of the variations in R values with SNR for CCD 17 and CCD 22.

To capture the trend of R values as a function of SNR, we applied a polynomial fit to the measured R values for each CCD, with these fits shown as the cyan lines in the top panels

of Figure 9. Using the polynomial parameters obtained from the fit, we corrected the measured R values. The steps for this correction were: For each measured R value, we calculated its deviation from the fitted polynomial value at the corresponding SNR. These deviations were then scaled according to the standard deviations of the target distribution (the directly interpolated PSF model values). The scaled deviations were added to the mean of the target distribution to obtain the corrected R values. This process ensures that the corrected R values align more closely with the target distribution, accounting for systematic biases observed in the original measurements. Additionally, for each measured R value, we calculated the ratio of its measured standard deviation to the

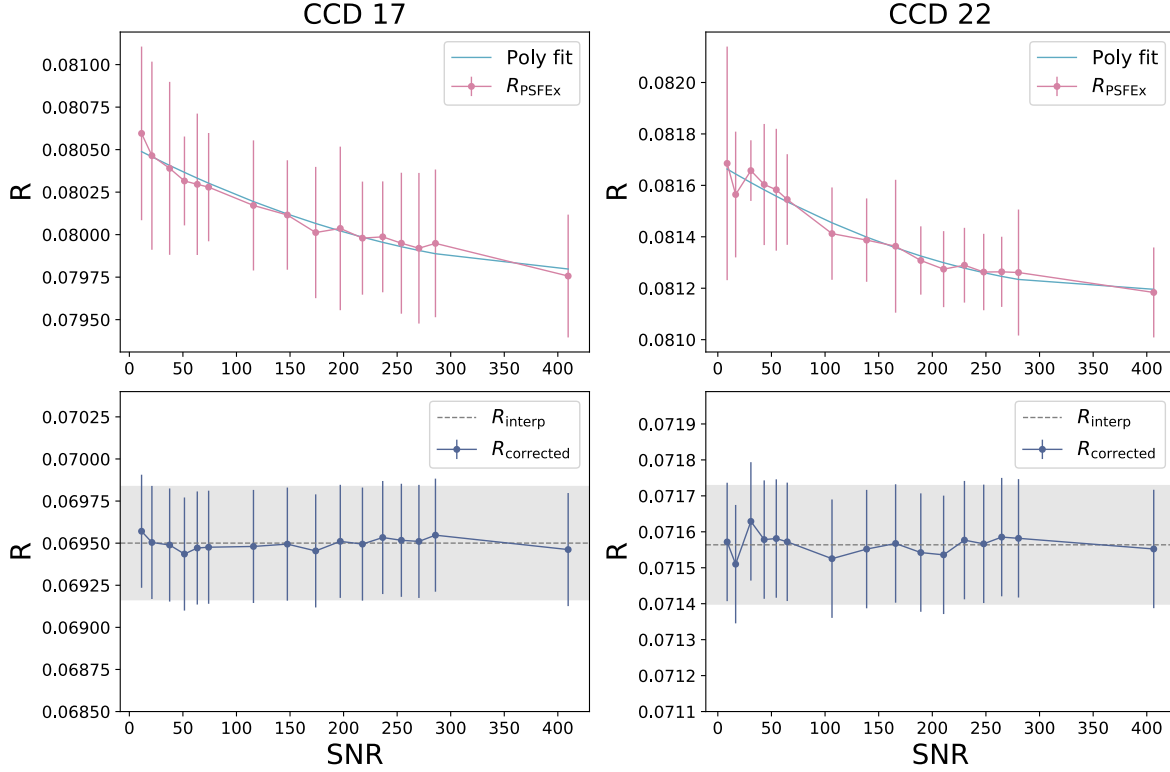


Figure 9. Variations in the radius parameters (R) of PSF models with SNR for CCD 17 and CCD 22 of the CSST. Top panels: The R values obtained from PSF models using PSFEX (pink line), with error bars indicating standard deviations. Polynomial fits are depicted with a cyan line across varying SNR levels. Bottom panels: The corrected R values are depicted by the navy line, with error bars indicating standard deviations. The gray dashed line represents R values from direct interpolation, and the gray shaded areas show the 1σ uncertainty range for these values.

target standard deviation. The measured standard deviations were then scaled to match the target standard deviation. This adjustment aligns the uncertainty of the corrected R values with the uncertainty of the target R values obtained through direct interpolation, thereby ensuring consistency in measurement precision. The corrected measured R values and their standard deviations are indicated by the navy points and error bars in the bottom panels of Figure 9. By combining these steps, we effectively correct both the R values and their associated standard deviations, improving the accuracy and reliability of the reconstructed PSF parameters. As shown in Figure 9, after correction, the measured R values and their standard deviations for different SNRs align well with the values obtained from direct interpolation, generally falling within one standard deviation.

4.2.2. Correction of Ellipticity

For the ellipticity parameters $e1$ and $e2$, we observed no discernible trend with varying SNRs across all CCDs. Instead, the distributions of these parameters remained concentrated around specific peak values. This behavior is illustrated in the top panels of Figure 10, where the pink histograms represent the distributions of $e1$ and $e2$ for CCD 22, as measured from

the PSF models reconstructed using PSFEX. In contrast, the gray histograms denote the distributions of $e1$ and $e2$ obtained from direct interpolation of the PSF models. As shown, there are noticeable deviations between the measured distributions and those derived from direct interpolation.

To correct the measured ellipticity parameters, we followed a systematic correction process. For each measured $e1$ and $e2$ value, we calculated its deviation from the mean of the measured distribution. These deviations were then scaled by the ratio of the standard deviation of the target distribution to that of the measured distribution. Finally, the scaled deviations were added to the mean of the target distribution to obtain the corrected $e1$ and $e2$ values. As shown in the bottom panels of Figure 10, the navy histograms represent the distributions of the corrected $e1$ and $e2$ values. It is evident that the corrected measurements align more closely with the values obtained from direct interpolation of the PSF models, significantly reducing the observed systematic biases. The Kolmogorov–Smirnov (KS) test further supports this improved alignment. Specifically, for $e1$, the KS statistic is 0.062 and the p -value is 0.105, while for $e2$, the KS statistic is 0.054 and the p -value is 0.209. The small KS statistics imply that the maximum deviation between the cumulative distribution functions of the corrected measurements and the directly interpolated values is

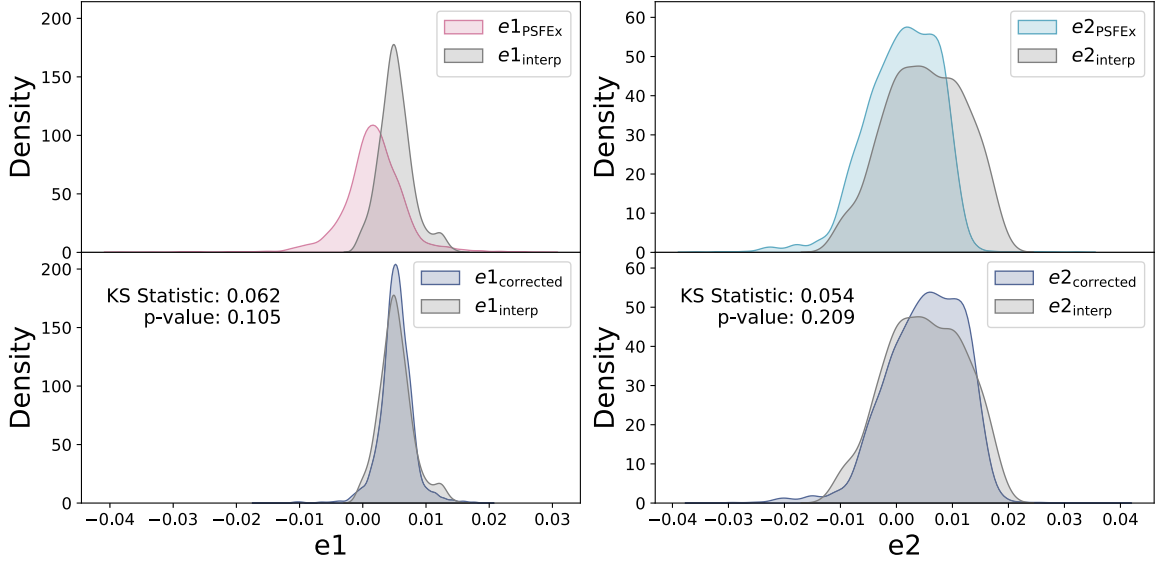


Figure 10. Ellipticity parameter distributions (e_1 , e_2) of PSF models for CCD 22 of CSST. Left panels: The top plot contrasts the e_1 distributions obtained through PSFEx (pink) with those derived from direct interpolation (gray), while the bottom plot displays the distribution after correction (navy) alongside the interpolated distribution (gray). Right panels: Similar analysis for e_2 , comparing the original PSFEx distribution (pink), the direct interpolation (gray), and the corrected distribution (navy). The corresponding KS statistic and p -value for both e_1 and e_2 are provided.

minimal. Moreover, the high p -values support the hypothesis that the corrected measurements are statistically similar to the directly interpolated values. These results indicate a strong alignment between the corrected and target distributions. In summary, the correction process effectively reduces the systematic biases inherent in the PSF reconstructed using the PSFEx method, leading to more accurate and reliable ellipticity parameter measurements.

4.3. A Combined Approach for Improving PSF Modeling Accuracy

The systematic biases in the parameter measurements of PSF reconstruction can be addressed through parameter correction methods, while the influence of unresolved binaries on PSF modeling requires careful source selection. For the latter, we applied deep learning techniques for source selection, with the process specifically designed to filter out unresolved binaries that could introduce significant distortions into the PSF models reconstructed. In this section, we will combine these two factors to evaluate the effectiveness of parameter correction and source selection in improving the accuracy of PSF model parameters, specifically the radius parameter and the ellipticity components.

To conduct this analysis, we used CCD12 of CSST as a representative example and generated a data set with an 80% binary fraction. We then analyzed the impact of different combinations of source selection and parameter correction on the PSF parameters. As shown in Figure 11, the distributions of the measured R values are presented for four different scenarios. The gray histograms represent the R values obtained

directly from interpolation of the original PSF matrix, serving as the true values for comparison. With no source selection and no parameter correction (top left), the distribution of R values shows the greatest deviation from the true values, indicating significant inaccuracies in the PSF reconstruction. When source selection is applied without parameter correction (top right), the distribution moves closer to the true values, suggesting that filtering out unresolved binaries helps to reduce some of the errors, but discrepancies still remain. Similarly, when parameter correction is applied without source selection (bottom left), the distribution also shifts closer to the true values, but a noticeable gap remains. The combination of source selection and parameter correction (bottom right) provides the most accurate R values, closely aligning with the reference values. The KS statistic and p -value further support this observation, showing minimal deviation between the corrected measurements and the reference values. Specifically, for the case with both source selection and parameter correction, the KS statistic is 0.060 and the p -value is 0.468, indicating strong alignment between the corrected and reference distributions.

Figures 12 and 13 present the distributions of the measured e_1 and e_2 values of PSF models for CCD 12 of CSST under various combinations of source selection and parameter correction scenarios, illustrating their impact on the ellipticity measurements compared to the true values (gray histograms). In both figures, the e_1 and e_2 distributions with no source selection and no parameter correction (top-left panels) exhibit the most significant deviations from the true values, similar to the measured R values. Applying only source selection (top-

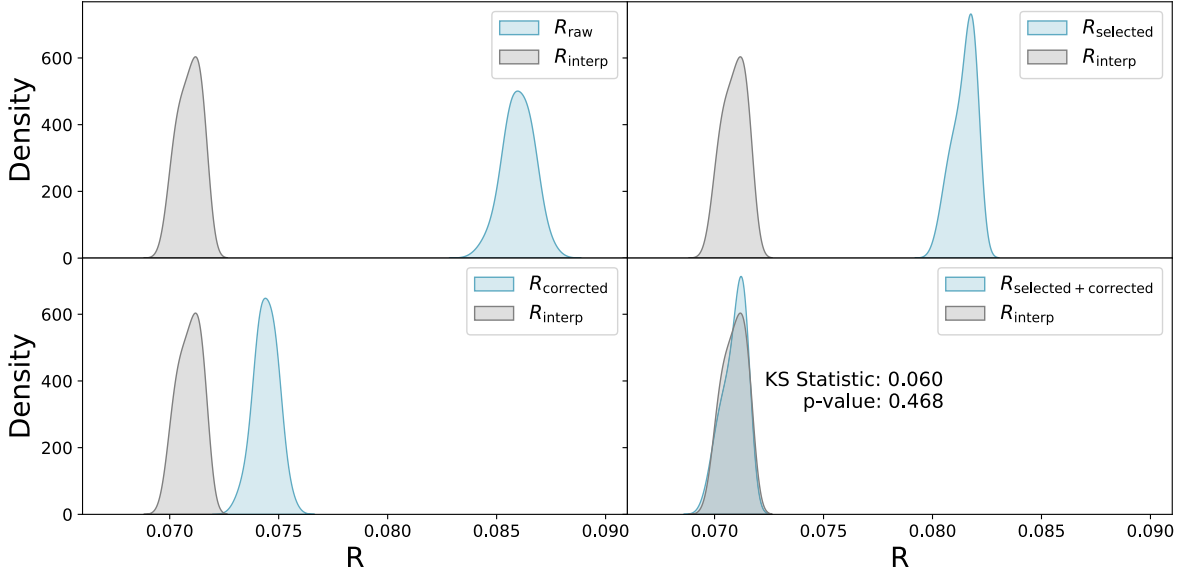


Figure 11. The distributions of measured R values of PSF models for CCD 12 of CSST under various combinations of source selection and correction methods. Panels represent cases of no selection and no correction (top left), selection only (top right), correction only (bottom left), and both selection and correction applied (bottom right). The gray distributions in all plots represent the R values obtained directly through interpolation of the original PSF matrix, serving as a reference for comparison.

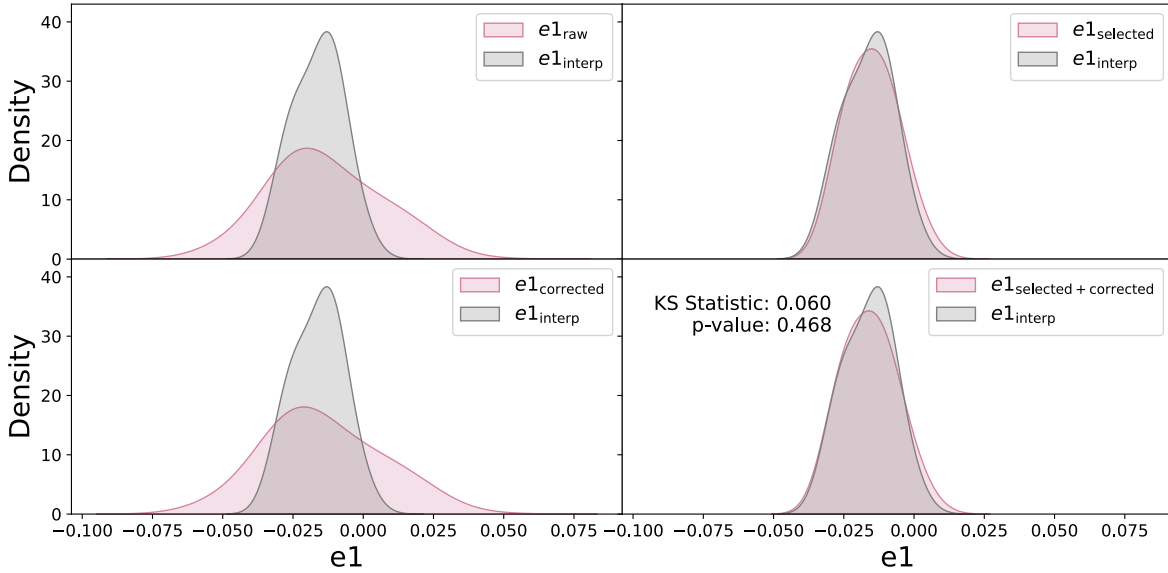


Figure 12. The distributions of measured e_1 values of PSF models for CCD 12 of CSST under various combinations of source selection and correction methods. Panels represent cases of no selection and no correction (top left), selection only (top right), correction only (bottom left), and both selection and correction applied (bottom right). Gray distributions show reference e_1 values obtained by interpolating the original PSF matrix.

right panels) brings the distributions closer to the true values, indicating that filtering out unresolved binaries significantly improves the accuracy of e_1 and e_2 measurements. The most accurate results for both e_1 and e_2 are achieved when source selection is combined with parameter correction (bottom-right panels). The KS-test results further support this observation, indicating minimal deviation between the corrected measurements and true values. The KS statistic for e_1 is 0.060 with a p -

value of 0.468, while for e_2 , the KS statistic is 0.065 with a p -value of 0.367.

These results demonstrate that the combined approach of source selection and parameter correction significantly enhances the accuracy of the reconstructed PSF parameters for CSST, particularly in terms of R , e_1 , and e_2 . By effectively filtering out unresolved binaries through source selection, we reduce one of the main sources of measurement bias, while

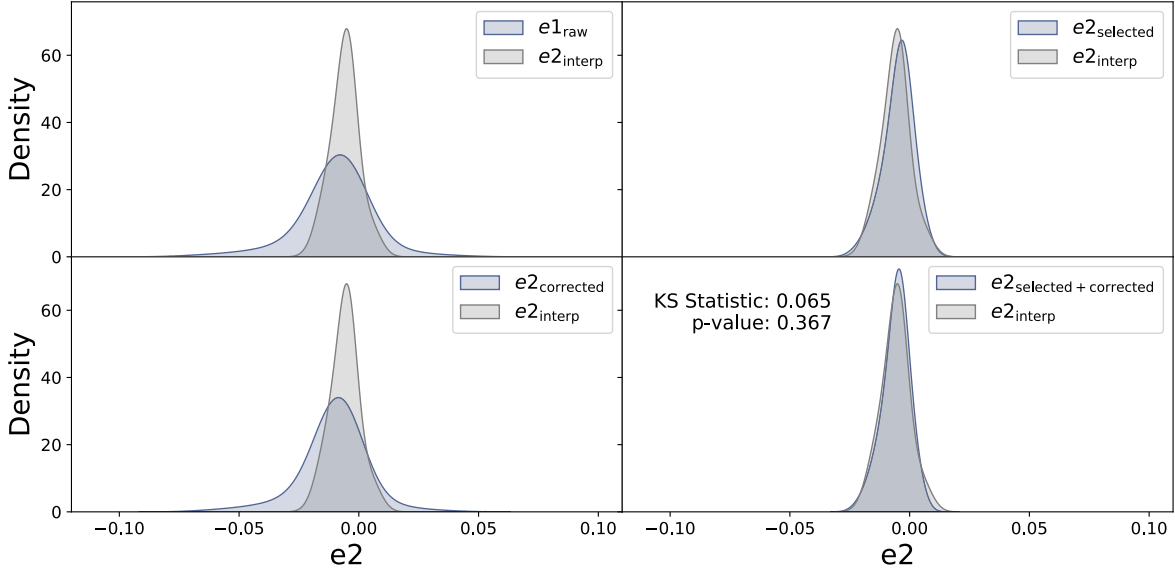


Figure 13. The distributions of measured e_2 values of PSF models for CCD 12 of CSST under various combinations of source selection and correction methods. Panels represent cases of no selection and no correction (top left), selection only (top right), correction only (bottom left), and both selection and correction applied (bottom right). Gray distributions show reference e_2 values obtained by interpolating the original PSF matrix.

parameter correction further improves the precision of the PSF shape parameters. The high p-values from the KS tests indicate that the distributions of the corrected PSF parameters are statistically consistent with the true values. This underscores the importance of integrating source selection and parameter correction to improve PSF modeling accuracy, particularly for complex data sets in large-scale surveys like CSST.

5. Discussion and Conclusion

In this study, we conducted a comprehensive analysis of PSF modeling for the upcoming CSST mission, focusing on addressing systematic errors and the contamination introduced by unresolved binaries in the PSF reconstruction process. We used the original CSST PSF matrix and followed the corresponding CSST simulation process to generate simulated CSST images. The impact of image size and sampling rate on PSF modeling was then evaluated with both the PSFEX and ePSF methods. The analysis showed that image size has a minimal effect on the accuracy of PSF modeling, as the PSF parameters (R , e_1 , and e_2) remain consistent across different image sizes. Across different sampling rates, both e_1 and e_2 exhibit deviations, while R tends to be slightly overestimated compared to the true value. However, changes in the sampling rate only introduce minor variations in the reconstructed PSF parameters, suggesting that systematic biases in the PSF reconstruction process play a more significant role in determining the accuracy of PSF modeling.

We conducted tests on all 18 CCDs and found that e_1 and e_2 exhibit systematic deviations that vary across different CCDs. In the shorter-wavelength bands (NUV and u), these deviations

tend to be larger, with more significant scatter and uncertainty. As the wavelength increases, the deviations gradually decrease in magnitude, and the uncertainties become smaller. Meanwhile, R is consistently overestimated across all CCDs, with the largest deviations up to 15% observed in the NUV and u bands. As the wavelength increases, the deviations gradually decrease, reaching the lowest value of approximately 8% in the y -band. This overall trend indicates more stable PSF reconstruction in longer-wavelength bands. To mitigate systematic biases, we applied empirical corrections to the measured PSF parameters across all 18 CCDs of CSST. For R , a polynomial fit was applied to capture its trend with varying SNRs, and corrections were made by adjusting the measured values and their standard deviations to match those obtained from direct interpolation. Similarly, for e_1 and e_2 , we corrected the measured values by scaling deviations relative to the target distributions, reducing systematic biases. The corrected parameters aligned closely with the values from direct interpolation, as confirmed by KS tests, which demonstrated improved consistency and reduced biases in the reconstructed PSF parameters.

Furthermore, we explored the impact of unresolved binaries on PSF modeling accuracy by examining how binary fraction and separation affect the reliability of the model. Our analysis revealed that as the fraction of unresolved binaries increases, both the ellipticity components (e_1 , e_2) and the radius parameter (R) of the PSF models exhibit significant variations, with larger binary fractions leading to greater distortions. Additionally, we found that the separation between binary stars plays a critical role, with larger separations introducing more pronounced distortions in the PSF models, particularly in

the radius parameter (R). These findings highlight the significant impact of contamination from unresolved binaries and underscore the importance of selecting clean, single-star sources to ensure accurate and reliable PSF reconstruction. We developed a neural network-based approach for source selection to enhance PSF reconstruction by filtering out unresolved binaries. Built on a modified *EfficientNet* architecture, the model achieved high performance on CSST simulated images, reaching an accuracy of 89.1% after training. The ROC curve, with an AUC score of 0.937, demonstrates the model’s effectiveness in distinguishing unresolved binaries from single stars, while the confusion matrix confirms its balanced classification performance. These results indicate that our method can reliably select clean single-star samples for PSF reconstruction.

Finally, we combined source selection and parameter correction to improve the accuracy of PSF model parameters for CSST. By applying deep learning-based source selection to filter out unresolved binaries, we effectively reduced the most significant measurement biases, while parameter correction further enhanced the precision of the reconstructed PSF parameters. This combined approach results in parameter distributions showing improved consistency with the true values, as confirmed by KS-tests with high p -values. By optimizing PSF modeling techniques for large-scale surveys such as CSST, this method helps minimize systematic errors in scientific studies that rely on precise PSF shape and size measurements. Accurate correction of PSF ellipticity parameters enhances the statistical precision of weak lensing surveys, leading to tighter constraints on cosmological models and deeper insights into the large-scale structure of the Universe. Meanwhile, improving the accuracy of the PSF radius parameter is crucial for photometric and spectroscopic analyses, ensuring more reliable flux and size measurements of stars and galaxies.

Several areas for future development remain. The deep learning-based source selection method can be further optimized to handle more complex observational environments. While the current model is trained specifically to distinguish single stars from unresolved binaries under the assumption of a clean input catalog, real CSST observations may still contain residual contaminants after preprocessing, potentially affecting source classification performance. To improve the model’s generalizability and robustness against such out-of-distribution samples, future iterations may benefit from incorporating a broader “non-stellar” class during training. Furthermore, enhancements to model design that account for observational factors such as sky background fluctuations and crowded stellar fields will be essential to ensure reliable PSF modeling in realistic survey conditions.

Another key challenge is the reliance on a known “ground truth” PSF distribution when correcting systematic biases. While the empirical correction method performs well in

simulations, such ideal reference PSFs are not available in real observations. A practical way to construct a reference PSF distribution is through a combination of ground-based optical testing and on-orbit calibration. During the ground-based testing phase, high-resolution PSFs can be obtained under controlled conditions across different bands and focal plane positions, providing an initial reference distribution for PSF parameters. After launch, the CSST will periodically perform dedicated PSF calibration observations of selected calibration fields using large dithered imaging strategies. By slightly shifting the telescope pointing during different exposures, stars are sampled at different sub-pixel positions across the detector, effectively mitigating the undersampling problem. High SNR single stars in these dithered exposures can then be used to reconstruct a high-resolution “effective PSF” directly from the data. Furthermore, calibration fields can also be observed with varying exposure times under the same pointing, enabling the reconstruction of ePSFs across a range of SNRs and providing direct calibration of SNR-dependent PSF variations. This data-driven ePSF model thus serves as a high-fidelity reference for calibrating PSF parameters from single exposures, correcting for undersampling, detector-induced distortions, and SNR-dependent effects. In this way, realistic end-to-end validation of the empirical correction pipeline becomes possible, reducing reliance on idealized simulations.

In addition to data-driven strategies, further optimization of the PSF correction model can benefit significantly from physics-informed modeling. Current corrections primarily rely on empirical adjustments to address systematic biases, but a deeper understanding of the underlying physical origins of PSF distortions would enhance their effectiveness. Instrumental effects—including optical distortions from thermal and mechanical variations, pixelization errors from detector undersampling, flux-dependent brighter-fatter effects, and charge transfer inefficiency during CCD readout—should be explicitly modeled in terms of observable parameters. By incorporating physically motivated relationships among source brightness, detector position, wavelength, and observation time, future correction models will be better equipped to accurately characterize and predict PSF variations across spatial, temporal, and observational conditions, thereby reducing reliance on empirical calibration and ensuring robust PSF accuracy over the course of extended CSST observations. This physics-informed approach, combined with data-driven calibration using dithered observations, has the potential to significantly enhance PSF reconstruction precision, ensuring more reliable scientific measurements from future CSST observations.

Beyond these methodological improvements, our future work will integrate the correction results into a generative PSF modeling framework, enabling the construction of corrected PSF images based on the initial PSF model and its measured parameter deviations. In parallel, we will systematically

evaluate the impact of PSF correction on shear estimation, thereby quantifying the resulting improvement in weak lensing measurement precision. Moreover, the techniques developed here are expected to be applicable beyond CSST, offering a pathway for other large-scale surveys that demand precise PSF modeling for weak lensing and related cosmological analyses.

Acknowledgments

This work was supported by the science research grants from the China Manned Space Project (Nos. CMS-CSST-2021-A01 and CMS-CSST-2025-A19), the National Key R&D Program of China (No. 2021YFA1600400/1), the CAS Project for Young Scientists in Basic Research (No. YSBR-062), and the Ministry of Science and Technology of China (No. 2020SKA0110100).

ORCID iDs

Peng Wei  <https://orcid.org/0000-0003-2477-6092>

Cheng-Liang Wei  <https://orcid.org/0000-0001-5912-7522>

Juan-Juan Ren  <https://orcid.org/0000-0003-3243-464X>

References

- Anderson, J. 2016, Empirical Models for the WFC3/IR PSF Instrument Science Report WFC3 2016-12, Space Telescope Science Institute [42](#)
- Anderson, J., & King, I. R. 2000, *PASP*, **112**, [1360](#)
- Beltramo-Martin, O., Ragland, S., Fétick, R., et al. 2020, *Proc. SPIE*, **11448**, [114480A](#)
- Bernstein, G., & Jarvis, M. 2002, *AJ*, **123**, [583](#)
- Bertero, M., & Boccacci, P. 2005, *A&A*, **437**, [369](#)
- Bertin, E. 2011, in ASP Conf. Proc. 442, Astronomical Data Analysis Software and Systems XX, Vol. 442, ed. I. N. Evans et al. (San Francisco, CA: ASP), [435](#)
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, [393](#)
- Born, M., & Wolf, E. 2013, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light (Amsterdam: Elsevier)
- Bradley, L., Sipocz, B., Robitaille, T., et al., 2016 Photutils: Photometry tools, Astrophysics Source Code Library, record, [ascl:1609.011](#)
- Braat, J. J. M., van Haver, S., Jansse, A. J. E. M., & Dirksen, P. 2008, *PrOpt*, **51**, [349](#)
- Cao, Y., Gong, Y., Meng, X.-M., et al. 2018, *MNRAS*, **480**, [2178](#)
- Erben, T., Van Waerbeke, L., Bertin, E., Mellier, Y., & Schneider, P. 2001, *A&A*, **366**, [717](#)
- Guinot, A., Kilbinger, M., Farrens, S., et al. 2022, *A&A*, **666**, [A162](#)
- Hanisch, R. J., White, R. L., & Gilliland, R. L. 1997, Deconvolution of hubble space telescope images and spectra, Deconvolution of Images and Spectra, Vol. 2 (Boston, MA: Academic Press)
- Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, *MNRAS*, **421**, [2355](#)
- Hirata, C., & Seljak, U. 2003, *MNRAS*, **343**, [459](#)
- Hoffmann, S. L., & Anderson, J. 2018, AAS Meeting, **231**, [150.36](#)
- Hu, Y., Deng, L., de Grijs, R., & Liu, Q. 2011, *PASP*, **123**, [107](#)
- Jarvis, M., Bernstein, G., Amon, A., et al. 2021, *MNRAS*, **501**, [1282](#)
- Jee, M., Blakeslee, J., Sirianni, M., et al. 2007, *PASP*, **119**, [1403](#)
- Jee, M. J., & Tyson, J. A. 2011, *PASP*, **123**, [596](#)
- Kaiser, N., Squires, G., & Broadhurst, T. 1995, *ApJ*, **449**, [460](#)
- Krist, J. 1993, in ASP Conf. Ser. 52, Astronomical Data Analysis Software and Systems II, Vol. 52, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (ASP), [536](#)
- Krist, J. 1995, in ASP Conf. Ser. 77, Astronomical Data Analysis Software and Systems IV, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes, [349](#)
- Krist, J. E., Hook, R. N., & Stoehr, F. 2011, *Proc. SPIE*, **8127**, [81270J](#)
- Liaudat, T., Bonnin, J., Starck, J.-L., et al. 2021, *A&A*, **646**, [A27](#)
- Liaudat, T. I., Starck, J.-L., & Kilbinger, M. 2023, *FrASS*, **10**, [1158213](#)
- Libralato, M., Argyriou, I., Dicken, D., et al. 2024, *PASP*, **136**, [034502](#)
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, [A2](#)
- Lupton, R., Gunn, J. E., Ivezić, Z., Knapp, G. R., & Kent, S. 2001, in ASP Conf. Proc. 238, Astronomical Data Analysis Software and Systems X, ed. F. R. Jr. Harnden, F. A. Primini, & H. E. Payne (San Francisco, CA: ASP), [269](#)
- Mandelbaum, R. 2018, *ARA&A*, **56**, [393](#)
- Mandelbaum, R., Hirata, C. M., Seljak, U., et al. 2005, *MNRAS*, **361**, [1287](#)
- Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, *PASJ*, **70**, [S25](#)
- Mighell, K. J. 2005, *MNRAS*, **361**, [861](#)
- Ngolè, F., Starck, J.-L., Okumura, K., Amiaux, J., & Hudelot, P. 2016, *InvPr*, **32**, [124001](#)
- Rhodes, J., Refregier, A., & Groth, E. J. 2000, *ApJ*, **536**, [79](#)
- Rowe, B. T., Jarvis, M., Mandelbaum, R., et al. 2015, *A&C*, **10**, [121](#)
- Schmitz, M. A. 2019, PhD thesis, Université Paris Saclay (COMUE), <https://www.theses.fr/2019SACL3539>
- Stetson, P. B. 1987, *PASP*, **99**, [191](#)
- Tan, M. 2019, in Proc. of the 36th Int. Conf. on Machine Learning, **97**, (PMLR)
- Wu, Y., Li, J., Liu, C., et al. 2023, *ApJS*, **268**, [37](#)
- Zhan, H. 2011, *SSPMA*, **41**, [1441](#)
- Zuntz, J., Sheldon, E., Samuroff, S., et al. 2018, *MNRAS*, **481**, [1149](#)