GEMS: GALAXY FITTING CATALOGS AND TESTING PARAMETRIC GALAXY FITTING CODES: GALFIT AND GIM2D

Boris Häussler,[1] Daniel H. McIntosh,[2] Marco Barden,[1] Eric F. Bell,[1] Hans-Walter Rix,[1] Andrea Borch,[3]
Steven V. W. Beckwith,[4,5] John A. R. Caldwell,[6] Catherine Heymans,[7] Knud Jahnke,[1] Shardha Jogee,[8]
Sergey E. Koposov,[1] Klaus Meisenheimer,[1] Sebastian F. Sánchez,[9] Rachel S. Somerville,[1]
Lutz Wisotzki,[10] and Christian Wolf[11]

ABSTRACT

In the context of measuring the structures of intermediate-redshift galaxies with *HST* ACS surveys, we tune, test, and compare two widely used fitting codes (GALFIT and GIM2D) for fitting single-component Sérsic models to both simulated and real galaxy data. Our study focuses on the GEMS survey with the sensitivity of typical *HST* survey data, and we include our final catalog of fit results for all 41,495 objects detected in GEMS. Using simulations, we find that fitting accuracy depends sensitively on galaxy profile shape. Exponential disks are well fit and have small measurement errors, whereas fits to de Vaucouleurs profiles show larger uncertainties owing to the large amount of light at large radii. Both codes provide reliable fits with little systematic error for galaxies with effective surface brightnesses brighter than that of the sky; the formal uncertainties returned by these codes significantly underestimate the true uncertainties (as estimated using the simulations). We find that GIM2D suffers significant systematic errors for spheroids with close companions owing to the difficulty of effectively masking out neighboring galaxy light; there appears to be no work-around to this important systematic in GIM2D's current implementation. While this crowding error affects only a small fraction of galaxies in GEMS, it must be accounted for in the analysis of deeper cosmological images or of more crowded fields with GIM2D. In contrast, GALFIT results are robust to the presence of neighbors because it can simultaneously fit the profiles of multiple companions as well as the galaxy of interest. We find GALFIT's robustness to nearby companions and factor of $\gtrsim 20$ faster runtime speed are important advantages over GIM2D for analyzing large *HST* ACS data sets.

Subject headings: catalogs — galaxies: general — galaxies: photometry — galaxies: statistics — methods: data analysis — surveys

Online material: machine-readable table

1. INTRODUCTION

One of the central goals of observational exploration of galaxy evolution is to understand how the structures of galaxies evolve with cosmic time. A powerful tool in this context are large look-back surveys, in which the time evolution of the distribution of galaxy structural properties can be quantified. The key to unlocking the potential of these surveys is the development of quantifiable, well-understood, and repeatable ways to measure and describe galaxy structures. Using such methods, the evolution of the structure of disk galaxies (Lilly et al. 1998; Simard & Pritchet

1998; Ravindranath et al. 2004; Barden et al. 2005; Trujillo & Pohlen 2005; Sargent et al. 2007) and spheroid-dominated galaxies (e.g., Schade et al. 1997, 1999; McIntosh et al. 2005; Trujillo et al. 2004, 2006) has been quantified over the last 10 Gyr of cosmic time, since $z = 3$. In this paper, we exhaustively test and tune two parametric galaxy fitting codes, GALFIT (Peng et al. 2002) and GIM2D (Simard 2002), that are commonly used in the literature. With these tests we determine the best-fitting setups for each code, quantify the sources of random and systematic uncertainty, and present parametric fits for 41,495 objects in the *Hubble Space Telescope* (*HST*) Galaxy Evolution from Morphology and SEDs (GEMS) (Rix et al. 2004) data set.

There are two main approaches toward describing galaxy structure from the two-dimensional information contained in image data. Nonparametric methods provide estimates of total brightness, galaxy half-light size, and structure, using metrics which do not depend on a galaxy having a structure well-described by any particular functional form (e.g., Petrosian 1976; Abraham et al. 1996; Bershady et al. 2000; Lotz et al. 2006). The main disadvantage of nonparametric methods is that they are reasonably sensitive to the depth of the images; because there is no parametric form for extrapolating to account for the faint outer parts of galaxies, one can underestimate flux and/or size in poorly posed cases (Blanton et al. 2003a). Parametric methods, in contrast, choose particular functional forms (sometimes reasonably complicated) with which to fit the galaxy light distribution. These have substantially less flexibility than nonparametric fitting codes but have the advantage that light at large radii can be accounted for

---

[1] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany.

[2] Department of Astronomy, University of Massachusetts, 710 North Pleasant Street, Amherst, MA 01003.

[3] Astronomisches Recheninstitut, Mönchhofstrasse 12–14, 69120 Heidelberg, Germany.

[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218.

[5] Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218.

[6] University of Texas, McDonald Observatory, Fort Davis, TX 79734.

[7] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada.

[8] University of Texas at Austin, 1 University Station, C1400, Austin, TX 78712-0259.

[9] Centro Astronomico Hispano Aleman de Calar Alto, C/Jesus Durban Remon 2-2, Almeria E-04004, Spain.

[10] Universität Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany.

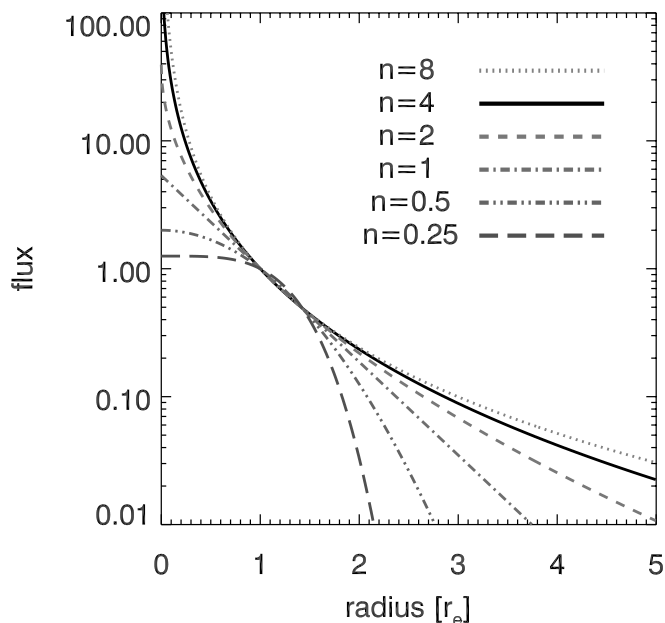[11] Department of Physics, Denys Wilkinson Building, University of Oxford, Keble Road, Oxford OX1 3RH, UK.

Fig. 1.—Sérsic profiles for different values of the Sérsic index n, normalized to have the same flux at r_e. One can see that profiles with high Sérsic indices $n \gtrsim 2$ have more flux at larger radii; thus, a good estimate of the background sky level is particularly important for precise fitting of high-n galaxies.

reasonably well by the natural extrapolation of the best-fitting model profile (under the assumption that the parametric form chosen does, in fact, describe the light profile in the outer parts of galaxies reasonably well). Besides robust estimates of galaxy size, parametric methods provide measures of galaxy structure that may shed light on relative contributions of physically distinct and meaningful components such as spheroids, disks, and stellar bars.

One particularly useful and flexible profile for parametric galaxy fitting is a single-component Sérsic (1968) model, which describes the radial surface brightness profile of a galaxy by the Sérsic function given by

$$\Sigma(r) = \Sigma_e \exp\left\{-\kappa\left[(r/r_e)^{1/n} - 1\right]\right\}, \qquad (1)$$

where r_e is the radius of the galaxy (note that for a Sérsic fit r_e is equivalent to the half-light radius r_{50}), Σ_e is the surface brightness at r_e, and the Sérsic parameter n describes the profile shape (the parameter κ is closely connected to n). Together with position (x and y), axis ratio b/a, and position angle, this profile has seven free parameters. The Sérsic profile represents a more general form of the exponential light profiles seen in galactic disks ($n = 1$) and the $R^{1/4}$-law (de Vaucouleurs law) profiles typical of luminous early-type galaxies ($n = 4$) (e.g., de Vaucouleurs 1948; Freeman 1970); fitting with this profile has been explored in detail in several works (e.g., Simard 1998, 2002; Graham et al. 2005; Trujillo et al. 2001). Figure 1 shows some examples of Sérsic profiles with different n. Many authors have used a constant value of $n = 2.5$ or 2.0 to crudely distinguish early-type (bulge-dominated) from late-type (disk-dominated) galaxies (e.g., Blanton et al. 2003b; Shen et al. 2003; Hogg et al. 2004; Bell et al. 2004; Barden et al. 2005; McIntosh et al. 2005; Ravindranath et al. 2004). Furthermore, fitting galaxies with a Sérsic profile gives an estimate of size and therefore is very useful for the examination of the evolution of galaxy scaling relations.

The goal of this paper is to describe our efforts to optimize the estimation of single-component Sérsic profile fits to the galaxies

in the GEMS survey (Rix et al. 2004). To date, this has been our primary method for quantifying galaxy structure.[12] We compare the performance of the GALFIT and GIM2D automated galaxy fitting codes, which are suitable for fitting large data sets such as GEMS, STAGES (M. Gray et al. 2007, in preparation), the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004), and COSMOS (Scoville et al. 2007). We address the reliability and limitations of these codes through thorough testing, using simulated and real galaxies. We describe the details of the simulations used throughout this paper in § 2. In § 3 we explore different setups, converging on "best-fitting" setups for each fitting code. Section 4 summarizes the results from our testing of these best-fitting setups using both simulated (§§ 4.1 and 4.2) and real galaxies (§ 4.3). We compare our findings with those of a recent paper on the same topic (using the codes GALFIT, GIM2D, and GASPHOT) by Pignatelli et al. (2006) in § 5 and publish a catalog of GALFIT fitting results for all 41,495 detected galaxies from the GEMS survey in § 6. All results from this paper and a number of other catalogs and images useful for testing galaxy fitting codes are presented on the GEMS Web site.[13]

2. SIMULATIONS

Galaxy simulations are an invaluable tool for understanding the performance of quantitative fitting pipelines. In this section we describe the set of simulations that are extensively used for this paper; the results obtained from fitting these simulations are discussed in § 4.

In this paper, we focus on simulations of two different galaxy light profiles: purely exponential profiles ($n = 1$) representing the luminosity profile of a galactic disk (we will call these galaxies "disks" throughout this paper) and Sérsic profiles with a Sérsic index of $n = 4$ representing a de Vaucouleurs luminosity profile of a galactic bulge/elliptical galaxy (we will call them "spheroids"). Profiles having Sérsic indices between these two values of 1 and 4 are not presented here because $n = 1$ and $n = 4$ simulations span the range of observed behavior: exponential profiles being the "easiest" to fit, de Vaucouleurs being the "hardest." Nonetheless, extensive simulations of intermediate profiles (200,000 objects) have been produced; the catalog of GALFIT fitting results for this sample can also be downloaded from the GEMS Web site.

This section is arranged as follows. The simulation of individual galaxies is described in § 2.1. Section 2.2 describes the construction of simulated GEMS frames from the individual galaxy simulations, including the addition of realistic noise.

2.1. Simulation of Individual Noise-free Galaxies and Oversampling

Galaxies were simulated using a custom-built IDL routine. Most available standard routines (like MKOBJECT in IRAF, create/image in MIDAS, and similar tasks in other programs) compute the correct flux value for the *center* of the pixel, but due to curvature of the profile, taking this as the mean flux value for the *whole* pixel is incorrect. The higher the curvature is (within a certain pixel), the more one underestimates the true pixel value. This implies progressively larger inaccuracies for higher Sérsic indices.

While it is possible to analytically integrate the profile across a pixel to obtain an exact answer, this procedure is very CPU-intensive. We adopt a hybrid approach. We use IDL's dist_ellipse routine in conjunction with equation (1) to compute

---

[12] Bulge-disk composite galaxies were *not* simulated for this paper; bulge-disk decomposition will be addressed in a future paper.

[13] See http://www.mpia.de/GEMS/gems.htm.

TABLE 1
SIMULATION PARAMETERS: DISK ($n = 1$) GALAXIES

Parameter	Min	Max	Distribution
mag (mag)	20	26.5	Uniform
r_e (pixel)	2	316	Uniform in logarithmic space; $r_e < 10^{7.36 - 0.233 \text{mag}}$, mag being chosen magnitude for object
b/a	0.18	1	Uniform in $\cos(i)$, i being inclination angle corrected for intrinsic thickness: $b/a = \{\cos^2(i) + [0.18 \sin(i)]^2\}^{1/2}$; intrinsic thickness 0.18 following Pizagno et al. (2005); Ryden (2006), and others
PA (deg to image)	0	180	Uniform
Sérsic index n	1	1	Fixed

Sérsic galaxy models which, as in the above cases, are only correct for the center of the pixel. In order to increase accuracy, the inner parts of our simulated profiles (100×100 pixels up to 200×200 pixels, depending on object size) have been oversampled by a factor of 10, and the very inner parts (10×10 up to 20×20 pixels) are oversampled by a factor of 100. This was done by creating the images by a factor of 10 (or 100, respectively) bigger and then rebinning the image while holding the total flux constant. In this way, it is possible to create a final profile accurate to better than 0.03% at all radii (much smaller than the Poisson noise added later in the simulation process) with a factor of 100 gain in speed compared to the analytical integration—an important gain when simulating large samples of galaxies.

2.2. Simulation of Crowded Images

To realistically test galaxy extraction and fitting codes requires the creation of images with large numbers of simulated galaxies distributed as in real data. Such images were created by providing a catalog of simulated galaxy input parameters to the simulation code, which simulated galaxies at the location, luminosity, size, orientation, and axis ratio b/a specified in this catalog. In this step, galaxies were put in an empty image of the same size as the final image.

To choose the range of galaxy parameters for the simulated galaxies, we first fitted all GEMS galaxies with GALFIT and determined the parameter range covered by the real galaxy sample. Given these results, we chose a wider range of parameter space for the simulations, in order to test detection efficiency and completeness, and to push both parametric fitting codes to their limits. The simulations have a random distribution in size between 2 and 316 pixels (uniformly distributed in logarithmic space) and magnitude between 20 and 27 (uniform). With this distribution of parameters, there were a relatively large number of large and low surface brightness galaxies (stringently testing the detection efficiency and fitting codes); we discuss this point in more detail in § 4.2. The exact distributions of simulation parameters are given in Tables 1 and 2.

After simulating the galaxy profiles and putting them in an empty image, this final image was convolved with a real F850LP-band point-spread function (PSF) derived from the GEMS data set (Jahnke et al. 2004). Next, an appropriate amount of noise

had to be added to the images. Owing to the multiple-frame dither characteristic of *HST* imaging surveys, the noise is somewhat correlated pixel-to-pixel. Thus, strictly speaking, galaxies should be simulated in individual dithers and then dithered together using exactly the same routines as were used to combine the GEMS frames. We took an intermediate approach: Poisson noise with the same rms as the GEMS noise was added to the simulated galaxy frame; then a real "sky" frame was added to the simulated frame to accurately account for real fluctuations and correlated noise in observed *HST* sky backgrounds. We have confirmed through tests with GALFIT that this (much less CPU- and work-intensive) hybrid approach yields a scatter which is negligible compared to random fitting uncertainties. The sky frame was constructed by adding GEMS F606W- and F850LP-band frames (to increase image depth and to make sure that objects appear in neither of the two bands) and visually checking those images to identify patches of 500×500 pixels in size without objects detected by SExtractor. The chosen patches were cut from the F850LP-band images and pasted together to form an empty image of the same size as an original GEMS image.

The result of the simulation process was a simulated image with noise properties very similar to a real GEMS image that contained 800 simulated light profiles with different magnitudes, sizes, position angles, and values of b/a. Different sets of simulations were created in this way: one set contains disklike $n = 1$ galaxies only (for results see § 4.1.1), and the other contains spheroidal $n = 4$ profiles only (see § 4.1.2). Not all 800 galaxies were recovered by SExtractor. Roughly 80% of the objects were recovered, depending on the simulated profile shape and the distribution of galaxy parameters in the particular image (see Fig. 2). In particular, very large and low surface brightness galaxies were not detected (see Fig. 2 and Rix et al. 2004). Due to the fact that spheroids are easier to detect due to their centrally concentrated light profiles, fewer galaxies were recovered in the disk sample.

By design, these simulated tiles are artificial in two ways. First, the galaxy input parameters span a wider range in parameter space than real galaxies. Second, the simulated tiles are significantly more crowded than the actual data itself, about a factor of 7–8 overdense in galaxies with Sérsic index $n > 2.5$ compared to a typical Advanced Camera for Surveys (ACS) image from GEMS (see Fig. 3). They contain many more low surface

TABLE 2
SIMULATION PARAMETERS: SPHEROIDAL ($n = 4$) GALAXIES

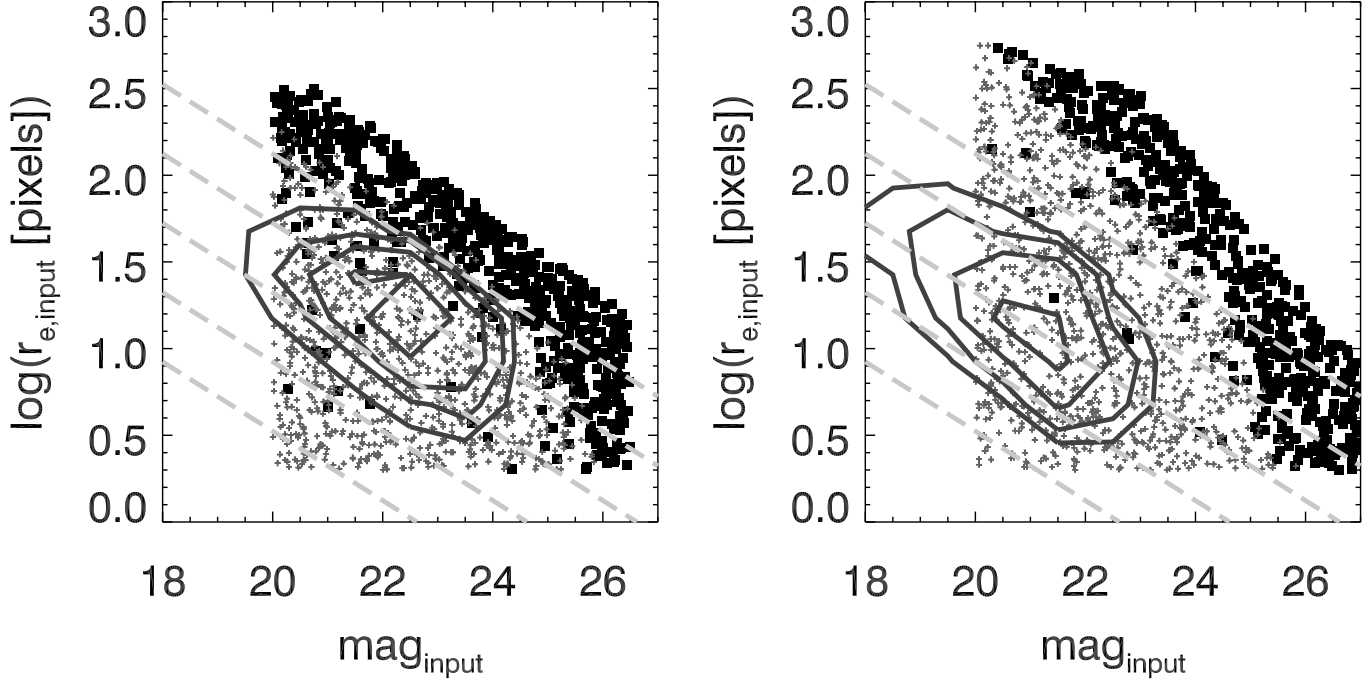| Parameter	Min	Max	Distribution
mag (mag)	20	27	Uniform
r_e (pixels)	2	630	Uniform in logarithmic space; $r_e < 10^{4.79 - 0.1 \text{mag}}$; $r_e < 10^{11.49 - 0.392 \text{mag}}$
b/a	0.45	1	Uniform in $\cos(i)$, i being inclination angle; corrected for intrinsic thickness: $b/a = \{\cos^2(i) + [0.45 \sin(i)]^2\}^{1/2}$
P.A. (deg to image)	0	180	Uniform
Sérsic index n	4	4	Fixed

Fig. 2.—F850LP magnitudes and sizes for the full set (*all symbols*) of 1600 simulated $n = 1$ galaxies (*left*) and $n = 4$ galaxies (*right*), and the subsets that were detected by SExtractor (*small gray plus signs*). The black squares indicate galaxies that were missed during object detection. The indicated contours show the magnitude-size space populated by actual GEMS galaxies used for disk galaxies by Barden et al. (2005, *left*) and for spheroidal galaxies by McIntosh et al. (2005, *right*); the contours show the areas of parameter space where the reliability of the fitting routines becomes especially important. Whereas real $n \geq 2.5$ galaxies lie in the area where all galaxies are detected, we did use $n \leq 2.5$ galaxies that are close to the edge of detectability for our analysis. The different behavior of the nondetected galaxies in both samples reflects the fact that, due to their bright central peak, galaxies with a high Sérsic index are easier to detect than galaxies with low n. To guide the eye, we plot long-dashed lines of constant surface brightness of 17, 19, 21, 23 and 25 mag arcsec^{-2} from bottom to top.
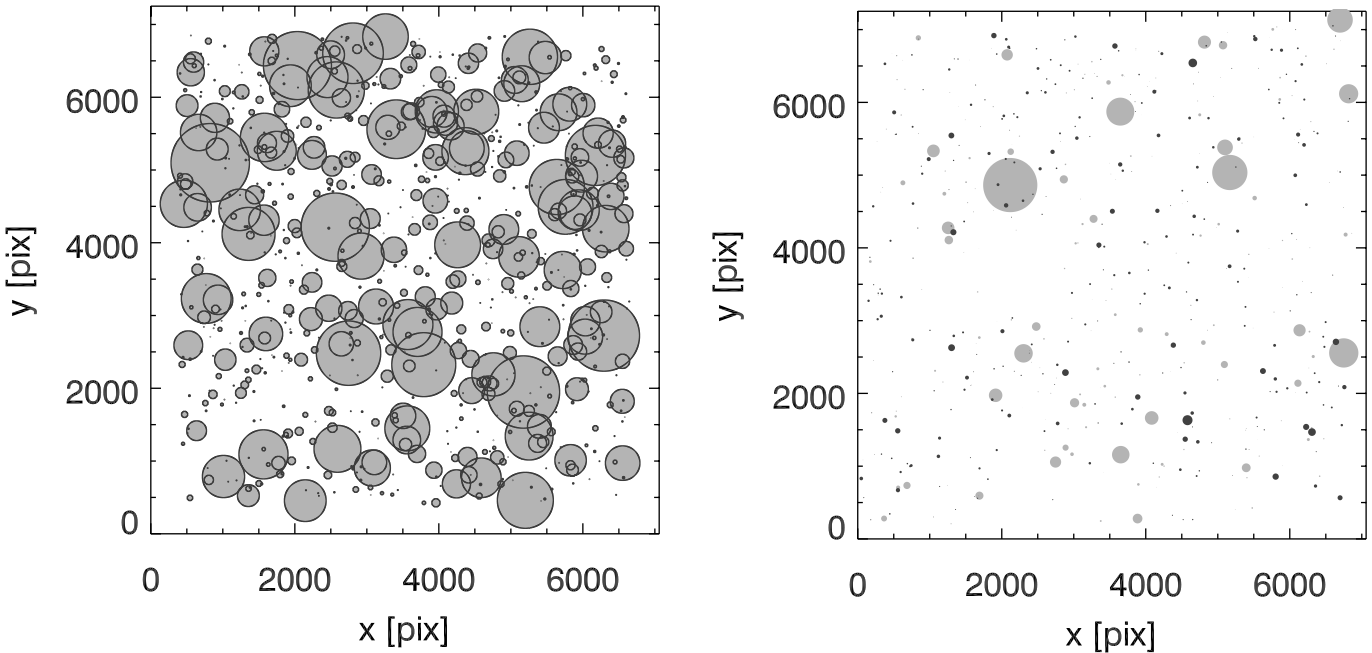

Fig. 3.—Density of sources in simulation images. *Left*: One of the two simulated images with galaxies having Sérsic index of 4. We plot a circle with radius r_e (simulated) at the correct (x, y) location of each simulated galaxy. In total there are 800 simulated $n = 4$ galaxies placed in a single ACS tile, ~250 of which have too low surface brightness to be detected by SExtractor; they could nonetheless influence the fitting results by contributing to the image background. This simulation represents an extreme case for testing the limits of profile fitting with GIM2D and GALFIT. *Right*: Sources in a simulation of a typical GEMS ACS image (GEMS tile 04) using GALFIT fitting results. There are 523 total simulated galaxies, 374 with $n = 1$ (*dark gray*), and 148 with $n = 4$ (*light gray*). Stars and the few objects (in total, 46 objects) that ran into a fitting constraint were excluded from this simulation. One can easily see from this plot that real galaxies are significantly less crowded than the completely artificial simulations used in §§ 4.1.1 and 4.1.2.

brightness galaxies (detected and undetected), adding a complex layer of extra flux to the background. This makes the simulations more difficult to analyze than real data; this was intentional since we wanted to push both codes to their limits. In a third set of simulations we mixed the two types of profile (see § 4.2) to estimate the effects of deblending given a more realistic mix of $n = 1$ and $n = 4$ galaxies.

To test and compare the two different 2D-fitting routines, the simulated images were treated as "real" images, i.e., we used exactly the same data pipeline for fitting that was used for the real GEMS data analysis. Therefore, all effects which we can see in the results from simulations should be present in real data as well, although mixed with many other effects like bulge/disk composite profiles, nonsmoothness, lumpiness, and/or spiral features of real galaxies.

3. GALAXY FITTING: DESCRIPTION, BASIC CONSIDERATIONS, AND BEST-FITTING SETUPS

For the GEMS analysis, we have used two widely employed parametric fitting codes for quantitatively describing galaxy structure and morphology: GALFIT and GIM2D. In this section, we describe both codes and the procedures used to parametrically fit both the real GEMS data and the simulations described in § 2. The basic considerations for code setup and application to real data, and the tests which we have performed on simulated data, are useful in general to other workers in galaxy image fitting. These basic considerations for setup and application of these (and most) fitting codes are (1) sky estimation, (2) initial parameter guesses, (3) postage stamp construction, and (4) deblending and/or masking of neighboring sources. We describe in detail the setups and various tests we carried out in order to optimize these setups in § 3.1 for GALFIT and § 3.2 for GIM2D.

The initial conditions and setups for both GALFIT and GIM2D are determined using SExtractor output images and catalogs. We use SExtractor (ver. 2.2.2; Bertin & Arnouts 1996) for image parsing and catalog creation. SExtractor detects, deblends, measures, and classifies objects, giving estimates of magnitude, size, b/a, position angle, and a star-galaxy classification. In GEMS, we found that no single SExtractor setup satisfactorily detected and deblended both bright, well-resolved galaxies and faint galaxies near the detection limit. Accordingly, our best setup is to run SExtractor twice: once to detect the bright objects without splitting them up (what we call the "cold" version) and once to detect the faint objects ("hot" version). The two versions are then combined to give one single catalog containing all objects. The procedure is described in more detail in Rix et al. (2004) and Caldwell et al. (2007). We do not use the SExtractor output catalogs directly for science; instead, these values are used as initial estimates for galaxy fitting codes and their setup. In §§ 3.1 and 3.2 we will describe which parameters are taken as starting guesses and how these values are used for the two parametric galaxy fitting codes used in this work: GALFIT and GIM2D.

3.1. *GALFIT*

GALFIT is a two-dimensional galaxy-fitting software package written by Peng et al. (2002). We used GALFIT version 2.0.3b from 2005 February 2 for this analysis. GALFIT was designed to extract structural components from galaxy images. Compared to other fitting techniques it has two main advantages. It uses a Levenberg-Marquardt downhill-gradient (Press et al. 1997) method to derive the best fit and therefore is relatively fast, being able to fit roughly 3000 galaxies per day on a dual 2.4 GHz Linux processor (when running four threads simultaneously to efficiently use all the CPU time). Furthermore, due to its speed and design,

it is able to fit an image containing an arbitrary number of galaxies simultaneously, making it possible to fit neighboring objects. The main disadvantage of GALFIT, in theory, is that it is possible that it converges on fit solutions that represent a local minimum instead of giving the global minimum. Our experience with GALFIT is that in single-component, but multiobject, fits this happens relatively rarely, if at all, both through the simulations (§ 4.1) and through comparison of fitting results for real galaxies from GALFIT and GIM2D (§ 4.3).

During the fitting process, the model is convolved with a specified PSF to model the image seeing and then compared to the input image. It is possible to fit the background sky level during the fitting process, although in this paper we use this capability for testing purposes only (see § 3.1.2).

In § 3.1.1, we will explain the basic setup procedure of GALFIT in detail, e.g., cutting postage stamps, estimating the sky background, deciding on how galaxies should be deblended, and setting up the initial parameters for GALFIT. We developed automated routines for this purpose, and we describe their most important features in this section. As sky background is of critical importance, we discuss this issue in some detail in § 3.1.2.

3.1.1. *GALFIT Setup and GALAPAGOS*

GALFIT is designed to fit one galaxy of interest at a time. Therefore, we created an individual postage stamp for each galaxy of interest. These postage stamps were created and initial GALFIT parameter files were produced, by an IDL program, GALAPAGOS (Galaxy Analysis over Large Areas: Parameter Assessment by GALFITting Objects from SExtractor; for further details about GALAPAGOS and details of the procedure, see M. Barden et al. 2007, in preparation). For every object in the SExtractor catalog GALAPAGOS did the following.

1. GALAPAGOS determined the size of the required postage stamp for each object. This was done using different object sizes and angles given by SExtractor:

$$X\text{size} = 2.5a\text{kron}[|\sin(\theta)| + (1 - \text{ellip})|\cos(\theta)|], \quad (2)$$

$$Y\text{size} = 2.5a\text{kron}[|\cos(\theta)| + (1 - \text{ellip})|\sin(\theta)|], \quad (3)$$

where a is the SExtractor output parameter A_IMAGE, kron is KRON_RADIUS, θ is THETA_IMAGE, and ellip is ELLIPTICITY. Extensive testing showed that this algorithm for producing postage stamps was a good compromise between the conflicting needs of having enough sky pixels present in the postage stamp to give a robust fit of the object, while keeping the postage stamps small enough to be fit in reasonable amounts of CPU time.

2. In the next step, GALAPAGOS decided from this postage stamp and the aperture map, which secondary objects had to be deblended and fitted simultaneously and which objects were simply masked out during the fitting process. For this it created a second map in which SExtractor aperture ellipses were increased in linear size by a factor of 1.5 (a factor of 2.25 larger area). Every object whose ellipse overlapped with the ellipse of the primary object was fitted simultaneously using a single Sérsic profile; every other object with pixels in the postage stamp was masked out during the fit, using this expanded ellipses as the mask.[14] This way time-consuming fits, with 10 or more objects to be simultaneously fitted, were avoided in most cases. In total for around 48%, 31%, and 46% of the fits, at least *one* secondary

---

[14] In many other fitting routines the SExtractor segmentation map is used for masking; our masks are considerably more conservative.

Parameter	Starting Guess from SExtractor
Mag	MAG_BEST
r_e	0.162*FLUX_RADIUS$^{1.87}$
b/a	1− ELLIPTICITY
P.A.	THETA_IMAGE
n	1.5
x, y	The postage stamp is centered on the primary object; positions of secondary objects can be derived from SExtractor

object had to be taken into account (for $n = 1$ simulations, $n = 4$ simulations, and real galaxies, respectively). In the most crowded situations we find that we needed to simultaneously fit a maximum of 9, 7, and 12 profiles.

3. After this step the sky background was estimated. For this, GALAPAGOS used the aperture map on the whole science frame (and not only the postage stamp) and estimated the mean value of all pixels that lay within six consecutive elliptical annuli, each with a width of 60 pixels (measured along the semimajor axis; corresponding to 1.8″ using the GEMS data with 0.03″ pixel^{-1}). These six annuli partially overlap, with a spacing of 30 pixels between successive annuli. The annuli were centered on the primary fitting galaxy (pixels belonging to a secondary object were ignored in this step). The innermost area is masked out during this process (the factor of 1.5 magnified aperture ellipse enlarged by a further 30 pixels). These annuli "marched outward" together in radius in steps of 30 pixels until the gradient of the mean values within the last six rings (180 pixels) was larger than −0.05; the change in the sky value, given that the mean GEMS F850LP sky background is around 18 counts, was then well below 0.3% within this radial range. The sky was then determined as the mean value of the outermost six annuli. This made the area where the sky is determined to be an ellipse between 35 and 215 pixels in semimajor axis for the smallest objects (between $15r_e$ and $30r_e$) and an ellipse of width of 180 pixels at around $(4-6)r_e$ for the bigger objects (for details see M. Barden et al. 2007, in preparation). We call this sky estimate the "isophotal sky" in what follows, and testing shows that for fitting with GALFIT the isophotal sky provides significantly better fitting results than using sky values from, e.g., SExtractor (see § 3.1.2).

4. In the same step, by dividing the elliptical individual annuli into eight octants, GALAPAGOS was able to detect sky gradients within an annulus as a function of position angle. Such cases were relatively rare and were due to nearby bright objects that did not reach into the postage stamp themselves (especially bright high Sérsic index objects with strong wings). GALAPAGOS then identified these objects in the SExtractor catalog automatically, and these objects were fitted simultaneously to eliminate this sky gradient (GALFIT can fit profiles that are centered outside of the postage stamp). In the very rare cases that an identification was not possible even though a strong gradient was present (i.e., the object lay outside of the original GEMS tile), we fitted an artificial object centered outside the postage stamp in the correct direction to achieve the same result. In total, 15.2% of the fits in the simulated disk sample needed an additional identified profile centered outside of the postage stamp; 1.5% needed an artificial, not identified profile (4.3% and 0.6% for simulated spheroidal galaxies). For real galaxies only 3% of the fits needed an identified object; 0.4% needed an artificial profile. Recall that the simulated images contained a large number of galaxies not recovered by SExtractor; these galaxies contributed to the background sky only. These galaxies can lead to "sky" gradients

found by GALAPAGOS. This effect should be, and is, more pronounced in the sample in which fewer galaxies are recovered.

5. The last step for setting up GALFIT was the determination of the starting guesses for the different fitting parameters from SExtractor and writing them to a GALFIT start file automatically (see Table 3). We decided to fit single Sérsic profiles to all galaxies (with a starting value of 1.5 for the Sérsic index). Starting magnitudes were given by SExtractor MAG_BEST; sizes were given using FLUX_RADIUS. (We used the formula $r_e = 0.162R_{flux}^{1.87}$, where R_{flux} is FLUX_RADIUS. This formula was determined empirically using simulations.) The axis ratio b/a was derived by taking the SExtractor ELLIPTICITY, the position angle by THETA_IMAGE. Furthermore, the position of the objects within its postage stamp was required, which was directly given by the cutting process of the postage stamps (the object is centered within its postage stamp; see step [1]). The parameter diskiness/boxiness in GALFIT was *fixed* to 0 (no boxiness/diskiness) for all our fits. Furthermore, as described above, the estimated sky value from step (3) *was held fixed during the fit*. Each object that had to be deblended during the fitting process was included (from step [2]) with its appropriate starting values; all other objects were masked out (using a mask image with the SExtractor apertures enlarged by a factor of 1.5, which tells GALFIT which pixels it should use and which pixels it should ignore during the fit). Finally, the GEMS PSF (see K. Jahnke et al. 2007, in preparation) was provided to GALFIT.

We adopted a set of fitting constraints for GALFIT which prevented the code from exploring unphysical (and time-consuming) areas of parameter space. We used $0.2 < n < 8$, $0.3 < r_e < 500$ (pixels of 0.03″ in size), and fixed the fit magnitude to be within 5 mag of the SExtractor MAG_BEST.

We used exactly the same constraints for real galaxies, also using the same setup procedure. Whenever we state that GALFIT fitted "successfully," we mean that GALFIT returned a result (it did not crash during the fit) *and* the fit did not run into any of the constraints given above.

3.1.2. *GALFIT Sky Test*

The estimate of the sky background is of critical importance in determining parametric or nonparametric descriptions of galaxy surface brightness profiles (e.g., de Jong 1996). While in principle it is possible to fit the sky level as an extra parameter, such a procedure requires that the surface brightness profile being used is an *accurate* description of the real galaxy light profiles. An alternative is to estimate the sky level as carefully as possible prior to the fit and hold it fixed.

In this section we quantify the effect of different assumptions/ estimates of the sky level for GALFIT (the results of the equivalent test for GIM2D are shown in § 3.2.3). We test three setups: (1) the isophotal sky, (2) the sky value determined by SExtractor, and (3) allowing sky to be a free parameter, to be estimated by GALFIT.
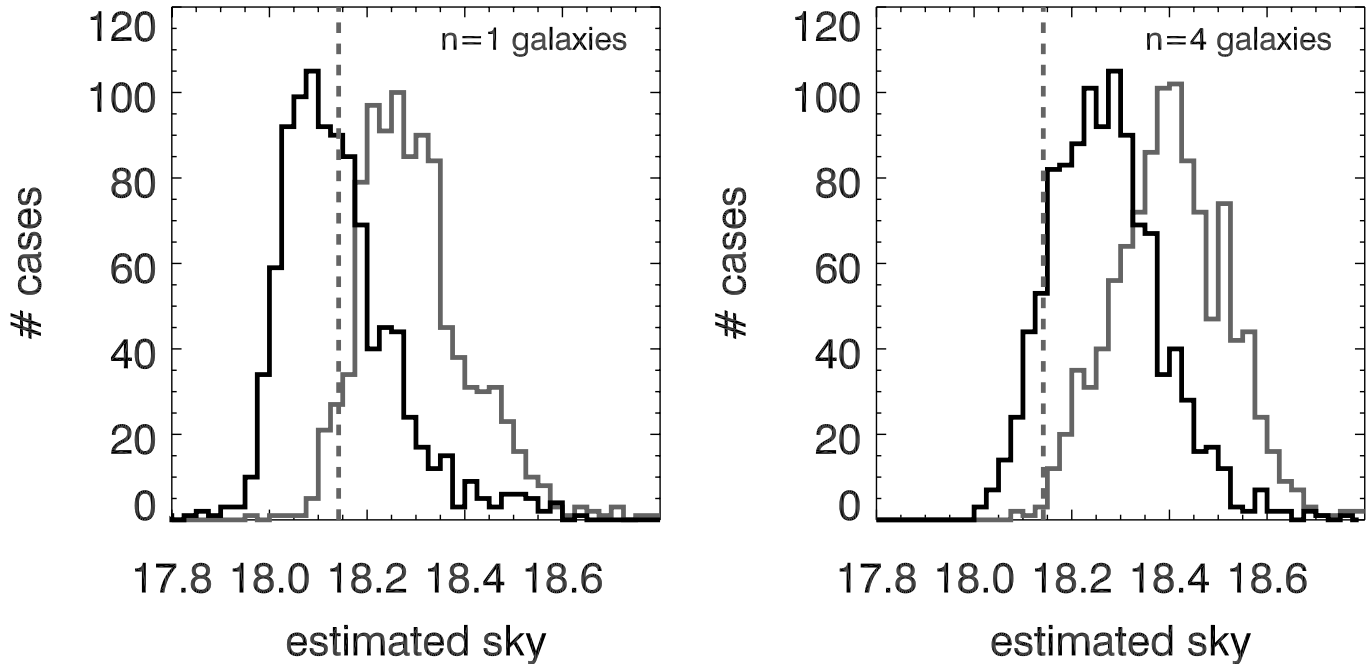
Fig. 4.—Recovered sky values for both used estimators, SExtractor (*gray*) and isophotal (determined within GALAPAGOS, *black*) for each of the simulated samples. The vertical dashed line indicates the true value. One can clearly see that both methods tended to overestimate the sky value for the spheroid sample, mainly due to the large wings of galaxies in this sample contributing to the sky level.

In Figure 4 one can see the difference between the sky values derived by the two sky estimation methods, GALAPAGOS and SExtractor, for the two different samples of simulated $n = 1$ and $n = 4$ galaxies, respectively. Because the simulations were added to a sky frame composed of empty patches of real sky, the *true* sky values were known to be 18.14 ± 0.03, indicated by the vertical dashed line in both plots. SExtractor recovers a mean value of 18.29 ($\sigma = 0.10$) for disk galaxies and 18.40 ($\sigma = 0.11$) for spheroidal galaxies. The isophotal estimator in GALAPAGOS gives a mean value of 18.13 ($\sigma = 0.10$) for disk galaxies and 18.26 ($\sigma = 0.11$) for spheroidal galaxies. Although all distributions have around the same width, one can see that both methods recover the sky better for the low Sérsic index sample. Furthermore, in both samples, the isophotal estimator gives back rather more accurate sky values.

That SExtractor recovers a sky value that is slightly too high has been noted before, e.g., by the GOODS team,[15] and was the reason why we decided to write our own isophotal sky estimator.

That the sky is easier to estimate for the $n = 1$ simulations than for the $n = 4$ simulations can be partly explained by our simulation of a number of large, low surface brightness galaxies which escape detection by SExtractor and which inflate the sky surface brightness. Since galaxies with high Sérsic index n have more extended wings, the effect of contamination in the outskirts is larger for $n = 4$ simulations than for $n = 1$. There is a further effect for $n = 4$ galaxies: since the sky estimates provided by SExtractor and GALAPAGOS only probe out to $<6r_e$ for brighter galaxies, there is a residual contribution to the sky from the galaxy itself which becomes more serious as n increases. We compare the fitting results with the three different sky setups in Figure 5. We only show results for the sample of simulated $n = 4$ galaxies; the results for the $n = 1$ galaxies were qualitatively similar, but the systematic effects are much weaker, showing very little difference

---

[15] See http://www.stsci.edu/science/goods/catalogs/r1.0_zreadme/, ch. 5.1, Local sky background.

---

between the three different sky setups. The *y*-axis shows the deviation of the three key parameters, magnitude, r_e, and n, from their true values, and the *x*-axis shows the simulated mean surface brightness μ_{input} of the galaxies within an ellipse with semimajor axis r_e and the axis ratio b/a:

$$\mu = \mathrm{mag} + 2.5 \log\left[2(b/a)\pi r_e^2\right], \qquad (4)$$

where mag is the magnitude, b/a the axis ratio, and r_e the half-light radius of the object in arcseconds. The factor of 2 accounts for the fact that only half the light is within the half-light radius. The top axis shows the mean signal-to-noise ratio (S/N) per pixel corresponding to that average surface brightness μ, given by

$$\mathrm{S/N} = \langle\rho\rangle\left[\langle\rho\rangle + \langle\rho_{\mathrm{sky}}\rangle + \sigma_{\mathrm{sky}}\right]^{-1/2}, \qquad (5)$$

where $\langle\rho\rangle$ is the average count rate (in e^-) for galaxy pixels within r_e, $\langle\rho_{\mathrm{sky}}\rangle$ is the background flux (in e^-) within a pixel, and σ_{sky} is the uncertainty of the background sky estimate, obtained from the empty sky image.

At faint surface brightness levels, one can see that magnitudes are typically overestimated (i.e., are recovered too faint), sizes are systematically underestimated, and Sérsic indices are typically underestimated (Fig. 5). The effects are subtle and affect only galaxies much fainter than the sky surface brightness for the isophotal sky and the sky fit by GALFIT. These effects set in at much higher surface brightnesses (approximately 2 mag arcsec^{-2} above the sky level) for SExtractor-derived sky values.

Following these test results (Tables 4 and 5), and the general concern that galaxies may deviate from parametric descriptions of their light profiles in their outer parts, we choose to use the isophotal sky estimate for GALFIT analysis for this paper (and those used in other papers, e.g., Barden et al. 2005). Should one not have access to accurate sky values from GALAPAGOS or a similar routine, the test results show also that allowing GALFIT to estimate the sky levels for single Sérsic profiles is an acceptable
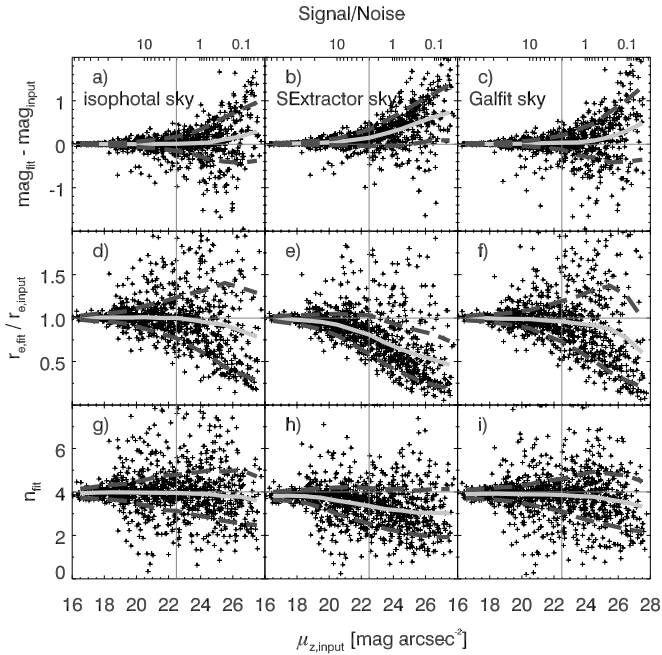
Fig. 5.—Fitting results of the spheroid galaxy sample when fitted with GALFIT using three different sky estimates (see § 3.1.2): isophotal sky from GALAPAGOS (*left*), SExtractor sky (*middle*), and the sky returned by GALFIT when allowed to fit it as a free parameter (*right*). The *x*-axis shows the simulated mean surface brightness within r_e defined by $\mu = \text{mag} + 2.5 \log\left[2(b/a)\pi r_e^2\right]$, where mag is the magnitude, b/a is the axis ratio, and r_e is the half-light radius of the object in arcseconds. The thin vertical gray line in the plot indicates the brightness of the sky background. The upper *x*-axis shows the mean S/N per pixel within r_e calculated by S/N = $\langle\rho\rangle[\langle\rho\rangle + \langle\rho_{\text{sky}}\rangle + \sigma_{\text{sky}}]^{-1/2}$, where $\langle\rho\rangle$ is the average counts in a galaxy pixels within r_e (basically μ), $\langle\rho_{\text{sky}}\rangle$ is the background flux within a pixel, and σ_{sky} is the uncertainty of the background sky estimation. Although this number is only a rough approximation, it gives a feeling about the mean S/N of the galaxies. The *y*-axis shows magnitude difference (fitted − simulated), size ratio (fitted/simulated), and Sérsic index–fitting results. Perfect parameter recovery is indicated as the horizontal thin dark-gray line. The thick light-gray line and the thick dashed dark-gray line indicate the mean value and 1 σ values for different surface brightness bins, respectively. The small crosses show the galaxies that were fitted "successfully," meaning that the fit returned a result and that it did not run into fitting constraints. As one can clearly see, getting a good estimation of the sky level is important. Both using the isophotal sky estimation (*left*) and using the sky level as a free parameter during the fit (*right*) return more reliable results than using SExtractor sky estimations.

alternative, provided that the surface brightness profiles of the galaxies of interest are well-approximated by a Sérsic profile over a wide range of radii.

3.2. *GIM2D*

GIM2D (Galaxy Image 2D) was written by Luc Simard (Simard 1998, 2002) as an IRAF package for the quantitative morpho-

logical analysis of galaxies. We use version 3.1 for the analysis in this paper. For a single Sérsic fit we work in seven dimensions with the bulge fraction parameter set to $B/T = 1$; thus, we find the best-fit model described by $f_{\text{tot}}, r_e, e, \phi_B, dx, dy$, and n. During the fit, the images are deconvolved with a given PSF. GIM2D uses the Metropolis algorithm to find a χ^2 minimum, which makes it less prone to settle on local minima. On the other hand, this algorithm is time consuming. Accordingly, to process large data sets, GIM2D ought to be run on many machines in parallel.

3.2.1. *GIM2D Setup*

As with GALFIT, GIM2D requires certain generic considerations for galaxy profile fitting: (1) postage stamp construction, (2) nearby companion masking, (3) background sky estimation, and (4) initial parameter guesses. We did not use GALAPAGOS to set up GIM2D's galaxy fit for two reasons: GIM2D is embedded into IRAF, whereas GALAPAGOS requires IDL; and the simultaneous fitting of galaxies is not supported in GIM2D, whereas much of GALAPAGOS's algorithm is devoted to making decisions about which galaxies are to be simultaneously fit. Therefore, GIM2D is set up by using a different procedure, which we describe in this section.

Starting with the combined hot/cold SExtractor output catalogs, a square postage stamp was cut from the large image, centered on each galaxy with size given by $4a_{\text{iso}} \times 4a_{\text{iso}}$, where a_{iso} is the major axis diameter of the SExtractor isophotal area in pixels (the minimum postage stamp size we allowed was 101×101 pixels). GIM2D masks out nearby objects using SExtractor segmentation maps: discussion of the consequences of this procedure is presented later in this section and in § 4.1.2. For sky estimation and defining the best part of the fitting parameter space to explore, GIM2D has several important setup parameters that allow the user to modify its behavior. In this section, we describe some of the most important—parameters that we find to critically affect the performance of the code.

The parameter dobkg specifies whether GIM2D determines the background itself (dobkg=yes) or fixes the sky to a user-defined value (dobkg=no). With dobkg=yes, GIM2D calculates the background prior to galaxy fitting directly from the postage stamp images of each source using only nonobject (sky) pixels as specified by the SExtractor segmentation map. As such, this method is closely dependent on extracting a large enough image to get a reliable sky measurement. Once determined, the sky value is held fixed during the fitting. If dobkg=no, GIM2D assumes that the postage stamps have background equal to zero; therefore, the user may use an external method to estimate the sky and subtract this from the input images. GIM2D does offer an option to fit the background offset (parameter db) as a free parameter during fitting, but this is not recommended when working with real galaxies

TABLE 4
Fitting of $n = 4$ Simulations Using GALFIT: Bright Subsample with $\mu_{\text{in}} < 22.5$ and $\text{mag}_{\text{in}} < 22.5$

Sky used	Sérsic n		r_{50} Ratio		Δ mag		e Ratio		ΔP.A.		Quality
	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	
Isophotal sky	3.99	0.27	1.00	0.05	0.00	0.03	0.99	0.04	0.17	1.54	0.06
SExtractor sky	3.79	0.29	0.96	0.07	0.03	0.04	1.00	0.04	0.17	1.56	4.66
GALFIT sky	3.94	0.24	0.99	0.06	0.01	0.03	1.00	0.04	0.16	1.55	0.38

Notes.—This table summarizes the results from using different sky estimators with GALFIT for bright galaxies; see Table 5 for results for faint galaxies. The columns give deviations (resistant mean values clipped at 3 σ) from the simulated value and scatter for the five key fitting parameters. The σ-values given are values computed iteratively for all galaxies within 3 σ. The last column gives the fit quality. This number is defined as Quality = $1000\{(\Delta n/4 - 1)^2 + (\Delta r_e - 1)^2 + (\Delta \text{mag})^2 + [\Delta(b/a) - 1]^2 + (\Delta \text{P.A.}/180)^2\}$, where Δ-values are given as the mean values in the table. This quantity is a fairly intuitive combination of the different fit parameters, indicating in broad terms which setups perform well (low values) and which perform poorly (high values). One can see that indeed using the isophotal sky as given by GALAPAGOS and using the sky level as a free parameter during the fit return much more reliable results than the SExtractor sky already for these bright galaxies. Using the isophotal sky seems to be the ideal setup.

TABLE 5
FITTING OF $n = 4$ SIMULATIONS USING GALFIT: FAINT SUBSAMPLE WITH $23.5 < \mu_{in} < 26.0$

	SÉRSIC n		r_{50} RATIO		Δ_{MAG}		e RATIO		ΔP.A.		
SKY USED	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	QUALITY
Isophotal sky...............	3.95	1.13	0.97	0.44	0.02	0.42	1.04	0.18	0.20	6.31	2.92
SExtractor sky............	3.05	0.98	0.61	0.27	0.41	0.39	1.05	0.18	−0.06	6.38	375.33
GALFIT sky................	3.78	1.16	0.94	0.47	0.09	0.44	1.04	0.19	0.07	6.12	16.93

NOTES.—Same as Table 4, but for faint galaxies. As one can clearly see from this table and as was expected, fitting faint galaxies is more difficult. Using the isophotal sky during the fit returns the best fitting results, slightly better than the internal estimation in GALFIT. Using the SExtractor sky returns significantly worse results.

with nonidealized profiles. We test the effect of different sky estimates in detail in § 3.2.3.

GIM2D, like GALFIT, has constraints which can be applied to limit the regions of parameter space searched for solutions. GIM2D starts with a user-specified parameter space, given by the initial value and minimum/maximum hard limits for each parameter to be fit. GIM2D has an option to automatically narrow the focus of the input parameter space by setting `initparam=yes`. With this setup option GIM2D uses FOCAS-like image moments based on information extracted from the SExtractor-created segmentation map to estimate the hard limits for the model parameter space.

Under all setups GIM2D starts in the Initial Condition Finder (ICF) mode, which explores the user-specified parameter space coarsely to find the best initial model guess. In practice, the ICF creates N_{ICF} models throughout the allowed parameter space, selects the best one, and then reduces the search volume by a factor equal to N_{ICF}. The final result from the ICF is used as the starting point by the Metropolis algorithm. The GIM2D Web site gives a default value of $N_{ICF} = 100$.

To find the *best-fitting* setup, we rigorously tested a large number of different setups of GIM2D. We do not discuss all of the different setups here; the most important ones are shown in Tables 7 (for bright galaxies) and 8 (for faint galaxies) and will be discussed in detail in the following sections starting with the recommended GIM2D setup (§ 3.2.2), sky tests (§ 3.2.3), other tests (§ 3.2.4), and concluding with the final adopted best-fitting setup (§ 3.2.5).

3.2.2. GIM2D Recommended Setup

In Figure 6, in the leftmost panels, we show fitting results for the setup that is recommended on the GIM2D Web site[16] (setup K in Table 7). This recommended setup, in particular, has 'dobkg'= yes and `initparam=yes`; i.e., GIM2D determines the sky level and fitting constraints from SExtractor output. As is clear from this plot, this setup produces unsatisfactory results even for fairly high surface brightness galaxies and where GEMS survey completeness is still quite high. The systematic errors are already ∼50% in r_e near the sky level. Fitting results are strongly systematically biased toward fainter magnitudes, smaller sizes, and lower concentrations.

As most galaxy surveys aim to push their analysis down to faint levels, the ideal performance of any fitting code is to provide parameter estimates that are free of significant systematic trends. Therefore, we deem the recommended setup to not be suitable for the GEMS survey. In an attempt to improve the GIM2D performance, we tried a number of different strategies, among them different settings of `initparam` and `dobkg`.

[16] The GIM2D description is http://www.hia-iha.nrc-cnrc.gc.ca/STAFF/lsd/gim2d/ and specifies the last program update of 2001 March 19.

Through extensive testing, we find that the best results are obtained when *both* `initparam=no` and `dobkg=no` (see rightmost panels in Fig. 6), and when the SExtractor local background is used (see § 3.2.3). Setting `initparam=no` and `dobkg=yes` produces very modest improvement. Setting `dobkg=no` and `initparam=yes` helps considerably, giving satisfactory results for galaxies with surface brightness higher than the sky surface brightness.[17] As can be seen in Figure 6, using a fixed background (`dobkg=no`) and setting `initparam=no` removes the strong systematic trend toward poorer fits for low surface brightness galaxies, albeit with large scatter at the faint end. As explained

[17] This is the setup that was used by McIntosh et al. (2005) for their study of the evolution of the early-type $n > 2.5$ galaxy luminosity–size and stellar mass–size relations. Their sample of $n > 2.5$ galaxies all had F850LP surface brightness brighter than 22.5 mag arcsec^{-2}, and inspection of the third row of panels in Fig. 6 and setup J in Table 7 shows that at these limits the GIM2D fitting results suffer from ≲10% biases.
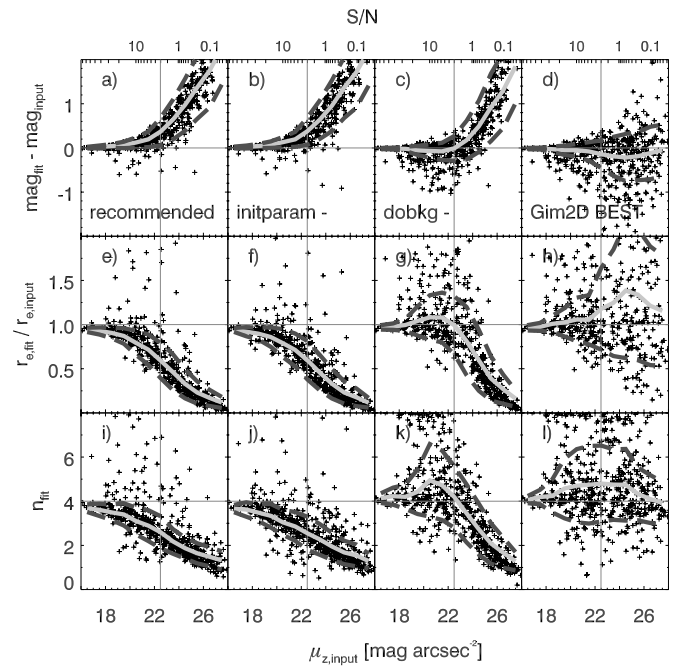

FIG. 6.—Comparison of the recommended GIM2D setup to setups where we used different settings of the GIM2D parameters `initparam` and `dobkg` [*from left to right*: setup K (`initparam`=+, `dobkg`=+; recommended), J (−, +), C (+, −), and A (−, −; best), see Table 7]. The x-axis shows the simulated surface brightness of the galaxies. The y-axes again show magnitude difference, size ratio, and Sérsic index–fitting results. The thick light-gray line and the thick dashed dark-gray line indicate the mean value and the 1 σ line for different surface brightness bins, respectively. One can see easily that our *best-fitting* GIM2D setup (see § 3.2.5) fits galaxies with much less systematic bias than the initial setup recommended on the GIM2D Web page. This is especially true for galaxies fainter than the surface brightness of the sky.

Parameter	Minimum	Maximum
Mag	20	27
r_e	0.3	500
Ellipticity.............................	0.0	1.0
P.A.	All	All
n..	0.2	8.0
Centering x...........................	0	3.0
Centering y...........................	0	3.0

in § 3.2.1, the initparam=no option allows GIM2D to explore the full range of parameter space when determining the best-fit solution. In contrast, setting this option to yes narrowly constrains the magnitude and r_e for objects for which the SExtractor segmentation map severely misses the total extent of galaxies below the sky brightness. It is worth noting that compared to these two parameter choices, other effects such as the precise fixed sky value used (see § 3.1.2) and the image size (§ 3.2.4) appear to produce only minor improvements.

The minimum and maximum limits of the parameter space that we allowed in our best setup are manually set to span more than the entire range of the simulations in terms of size, luminosity, and Sérsic index, more than the physically useful parameter range of real galaxies, so that the solutions are not "pinned up" against the boundary values artificially imposed on them (the actual values are given in Table 6). Fits that ran into any of the given fitting constraints were removed from the sample for the analysis in this paper.

From our findings, we *strongly recommend* that GIM2D users *avoid* the dobkg=yes option and *be cautious* of the surface brightness effects that arise when using initparam=yes.

3.2.3. *GIM2D Sky Test*

In the above, we showed that sky value estimation can dramatically affect GIM2D fits using the default (recommended)

setting. In this section, we repeat the sky analysis for GIM2D as carried out in § 3.1.2 for GALFIT. We tested GIM2D using the SExtractor local sky (setups A, B, and C in Table 7) and the isophotal sky that we used for GALFIT (setups D, E, F, G, and H). We carried out one test, setup I, where we used the "real" background of 18.14 counts, determined on the sky image used in the simulation process. Such a setup is of academic interest only, as for real galaxies it is impossible to measure such a sky value. Nonetheless, this test gives insight into the performance of GIM2D when the actual known sky value is used as an input for galaxy fitting. For setup J and K we tried fixing the sky background to the value determined directly by GIM2D (dobkg=yes). We show results from these different tests in Figure 7.

It is interesting that GIM2D performs somewhat better using SExtractor local sky, whereas we showed in § 3.1.2 that GALFIT performs somewhat better using the isophotal sky from GALAPAGOS. It is likely that the cause of this behavior is related to how GALFIT and GIM2D deal with nearby neighbors. Since GALFIT simultaneously fits neighboring galaxies which overlap with the galaxy of interest, the isophotal sky estimate better represents the background pedestal that is common to the neighboring sources. GIM2D, on the other hand, is unable to simultaneously fit neighbors and relies on masking neighbors using the SExtractor segmentation map. Thus, the "effective sky" for GIM2D includes flux from the outer parts of the galaxy itself and neighboring sources; as the SExtractor sky is derived from the same "sky" area used for fitting, it is a more appropriate value.

3.2.4. *Other Tests of GIM2D*

To determine the best-fitting setup to use with GIM2D, we performed 11 different tests (rows A–K) as shown in Tables 7 (bright galaxies) and 8 (faint galaxies). For the bright galaxies, we selected all N galaxies with $\mu_{sim} < 22.5$ magnitudes arcsec^{-2} and mag$_{sim} < 22.5$ (representing the sample of early-type galaxies from McIntosh et al. [2005], i.e., those important for surveys of early-type galaxy evolution) from the set of N_{tot} galaxies in the sample for which GIM2D returned a result. The faint sample included galaxies with 23.5 mag arcsec$^{-2} < \mu_{sim} < 26$ mag arcsec^{-2}. We

TABLE 7
FITTING OF $n = 4$ SIMULATIONS USING GIM2D: BRIGHT SUBSAMPLE WITH $\mu_{in} < 22.5$ AND mag$_{in} < 22.5$

SETUP	N/N_{tot}	SÉRSIC n Mean	σ	r_{50} RATIO Mean	σ	Δ_{MAG} Mean	σ	e RATIO Mean	σ	ΔP.A. Mean	σ	QUALITY
A..........	164/540	4.34	0.61	1.02	0.13	−0.04	0.07	1.01	0.05	0.00	2.2	7.02
B..........	168/533	4.37	0.43	1.02	0.09	−0.03	0.05	1.01	0.06	−0.16	1.5	8.38
C..........	165/549	4.32	0.56	1.01	0.10	−0.03	0.07	1.02	0.06	−0.10	1.8	6.31
D.........	161/533	5.57	1.49	1.14	0.22	−0.14	0.13	1.02	0.07	−0.20	2.0	154.57
E..........	164/531	4.84	0.76	1.11	0.17	−0.09	0.08	1.01	0.06	−0.17	2.0	44.42
F	165/539	4.75	0.67	1.10	0.17	−0.08	0.08	1.01	0.06	0.05	2.0	35.63
G..........	162/539	4.82	0.75	1.10	0.17	−0.08	0.07	1.01	0.06	0.34	2.2	42.02
H..........	161/545	4.81	0.76	1.12	0.16	−0.09	0.09	1.01	0.06	−0.03	1.6	40.63
I...........	163/533	5.68	1.40	1.27	0.30	−0.16	0.14	1.01	0.06	−0.03	1.7	177.28
J...........	167/551	3.30	0.49	0.82	0.14	0.09	0.09	1.00	0.05	−0.18	1.9	30.40
K..........	168/546	3.33	0.51	0.83	0.13	0.09	0.09	1.01	0.05	−0.01	1.6	28.32

NOTES.—This table summarizes the results from all GIM2D testing for bright galaxies; see Table 8 for results for faint galaxies. The expression N/N_{tot} gives the numbers of galaxies N selected from the total sample of N_{tot} that GIM2D returns a result for each setup. The following columns give deviations (resistant mean values clipped at 3 σ) from the simulated value and scatter for the five key fitting parameters. The σ-values given are values computed iteratively for all galaxies within 3 σ. The last column gives the fit quality as defined in Table 4. Explanation of the setups: (A) SExtractor local background, initparam=no, $N_{ICF} = 100$, and $4a_{iso}$ image sizes (*best* setup); (B) SExtractor local background, initparam=no, $N_{ICF} = 100$, and $2a_{iso}$ image sizes; (C) SExtractor local background, initparam=yes, $N_{ICF} = 100$, and $4a_{iso}$ image sizes; (D) isophotal background, initparam=no, $N_{ICF} = 100$, and $6a_{iso}$ image sizes; (E) isophotal background, initparam=no, $N_{ICF} = 25$, and $4a_{iso}$ image sizes; (F) isophotal background, initparam=no, $N_{ICF} = 100$, and $4a_{iso}$ image sizes; (G) isophotal background, initparam=no, $N_{ICF} = 400$, and $4a_{iso}$ image sizes; (H) isophotal background, initparam=yes, $N_{ICF} = 100$, and $4a_{iso}$ image sizes; (I) background = 18.14, initparam=no, $N_{ICF} = 100$, and $4a_{iso}$ image sizes; (J) dobkg=yes, initparam=no, $N_{ICF} = 100$, and $4a_{iso}$ image sizes; and (K) dobkg=yes, initparam=yes, $N_{ICF} = 100$, and $4a_{iso}$ image sizes (*recommended* setup).
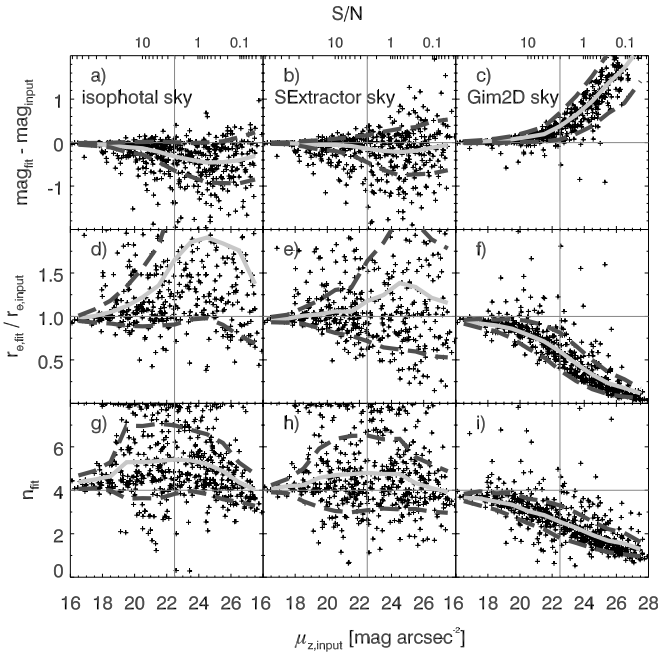
FIG. 7.—Same as Fig. 5, but for GIM2D results (setups F, A, and J in Table 7). As is immediately obvious from these plots, GIM2D performs best when using the SExtractor background held fixed during the fit.

then calculated the mean of the recovered value or ratios of the different fit parameters and the 68% confidence interval.

In our visual examination of the properties of the outliers in these distributions, we found that most of the non-Poisson scatter is caused by contamination of the outer isophotes of the object of interest by nearby neighbors. While this issue is discussed in more detail later in § 4.1.3, we illustrate this behavior by running GIM2D on three different postage stamp widths of $2a_{iso}$ (setup B), $4a_{iso}$ (setup A), and $6a_{iso}$ (setup D). These tests find that there is an increase in scatter for larger image size, consistent with the expectation of contamination. Using $2a_{iso}$ reduces the extreme outlier fraction somewhat from $4a_{iso}$. Yet, since such outliers are a small fraction of the objects, this change in postage stamp size had relatively little impact on the rms scatter (see Tables 7 and 8). The best-fit stamp cutouts we adopt here have sides equal to $4a_{iso}$. This seems to be the best compromise between a postage stamp large enough so that GIM2D includes enough of the important outskirts of the galaxies for fitting, but small enough that neighboring galaxies are reasonably rare and CPU requirements are

reasonable. For comparison, the stamp sizes used in GALFIT fits are nearly always larger in area than $4a_{iso} \times 4a_{iso}$, due to the requirement of simultaneously fitting neighboring galaxies. We will further quantify the effect of neighboring galaxies in § 4.1.2 using the set of simulated spheroids examined there.

We also tested whether the initial number of ICF models affected our fitting results. Holding all other setup choices constant, we compared the results from fits to the $n = 4$ simulations with the default value of $N_{ICF} = 100$ (setup F in Table 7) to results for $N_{ICF} = 25$ (setup E) and 400 (setup G). We found that the results are independent of the number of ICF models.

3.2.5. *GIM2D Best-fitting Setup*

As is apparent especially from Table 8 and Figure 6, the best combination of parameters for our simulations was given by setup A, using the SExtractor background (dobkg=no), initparam=no, and $4a_{iso}$ as image sizes. We choose this setup to be our *best* and use it throughout this paper to compare GIM2D results with GALFIT results.

4. GALFIT/GIM2D COMPARISON USING OPTIMIZED SETUPS

In this section, we discuss the results of testing our best setups of GIM2D and GALFIT. Section 4.1 describes the results obtained using the simulated images with artificial distributions of galaxy parameters as explained in § 2. Section 4.2 describes a very similar test using simulated galaxies having more realistic parameter distributions, as derived from real galaxies recovered from individual GEMS survey fields. Section 4.3 sums up the results of tests in which real images of different depths were fitted and the results intercompared.

4.1. *Results of Fitting Simulated Galaxy Images*

4.1.1. *Results of Pure Disk Simulations*

Figure 8 shows both GALFIT and GIM2D results for the set of simulated disk galaxies with an exponential $n = 1$ light profile. Of the 1600 galaxies simulated in this sample, 997 (62%) were recovered by SExtractor. Of these, 979 (98%) were successfully fitted by GALFIT, 12 (1.2%) ran into constraints, and six (0.6%) fits crashed. GIM2D fitted 870 (87%) successfully, 46 (5%) ran into fitting constraints, and 81 (8%) of the fits crashed. There are four (0.4%) galaxies for which both codes failed.

Plus signs in Figure 8 represent galaxies that were fitted by both codes. Gray squares show galaxies that were fitted by that code only; the other code failed to return a useful result either

TABLE 8
FITTING OF $n = 4$ SIMULATIONS USING GIM2D: FAINT SUBSAMPLE WITH $23.5 < \mu_{in} < 26.0$

|		SÉRSIC n		r_{50} RATIO		Δ MAG		e RATIO		Δ PA		
SETUP	N/N_{tot}	Mean	σ	Mean	σ	Mean	σ	Mean	σ	Mean	σ	QUALITY
A.........	151/540	4.63	1.45	1.38	0.78	−0.20	0.55	1.05	0.23	−1.10	7.4	24.83
B.........	149/533	4.52	0.92	1.44	0.67	−0.23	0.43	1.05	0.21	−0.13	7.0	17.35
C.........	154/549	2.46	0.86	0.50	0.23	0.57	0.48	1.01	0.19	−0.64	5.5	149.08
D.........	147/533	5.52	1.36	3.22	1.93	−0.76	0.57	1.00	0.20	−1.22	10.6	148.96
E.........	147/531	5.05	1.27	1.89	0.95	−0.47	0.49	1.02	0.16	−1.17	7.6	69.38
F	149/539	5.13	1.26	1.94	0.93	−0.47	0.46	1.03	0.20	−1.00	6.3	80.32
G.........	148/539	5.24	1.30	2.36	1.45	−0.52	0.51	1.02	0.16	−1.43	6.5	97.78
H.........	153/545	2.45	0.92	0.53	0.27	0.53	0.50	0.96	0.16	−0.88	8.1	150.72
I	146/533	5.38	1.19	2.84	1.51	−0.73	0.53	0.99	0.25	−0.88	7.0	122.88
J..........	155/551	1.79	0.44	0.28	0.11	1.08	0.49	1.02	0.19	−0.57	9.2	306.48
K.........	154/546	1.78	0.40	0.28	0.12	1.10	0.47	1.04	0.15	−1.86	7.5	310.23
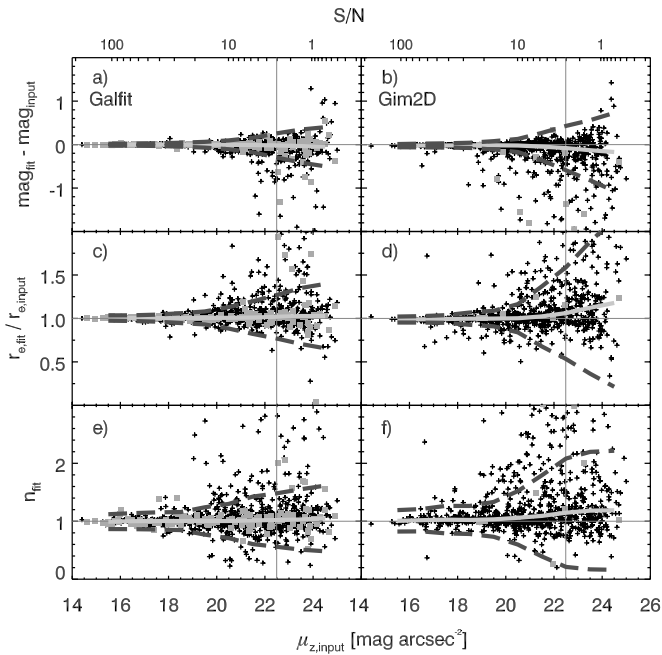
Fig. 8.—Fitting results for GALFIT (*left*) and GIM2D (*right*) for the set of simulated $n = 1$ galaxies. The *x*-axis again shows the input surface brightness; the thin vertical gray line indicates the brightness of the sky background. The *y*-axes are the same as in Fig. 5. The input value is indicated as the horizontal thin dark-gray line, in this case representing a Sérsic index of 1 for the sample of disk galaxies. The mean value of the deviations and a 3 σ line are indicated for different surface brightness bins. The small crosses show the galaxies that were fitted "successfully" by both codes. Gray squares indicate galaxies that were fitted by one code only and for which the other code did not return a meaningful result. The quality parameters as defined in Table 4 for this set of galaxies are 0.05/5.27 for GALFIT (bright/faint) and 1.01/251.67 for GIM2D, showing that GALFIT returns more reliable results for simulated disk galaxies in its optimal setup than GIM2D.
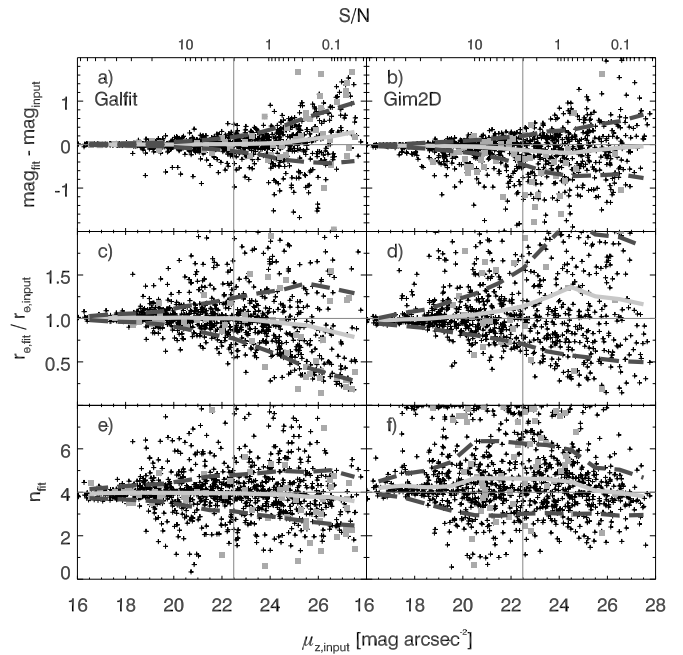
Fig. 9.—Code comparison for $n = 4$ galaxies. This figure is formatted in a similar way to Fig. 8, but for the sample of simulated $n = 4$ galaxies. The *X*-axis is shifted by 2 mag arcsec^{-2} compared to Fig. 8; furthermore, in this plot the gray dashed line represent the 1 σ limits. The value for σ in this sample is around 3 times as large as was the case for disk galaxies. The true value of 4 for the Sérsic index is again indicated as the horizontal thin line. The quality parameter for bright/faint galaxies is 0.06/2.92 for GALFIT and 8.09/122.00 for GIM2D using this sample of galaxies. Numbers given here are different from the numbers given in Tables 7 and 8, as a different, larger sample of galaxies was used for this analysis. From fewer galaxies in Table 8, the best setup A there returns a slightly higher mean Sérsic index which translates into a higher quality parameter. The numbers show clearly that GALFIT returns better results than GIM2D. Interestingly, the quality parameter for GIM2D is smaller for faint spheroids than for faint disks, suggesting that these galaxies are fitted more reliably. This might be due to the way the quality parameter was calculated in detail (using resistant mean offset values cut at 3 σ which removes large scatter; indeed the systematic offset is smaller in case of spheroidal galaxies, the scatter does not go into the quality number) but does not reflect the plots; the disk plots look better due to smaller scatter.

through running into one of the fitting constraints, or the fit crashed. The thick light-gray line and the thick dashed dark-gray line indicate, respectively, the mean value and the 3 σ lines for different surface brightness bins of all galaxies that were fitted using that code. The left column shows fitting results using GALFIT; the right column shows the results of the same set of simulations using GIM2D. The *x*-axis, showing the simulated surface brightness of the galaxies, is the same for all six plots. The three rows show the results for magnitude (Figs. 8*a* and 8*b*), size (Figs. 8*c* and 8*d*), and Sérsic index (Figs. 8*e* and 8*f*). The thin vertical line indicates the brightness of the sky background within the GEMS survey. This is also roughly the limit up to which real galaxies are used for science within GEMS. The *y*-axes show deviations of the fitting values from the true parameter values; the horizontal thin line indicating the ideal value, which, in the case of this galaxy sample, is simply the simulated value.

We will discuss here and in all other sections the behavior of the codes in three different surface brightness bins: first, the galaxies of highest surface brightness, which one clearly would want to be fitted well with any code; second, galaxies within a surface brightness bin of 1 mag around the surface brightness of the sky; and third, the faintest galaxies, much fainter than the sky surface brightness. The third are the galaxies that are obviously hardest to fit. Here the results from the two codes differ the most from each other.

To summarize our general findings, for $n = 1$ galaxies brighter than the sky's surface brightness, there is no significant mean off-set between the input and recovered values in Figure 8; however, the scatter in the GIM2D results is somewhat larger. For this sample of galaxies, this would mean that the final results would be

statistically unaffected by one's choice of fitting code, but for individual objects the reliability of the GALFIT results is slightly higher.

For galaxies around the sky surface brightness, there are small systematic trends and increased scatter for our setup of GIM2D: a size ratio of 1.06 (r_{fit}/r_{sim}, $\sigma \approx 0.18$) for GIM2D and a ratio of 1.02 for GALFIT ($\sigma \approx 0.08$). This trend continues toward fainter surface brightness, although at no point does the systematic size offset exceed 20%.

From Figure 8 (gray squares show objects fitted only by the respective code), one can easily see that GALFIT returns a result more often than GIM2D, although the fraction of galaxies with failed fits is small in both cases. It is interesting to note that the properties of galaxies with failed fits is somewhat different between the two codes: galaxies not fitted by GALFIT (those fitted only by GIM2D) are fainter than average (the parameters are discrepant even using GIM2D), whereas galaxies not fitted by GIM2D (those fitted only by GALFIT) are fitted almost as well by GALFIT as other galaxies with the same surface brightness.

4.1.2. *Results of Pure Spheroid Simulations*

Figure 9 shows the same plots as Figure 8 but for the simulated set of $n = 4$ profiles representing the light profile of a typical early-type galaxy. The total number of galaxies recovered in this sample out of 1600 simulated objects was 1091 (68%). Of these,

only two (0.2%) crashed in GALFIT, and 56 (5.1%) ran into constraints; GIM2D crashed on 31 galaxies (2.8%), and 36 (3.3%) additional fits ran into fitting constraints. Fifty-four (5.0%) galaxies were fitted by GIM2D that were not fitted by GALFIT, 63 (5.8%) galaxies were fitted by GALFIT and not fitted by GIM2D, and both codes crashed on four (0.4%) galaxies in common.

When comparing Figure 9 to Figure 8, one should be aware that the x-axis is shifted by 2 mag arcsec^{-2} toward fainter surface brightness and that the dashed lines for clarity indicate $1\,\sigma$ instead of $3\,\sigma$ as in Figure 8. It is clear that both codes recover the parameter values for $n = 4$ galaxies significantly less accurately than was the case for the $n = 1$ disks, resulting in a substantially larger scatter. This is due to two different effects. First, spheroidal profiles are in principle harder to fit due to the importance of the outskirts of the light profile; this makes using an appropriate sky estimate much more important for a successful fit. Second, due to the large amount of light in the faint wings of the galaxies, neighboring objects have a much bigger influence on the fit of the galaxy of interest than was the case for the exponential light profiles. This effect is particularly important for this simulated galaxy sample, because it was designed to have an unrealistically high number of large $n = 4$ galaxies.

As was the case for disk galaxies, both codes are basically indistinguishable for high surface brightness galaxies in a statistical sense. For galaxies with surface brightness close to that of the sky, our implementation of GALFIT recovers slightly better parameter values than GIM2D (size ratio of 1.00, $\sigma \approx 0.23$, and a somewhat asymmetric error distribution for GALFIT; size ratio of 1.14, $\sigma \approx 0.44$, and more asymmetric errors for GIM2D). The trend continues toward lower surface brightness, with the GIM2D showing increasingly important systematic offsets and a substantially increased scatter. The directionality and asymmetry of the scatter in all plots (GIM2D and GALFIT) are caused by neighboring contamination that is not fully removed, keeping in mind that 32% of the simulated galaxies escape detection by SExtractor because of their low surface brightness.

4.1.3. Deblending Effects

Given the significant differences in philosophy when it comes to the deblending techniques between GALFIT (multiobject fitting and masking) and GIM2D (masking only), we explore the recovery of input parameters as a function of the immediate environment of a galaxy for both codes. We analyze the subset of 390 (out of a total of 1033) $n = 4$ simulated galaxies for which GALAPAGOS decided that GALFIT needed to simultaneously fit two or more profiles. This has the advantage that only significant neighbors are included in this analysis and should be sufficient to demonstrate the magnitude of the influence of deblending on the quality of galaxy fitting with GALFIT and GIM2D.

The results are summarized in Figure 10, showing the difference between recovered and input magnitude as a function of the distance to the next neighbor (*left*) and as a function of the brightness of this neighbor (*right*).

Fitting neighboring objects simultaneously, GALFIT (Figs. 10a and 10b) is able to deblend these galaxies reliably, and the deviations of the fitting magnitudes is independent of both distance and brightness of the closest neighbor. For GIM2D (Figs. 10c and 10d), it is clear that the fitting residual is a strong function of both distance and brightness of the nearest neighbor. The closer and brighter a neighboring object is, the larger the magnitude deviation. In an attempt to disentangle the influence of distance and brightness, we try to correct for the systematics observed in Figures 10c and 10d by removing the offsets and the slope, showing the results in panels Figures 10e and 10f. It is clear that the
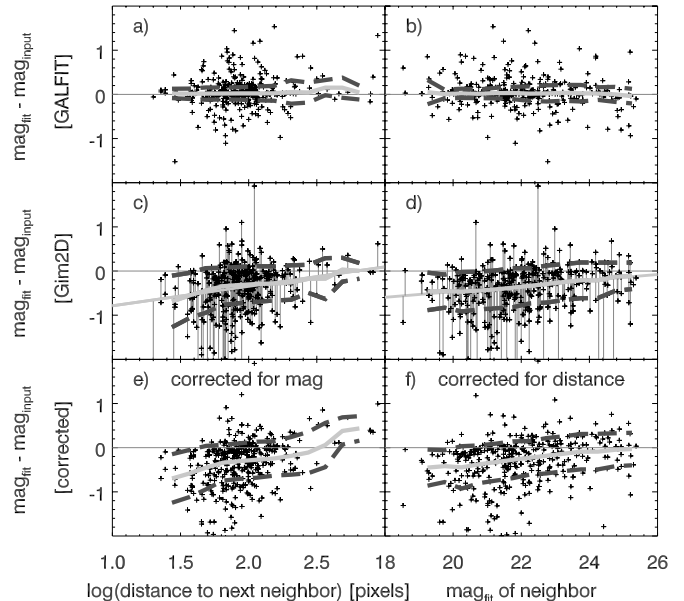

Fig. 10.—Impact of neighboring galaxies on fit results with GALFIT (*top*) and GIM2D (*middle and bottom*). The left column shows magnitude deviations from the simulated values as a function of the distance to the next neighbor; the right column shows the difference between the recovered and simulated values as a function of the brightness of the nearest neighbor. GIM2D shows strong systematic offsets as a function of both distance to the nearest neighbor and its brightness. In the lower two panels, we try to correct for the systematics observed in (c) and (d) by showing the distance dependence of the offset-magnitude relation residuals (e) and the magnitude dependence of the offset-distance relation residuals (f).

distance and brightness effects of the nearest neighbor cannot be easily corrected and thus can significantly impact the performance of GIM2D in recovering the true parameters for simulated $n = 4$ galaxies. For isolated galaxies, GIM2D does an excellent job of recovering the properties of $n = 4$ galaxies.

4.2. Results of Simulations Representing Simulated GEMS Tiles

Bearing in mind the importance of neighboring galaxies in determining the quality of fit, we repeated the above analysis using a sample of galaxies in which $n = 1$ and $n = 4$ galaxies were intermixed with realistic clustering, sizes, and magnitudes. To this end, simulations were produced from the GALFIT results of two real GEMS tiles using recovered values of magnitude, position, and size. The only parameter that was changed was the Sérsic index. Every galaxy with a real Sérsic index of 2.5 or smaller was simulated with a Sérsic index of 1 and all others were simulated with a Sérsic index of 4. These simulations have the advantage that they are better able to estimate the uncertainties of galaxy fits with GEMS data.

The results are shown in Figures 11 and 12. It is worth noting that the range in galaxy surface brightness is much smaller in these simulations, although we have left the x-axis the same as in the previous plots to facilitate comparison with these. We also show the surface brightness histograms of galaxies used in Barden et al. (2005; disk galaxies) and McIntosh et al. (2005; spheroidal galaxies) to show which areas of parameter space are especially important for scientific analysis of data.

Inspecting Figures 11 and 12, it becomes clear that GIM2D and GALFIT perform more similarly for galaxy populations with clustering and properties typical of medium-depth cosmological *HST* surveys than for purely simulated data. GALFIT shows increased scatter and mild systematic offsets compared to the earlier
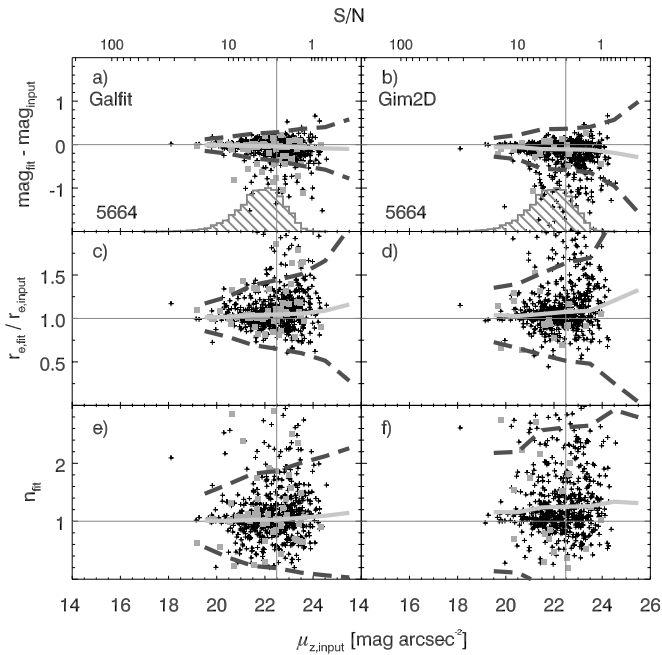
FIG. 11.—Fit results from resimulated GEMS tiles (see § 4.2) for disk galaxies. Same as Fig. 8, but for a sample of two GEMS tiles that were resimulated in order to create a more realistic distribution of galaxy parameters and object density (the results for the $n = 4$ galaxies in this sample are shown in Fig. 12). In the upper panel we plot the surface brightness histogram of the 5664 disk galaxies that were selected for analysis by Barden et al. (2005), showing where fitting accuracy is especially important.

simulations. In the case of the $n = 1$ galaxies the difference in behavior is especially pronounced: it is clear that the presence of realistically clustered $n = 4$ galaxies around $n = 1$ galaxies is a larger source of random error in galaxy fitting for both GALFIT and GIM2D than in pure $n = 1$ simulations. GIM2D shows very similar behavior compared to the earlier simulations, with still larger scatter and systematic offsets than GALFIT.
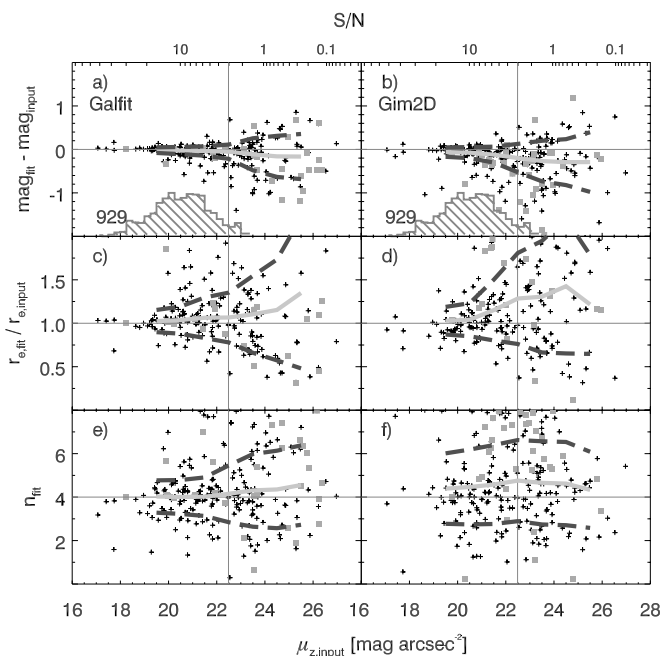

FIG. 12.—Same as Fig. 9, but for the $n = 4$ galaxies in the resimulated GEMS tiles. The histogram in the upper panels shows the surface brightness distribution of the 929 red-sequence galaxies that were selected for analysis in McIntosh et al. (2005).
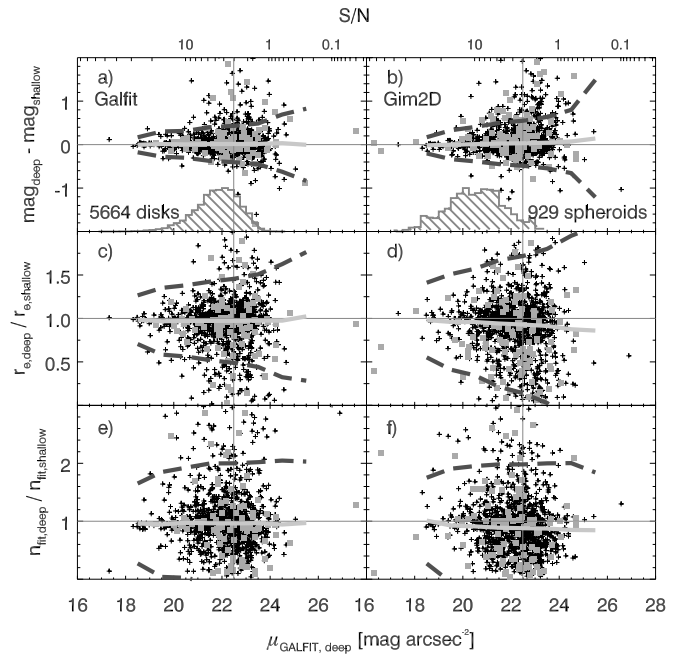

FIG. 13.—Comparison of fits to deep vs. shallow images. The left column shows results using GALFIT, and the right column shows results for GIM2D. Both codes were run on the same sample of real galaxies. Note that in this plot the x-axis shows the GALFIT surface brightness derived from the deeper GOODS data. The y-axis shows the deviations of the three key parameters of the galaxies between the "deep" and the "shallow" fit. In the uppermost plots we again plotted the histograms of the disk galaxy sample of Barden et al. (2005, *left histogram*) and the spheroid-dominated sample of McIntosh et al. (2005, *right histogram*) to highlight the area of parameter space where fitting and independence of the image depth are particularly important.

4.3. Results of Deep-Shallow Tests Using GOODS and GEMS Data

Simulations have the disadvantage that the galaxies have unrealistically simple structure and light profiles that are known a priori to be the same as the profiles used for fitting. Accordingly, in this section we test the performance of the codes on real galaxies. This goal is not straightforward to achieve, inasmuch as one does not know what the real parameters of a given galaxy are or indeed whether or not real galaxies are well described by the Sérsic light profile that was used during our analysis. Instead, we take an empirical approach and test whether the fitting results obtained for real galaxies are sensitive to the image depth by comparing fitting results from the same galaxies in the one-orbit depth GEMS survey and the overlapping five-orbit depth GOODS survey. If they were sensitive to the image depth, it would show that the Sérsic profile is of limited applicability in describing the light profile of real galaxies.

Inspection of Figure 13 shows clearly that *both* codes are reasonably self-consistent when fitting the same galaxies on images of different depth; i.e., neither GALFIT nor GIM2D depends strongly on image depth. While robustness to image depth does not imply that the fitting results are necessarily correct, it does give confidence that issues such as low surface brightness disks missing from shallow *HST* imaging, departures from Sérsic profiles at fainter surface brightness levels, etc., do not appear to seriously compromise the reliability of fitting parameters in one-orbit depth *HST* ACS data.

4.4. Error Estimations from GIM2D and GALFIT

It is interesting to consider if the internal error estimates from GALFIT and GIM2D are reasonable reflections of the more
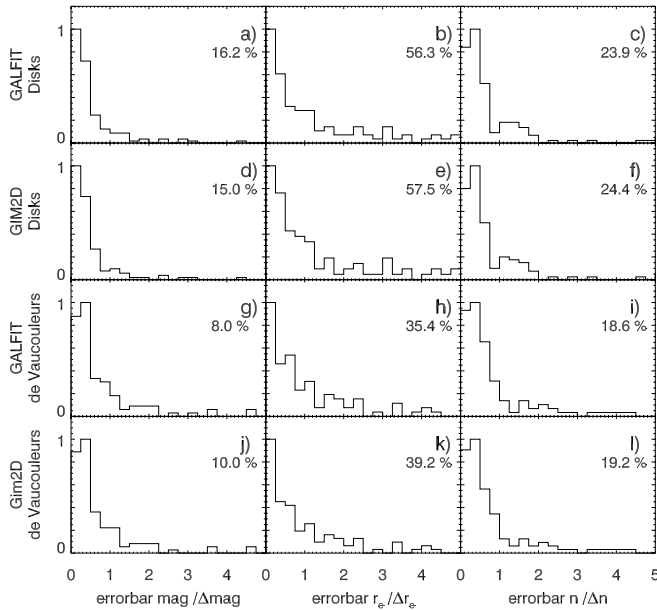
Fig. 14.—Comparison of estimated and real errors for two simulated samples of disk ($n = 1$) galaxies and spheroidal ($n = 4$) galaxies. Shown is the histogram of the error bars σ divided by the deviation from the input value Δ. We show σ/Δ instead of the more intuitive quantity Δ/σ, which would show a very wide distribution and effects not as obvious as here. Calculated as σ/Δ, in principle for 68% of all galaxies, this value should be >1. The number given in each plot shows the fraction of galaxies for which this is true. One can easily see that magnitude and Sérsic index error bars are dramatically underestimated by both GALFIT and GIM2D; the r_e uncertainties are significantly better estimated.

realistic uncertainties given by how well the codes recover input parameters for simulated galaxies. In Figure 14, we address this issue by exploring the distribution of the error estimate σ divided by the deviation of the fit result from the true value Δ. One can see a strong peak of values with $\sigma/\Delta \ll 1$, i.e., for these galaxies the deviation Δ is much larger than the error estimate σ.[18] Under the assumption that the error estimates are correct, σ/Δ should be >1 for 68% of the galaxies. Figure 14 shows that $\sigma/\Delta > 1$ for much less than 50% of the cases; i.e., both GALFIT and GIM2D substantially underestimate the true fit uncertainties, clearly indicating that the dominant contribution to fitting uncertainty is not shot and read noise. Instead contamination by neighbors, structure in the sky, correlated pixels, profile mismatch, etc., dominate the errors. Figure 14 shows no difference between the histograms of σ/Δ for GALFIT and GIM2D; i.e., GALFIT and GIM2D both underestimate the errors to a similar extent. Accordingly, in this work and all other GEMS works, we have *not* used the error estimates given by GALFIT or GIM2D on an object-by-object basis; we rely instead on the mean and width of the parameter distributions from Figures 11 and 9 at the surface brightness of the galaxy in question.

The uncertainties given in Table 9 are calculated and interpolated by using the surface brightness μ and the Sérsic index n and the results from the simulated data (see § 6 for details about this procedure).

4.5. Further Considerations

In the course of our preparation of Barden et al. (2005) and McIntosh et al. (2005), we found that there were two additional

---

[18] This behavior was the motivation for plotting σ/Δ instead of the more intuitive quantity Δ/σ.

practical considerations that potential users of GALFIT and GIM2D may wish to consider.

1. GALFIT is substantially less CPU intensive than GIM2D, reducing the cost and time of fitting large data sets.

2. GIM2D, at least in our implementation, failed to return a fitting result reasonably frequently, requiring manual intervention to restart the code. When trying to fit large data sets, we found this to be labor intensive. In contrast, GALFIT treated each fit as an individual task and therefore was run from shell scripts one fit after the other; if GALFIT does not return a fitting result, the script automatically starts the next fit, requiring no interaction from the user.

5. COMPARISON WITH PIGNATELLI ET AL. (2006)

In this paper we present an extensive and thorough test of the two different 2D galaxy fitting codes GALFIT and GIM2D. In this section, we compare our results with Pignatelli et al. (2006), who compared results from testing these two codes with their own one-dimensional profile fitter, GASPHOT.

Pignatelli et al. (2006) concluded that GASPHOT performed substantially better for significantly (realistically) blended objects than either GALFIT or GIM2D. In the course of our testing, we found a number of differences between our analysis and Pignatelli et al., that we felt may significantly affect their conclusions.

1. For the simulations examined in this paper, they used the IRAF task mkobject, which, as we described in § 2, is inexact for the inner pixels of a simulated galaxy light profile. According to initial tests, the differences in the profiles can lead to systematic errors of up to 10%–20% in the fitting parameters using GALFIT due to the lack of oversampling of the inner pixels when using at least our settings of artdata parameters, and depending on the exact profile parameters. As we used our own simulation script in IDL, improving the profile from IRAF might be possible by using different parameter settings.

2. In their paper, Pignatelli et al. (2006) allow the sky value to be fitted as a free parameter for all three codes. We argue in this paper that this is a nonoptimal way to run *any* galaxy-fitting code: not only would one be subject to errors from irregularities around a Sérsic profile but also the tests shown in this paper show that fitting the sky level as an additional parameter leads to significantly worse fits (especially in the case of GIM2D). Estimating a value for the sky *before* running the fitting codes and keeping this value fixed return more accurate galaxy parameter values.

3. Pignatelli et al. (2006) state that all automatic tools are likely to have problems with blended objects. Like them, we find that deblending is necessary when setting up fitting routines. Masking out blended objects, while better than doing nothing at all, still leads to significantly biased results: this appears to lie at the root of GIM2D's difficulties in fitting some simulations (Fig. 10). We find, furthermore, that if one fits multiple galaxies simultaneously (as is recommended when using GALFIT), GALFIT returns stable unbiased galaxy parameters, even in strongly blended cases (and in cases with intermixed $n = 1$ and $n = 4$ galaxies in which $n = 4$ galaxies play an important role and which was not tested in the paper by Pignatelli et al. [2006]; compare to § 4.2 in this work). Their argument that GALFIT does not deal well with blended galaxies is an artifact of the mode in which they chose to use GALFIT, in particular, the lack of simultaneous fitting of neighboring objects.

4. Pignatelli et al. (2006) only show the GIM2D results for $n = 4$ galaxies; according to our tests, these are the hardest galaxies to reliably fit, and showing only those galaxies leads to a

TABLE 9

GALFIT FITTING RESULTS FOR ALL GEMS GALAXIES

R.A. (deg) (1)	DECL. (deg) (2)	TILE (3)	SNUM (4)	GEMS ID (5)	POSX (pix) (6)	POSY (pix) (7)	SKY (counts) (8)	Mag (mag) (9)	r_e (pix) (10)	n (11)	b/a (12)	P.A. image/WCS (deg) (13)	f_{con} (14)	f_{sci} (15)
53.316325......	−28.059058	s9z01A	1	GEMSJ033315.92−280332.6	5752.64	721.57	18.120	13.91 ± 0.01	0.32 ± 0.01	6.63 ± 0.01	0.777 ± 0.011	3.8(1.6) ± 0.1	1	0
53.334888......	−28.062040	s9z01A	2	GEMSJ033320.37−280343.3	3786.09	368.55	18.120	15.57 ± 0.01	0.30 ± 0.02	3.39 ± 0.01	0.938 ± 0.001	84.0(81.8) ± 0.1	0	0
53.340850......	−28.062310	s9z01A	5	GEMSJ033321.80−280344.3	3154.69	337.72	18.056	21.58 ± 0.09	13.93 ± 0.11	0.37 ± 0.26	0.745 ± 0.056	21.4(19.2) ± 4.4	1	1
53.366943......	−28.062952	s9z01A	6	GEMSJ033328.07−280346.6	391.45	266.85	18.355	24.14 ± 0.10	5.21 ± 0.12	0.97 ± 0.27	0.719 ± 0.059	−47.8(−50.0) ± 4.6	1	1
53.340514......	−28.061939	s9z01A	9	GEMSJ033321.72−280343.0	3190.41	382.16	18.149	21.13 ± 0.32	67.08 ± 0.35	5.36 ± 1.02	0.558 ± 0.054	−60.0(−62.2) ± 5.8	1	0
53.370589......	−28.062339	s9z01A	10	GEMSJ033328.94−280344.4	5.61	341.22	18.426	23.22 ± 0.12	19.79 ± 0.14	0.66 ± 0.29	0.302 ± 0.064	−77.1(−79.3) ± 5.1	1	0
53.342879......	−28.061422	s9z01A	11	GEMSJ033322.29−280341.1	2940.04	444.77	18.098	23.05 ± 0.11	14.03 ± 0.13	1.00 ± 0.28	0.619 ± 0.063	−38.8(−41.1) ± 5.1	1	1
53.342659......	−28.060898	s9z01A	12	GEMSJ033322.24−280339.2	2963.58	507.56	18.171	20.05 ± 0.01	0.30 ± 0.17	1.17 ± 0.92	0.140 ± 0.001	52.4(50.1) ± 0.4	1	0
53.347315......	−28.061235	s9z01A	13	GEMSJ033323.36−280340.4	2470.43	468.25	18.223	22.96 ± 0.08	6.30 ± 0.10	1.18 ± 0.27	0.804 ± 0.051	57.5(55.2) ± 4.1	1	1
53.341322......	−28.059825	s9z01A	14	GEMSJ033321.92−280335.4	3105.38	636.05	18.144	21.56 ± 0.10	43.40 ± 0.12	0.45 ± 0.27	0.123 ± 0.059	50.7(48.4) ± 4.7	1	1

NOTES.—This table shows the first 10 objects of the GEMS fitting results published in the online version of this paper. For every object that was found by SExtractor we give R.A., decl., the GEMS tile name, the SExtractor number of that object, the GEMS ID (containing R.A. and decl.), as well as the x- and y-position on the GEMS tile, so that identification of objects is easily possible. Furthermore, we give the isophotal sky value that was used during the fit and the fitting results of the five key parameters, mag (in apparent F850LP-band magnitudes), r_e (in pixels of 0.03″), Sérsic index n, axis ratio b/a, and position angle P.A. in respect to the image (counted counterclockwise from vertical line according to the convention within GALFIT) and the WCS (in parentheses from north through east). Uncertainty estimates are derived from the simulations, following § 4.4. The last two columns show two different flags showing which fits ran into fitting constraints ($f_{con} = 0$) and which ones would make it into the final galaxy sample used for science according to the selection criteria given in Barden et al. (2005) ($f_{sci} = 1$). Table 9 is available in its entirety in the electronic edition of the *Astrophysical Journal Supplement*. A portion is shown here for guidance regarding its form and content.
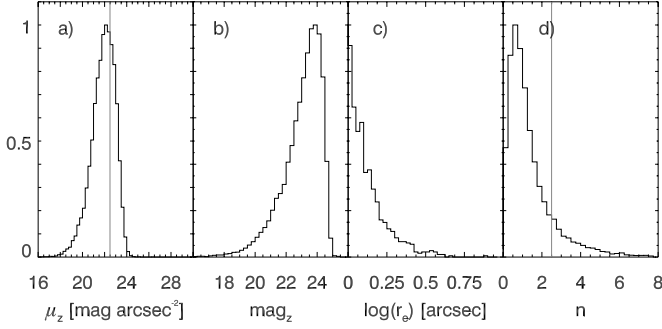
Fig. 15.—Histograms for 23,187 objects from the GEMS survey that were fitted without running into constraints when fitted by GALFIT and selected according to the selection criteria in Barden et al. (2005). From left to right we show surface brightnesses μ, apparent magnitudes, apparent sizes r_e (*logarithmic scale*), and Sérsic indices n. For comparison we plotted the surface brightness of the sky as a vertical line in the leftmost plot and the cut of $n = 2.5$ in the rightmost plot, which is frequently used to distinguish between disk- and bulge-dominated galaxies in an automated fashion.

false impression of the frequency and severity of GIM2D's difficulties with nearby neighbors. Also, it seems that Pignatelli et al. have used the standard setup for GIM2D, which, according to our tests, behaves poorly for faint galaxies; the influence of this decision on their fitting results is unknown.

6. GEMS GALFIT RESULTS

In this section, we present the GALFIT F850LP-band fitting results of all 41,495 GEMS objects that were found by SExtractor. We include fit results for all *unique* objects, be they stars or galaxies. Some objects appear on two or more GEMS frames; in this case the fit results for the images lying furthest from the frame edge was taken. Table 9 shows the first 10 objects in the catalog and gives the format of the catalog published in

the electronic version of this paper. It includes the following values:

1. R.A. (1), Decl. (2): right ascension and declination, given by SExtractor (J2000.0).
2. Tile (3): the GEMS tile in which the galaxies appears.
3. Snum (4): the SExtractor catalog number of this object.
4. GEMS ID (5): the identification of the galaxy within the GEMS project.
5. PosX (6), PosY (7): the position (pixels) of the galaxy in this given GEMS tile.
6. Sky (8): The background pedestal as returned by GALAPAGOS and used during the fit with GALFIT.
7. The GALFIT results (9)–(13): magnitude, half-light radius r_e, Sérsic index n, axis ratio b/a, and position angle (both with respect to the image, PA_{im}, and with respect to the World Coordinate System [WCS], defined north-to-east, PA_{WCS}), respectively, as well as their "uncertainties." These uncertainties are *not* the error bars returned by GALFIT; as shown in § 4.4, these error bars do not reflect the true uncertainty of the fit. We use a statistical method to derive the error estimates from our simulations. We first estimate from our simulations the scatter of the distribution (of the $n = 1$ and $n = 4$ galaxy sample, respectively) at the given surface brightness μ of the real object for $n = 1$ and $n = 4$ simulations. Then, we perform a linear interpolation between the $\sigma(n = 1, \mu = \mu_{obs})$ and $\sigma(n = 4, \mu = \mu_{obs})$ to estimate $\sigma(n = n_{obs}, \mu = \mu_{obs})$. We do not extrapolate; galaxies with $n < 1$ are given the value of the $n = 1$ sample, $n > 4$ galaxies the value of the $n = 4$ sample. We further adopt a minimum uncertainty for each fitting parameter (0.01 mag for mag, 0.01 pixels for r_e, 0.01 for n, 0.001 for b/a, and $0.1°$ for PA). In the table published online and on the GEMS Web site, the uncertainties are stored in extra columns.

8. f_{con} (14): A flag showing which fits ran into any of the fitting constraints (0: fit ran into constraint, 1: fit did not run into any of the constraints).
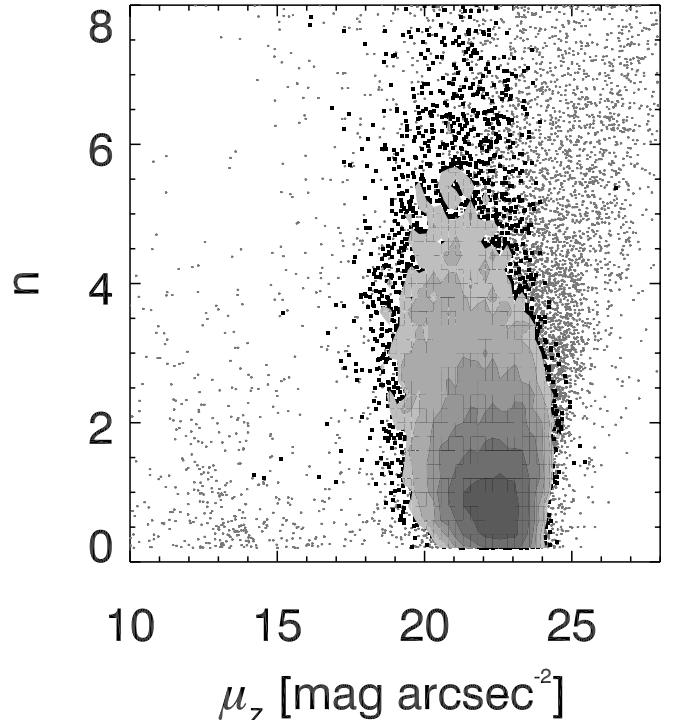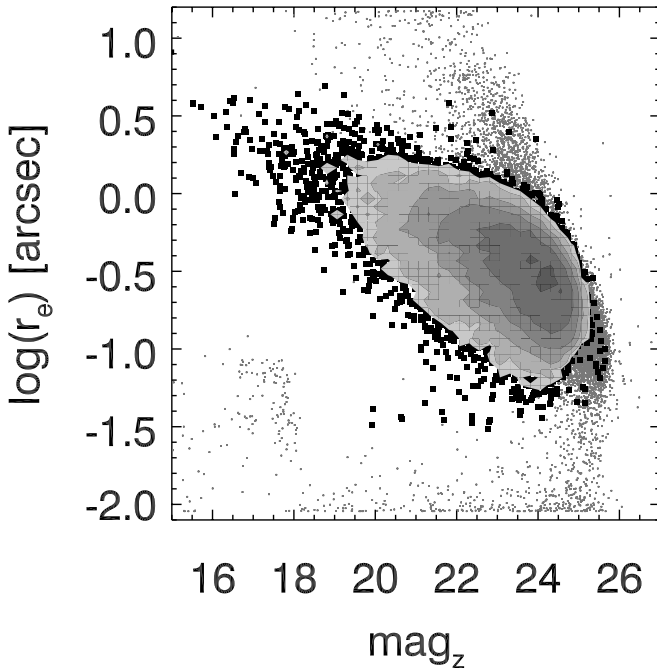

Fig. 16.—Parameter distributions for all galaxies in the catalog published in Table 9, excluding those in which the fit ran into any of the fitting constraints. Galaxies plotted in gray do not end up in the science sample according to the selection criteria used in Barden et al. (2005). Galaxies plotted in black and indicated by the contours pass this selection. In both plots one can see that the galaxies thrown out are mostly faint, low surface brightness galaxies.

9. $f_{\rm sci}$ (15): A flag showing which galaxies would be selected according to the selection criteria given in Barden et al. (2005) (0: object would not be selected for analysis, 1: object would be selected for analysis). The primary effect of the selection is to discard stars and very low surface brightness objects.

As is clear from Figures 15 and 2, the catalog has strongly varying completeness, primarily as a function of surface brightness. Many applications of the GEMS catalogs require a good understanding of these completeness properties. In Barden et al. (2005) and McIntosh et al. (2005) we used the simulations presented in this paper to quantify the effects of completeness. Accordingly, we have made extensive suites of simulation catalogs available to interested users on the GEMS Web site to allow detailed examination of systematic errors in fitting and sample completeness. These issues are discussed in substantially more detail in Rix et al. (2004), Barden et al. (2005), and McIntosh et al. (2005).

Figure 16 shows the parameter distribution of the subset of 34,638 objects for which the fit did not run into fitting constraints. Galaxies plotted in black and indicated by the contours would pass the selection in Barden et al. (2005). One can see that galaxies discarded (*gray*) are mostly faint, low surface brightness galaxies. Another important class of objects thrown out of the sample are objects with either very small sizes or relatively small sizes at high magnitudes. These are identified as stars (or saturated stars) by the automated selection criteria in Barden et al. (2005). Although all these objects are still included in Table 9, one should be very careful when using their fitting results. All these galaxies are indicated by $f_{\rm sci} = 0$.

Figure 15 shows histograms of the most important parameters (surface brightness μ, apparent magnitude, apparent size r_e, and Sérsic index n) for the subset of 23,187 objects that would be selected according to the cuts given in Barden et al. (2005).

7. CONCLUSIONS

In this paper, we have tuned and tested two parametric galaxy fitting codes, GALFIT and GIM2D, for fitting single Sérsic light profiles to both simulated and real data. Our conclusions are the following:

1. The performance of both GALFIT and GIM2D is a strong function of how the codes are set up; in particular, studies using different setups of parametric fitting codes may arrive at different conclusions about those codes if not properly or optimally used.

2. The recommended setup of GIM2D, using `dobkg=yes` and `initparam=yes` is unable to recover the input parameter values of simulated $n = 4$ galaxies that were fainter than the sky surface brightness. We *strongly* discourage users from using these settings, but to instead fix the background to the value local to each galaxy as given by SExtractor and to input very wide model parameter limits. This is very important if the SExtractor segmentation map does not represent the true extent of a galaxy, as was the case for galaxies below the sky surface brightness when using standard SExtractor configurations.

3. Both codes are able to fit (at least bright) $n = 1$ galaxies reasonably well with relatively little bias (Fig. 8). Concentrated $n = 4$ galaxies are substantially less straightforward to fit, owing to the large amount of light in the faint outer parts of the galaxies. For bright $n = 4$ galaxies, the behavior of GALFIT is better than that of GIM2D (Fig. 9); however, parameters returned by GIM2D are still not significantly biased. For galaxy populations and clustering typical of medium-depth cosmological *HST* surveys, there are no large differences between results obtained using GIM2D and GALFIT for these bright galaxies. For fainter galaxies, the performance of GALFIT is substantially better than that of GIM2D. In the set of realistically mixed simulations (Figs. 11 and 12) of $n = 1$ and $n = 4$ galaxies, representing resimulated GEMS tiles, one can see significantly different behavior of the two codes, especially in the recovery of the Sérsic index. GIM2D results are systematically biased to higher Sérsic indices, which, in automated galaxy classification using the Sérsic index to distinguish early- from late-type galaxies, would lead to systematic misclassification of a subsample of (faint) galaxies.

4. The error bars given by both codes underestimate the true uncertainty of the fit by a large factor. One has to use a different approach to derive more realistic error bars.

5. Our testing demonstrated that how a code treats neighboring galaxies can be of great importance. GIM2D only masks out neighbors, which in the tests we ran could lead to poor fitting results for strongly blended objects. GALFIT, in contrast, is able to simultaneously fit many objects and when used in that mode seems to be relatively robust to contamination by neighbors. For this reason, we caution users interested in strongly clustered galaxies against using GIM2D without extensive prior testing.

6. Both GALFIT and GIM2D are self-consistent and show no discernible dependence on image depth when comparing fitting results from GEMS and GOODS data.

7. Our tests on deep and shallow data show that real galaxies are indeed reasonably well described by general Sérsic light profiles.

8. GALFIT works best using an isophotal sky value given by GALAPAGOS. If this is not possible, using GALFIT to internally derive a sky value is significantly better than fixing the sky to a local value given by SExtractor.

9. On the whole, we would tend to recommend GALFIT for single Sérsic profile fitting in medium-depth *HST* ACS data, as GALFIT results are not only somewhat more reliable in the mean but also have lower scatter and less sensitivity to contamination by neighbors than GIM2D.

REFERENCES

Abraham, R. G., Tanvir, N. R., Santiago, B. X., Ellis, R. S., Glazebrook, K., & van den Bergh, S. 1996, MNRAS, 279, L47
Barden, M., et al. 2005, ApJ, 635, 959
Bell, E. F., et al. 2004, ApJ, 608, 752
Bershady, M. A., Jangren, A., & Conselice, C. J. 2000, AJ, 119, 2645
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Blanton, M. R., et al. 2003a, ApJ, 592, 819
———. 2003b, ApJ, 594, 186
Caldwell, J. A. R., et al. 2007, ApJS, in press (astro-ph/0510782)
de Jong, R. S. 1996, A&AS, 118, 557
de Vaucouleurs, G. 1948, Ann. d'Astrophys., 11, 247
Freeman, K. C. 1970, ApJ, 160, 811

Giavalisco, M., et al. 2004, ApJ, 600, L93

Graham, A. W., Driver, S. P., Petrosian, V., Conselice, C. J., Bershady, M. A., Crawford, S. M., & Goto, T. 2005, AJ, 130, 1535

Hogg, D. W., et al. 2004, ApJ, 601, L29

Jahnke, K., et al. 2004, ApJ, 614, 568

Lilly, S., et al. 1998, ApJ, 500, 75

Lotz, J. M., Madau, P., Giavalisco, M., Primack, J., & Ferguson, H. C. 2006, ApJ, 636, 592

McIntosh, D. H., et al. 2005, ApJ, 632, 191

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002, AJ, 124, 266

Petrosian, V. 1976, ApJ, 209, L1

Pignatelli, E., Fasano, G., & Cassata, P. 2006, A&A, 446, 373

Pizagno, J., et al. 2005, ApJ, 633, 844

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1997, Numerical Recipes in C (Cambridge: Cambridge Univ. Press)

Ravindranath, S., et al. 2004, ApJ, 604, L9

Rix, H., et al. 2004, ApJS, 152, 163

Ryden, B. S. 2006, ApJ, 641, 773

Sargent, M. T., et al. 2007, ApJS, 172, 434

Schade, D., Barrientos, L. F., & Lopez-Cruz, O. 1997, ApJ, 477, L17

Schade, D., et al. 1999, ApJ, 525, 31

Scoville, N., et al. 2007, ApJS, 172, 1

Sérsic, J. L. 1968, Atlas de Galaxias Australes (Córdoba: Obs. Astron.)

Shen, S., Mo, H. J., White, S. D. M., Blanton, M. R., Kauffmann, G., Voges, W., Brinkmann, J., & Csabai, I. 2003, MNRAS, 343, 978

Simard, L. 1998, in ASP Conf. Ser. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP) 108

———. 2002, ApJS, 142, 1

Simard, L., & Pritchet, C. J. 1998, ApJ, 505, 96

Trujillo, I., Graham, A. W., & Caon, N. 2001, MNRAS, 326, 869

Trujillo, I., & Pohlen, M. 2005, ApJ, 630, L17

Trujillo, I., et al. 2004, ApJ, 604, 521

———. 2006, ApJ, 650, 18