Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by

Diplom-Physiker Boris Häußler
born in Karlsruhe-Durlach, Germany

Oral examination: 7 February 2007

Galaxy Evolution in the last 9 Billion Years

Referees: Prof. Dr. Hans-Walter Rix

Prof. Dr. Immo Appenzeller

Zusammenfassung — Abstract

Galaxien-Entwicklung der letzten 9 Milliarden Jahre

Ich benutze *HST*-Bilder des Projektes GEMS und photometrische Rotverschiebungen von COMBO-17, um die Rolle von blauen elliptischen Galaxien (BSGs) und die Entwicklung von scheiben-dominierten Galaxien seit einer Rotverschiebung von $z \sim 1$ zu untersuchen. Um meine Galaxienauswahl treffen zu können, habe ich GALFIT und GIM2D, zwei oft angewandte 2-D Modellierungs-Programme, intensiv auf echten und simulierten Daten getestet. Ich zeige, daß GALFIT zuverlässigere Ergebnisse liefert als GIM2D, vor allem durch seine Fähigkeit, benachbarte Objekte gleichzeitig anzupassen. Beide Programme unterschätzen die Fehlerbalken der Parameter dabei stark.

Durch automatische Klassifikation sowie durch Klassifikation per Auge, gewinne ich eine Auswahl an BSGs und zeige, daß nur BSGs mit hoher Massendichte in der Lage sind, sich passiv in typische rote elliptische Galaxien zu entwickeln, und daß ca. 6% der massereichen elliptischen Galaxien blaue Färbung zeigen, in Übereinstimmung mit Vorhersagen aus Zusammenstoß-Raten von Galaxien.

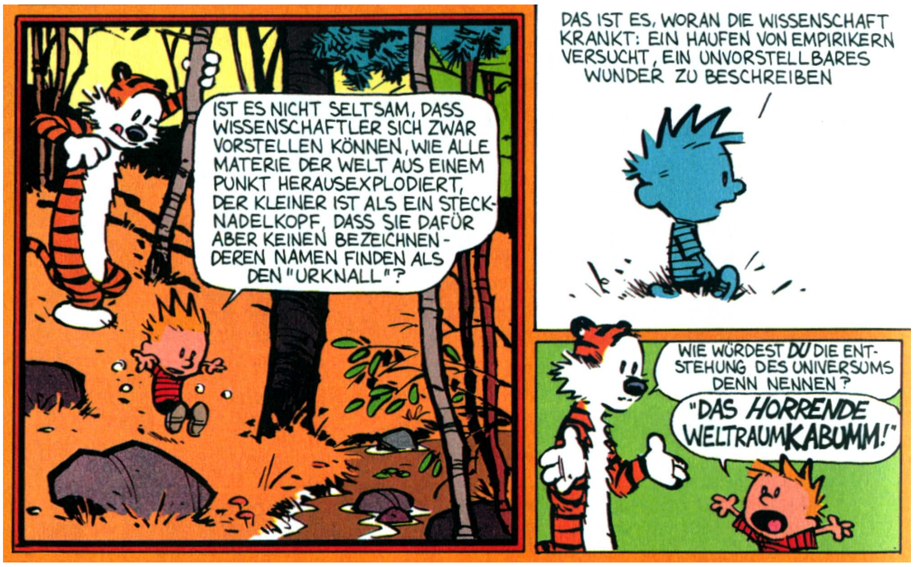
Für Scheiben-Galaxien finde ich eine starke Entwicklung der Helligkeit–Größe-Beziehung von Galaxien mit $M_V \lesssim -20$, entsprechend einer Zunahme der Helligkeit von ca. 1 mag arcsec^{-2} im V-Band seit $z \sim 1$. Nur schwache oder keine Entwicklung der Masse-Größe-Beziehung von Galaxien mit $\log(\mathcal{M}/\mathcal{M}_\odot) \gtrsim 10$ ist sichtbar im gleichen Zeitraum, konsistent mit einem mittleren Wachstum der Scheiben von Innen nach Außen. Scheiben werden umso größer, je massereicher sie werden.

Galaxy Evolution in the last 9 Billion Years

I use *HST* imaging from the GEMS survey and photometric redshifts from COMBO-17 to explore the role of blue spheroidal galaxies (BSGs) and the evolution of disk-dominated galaxies since $z \sim 1$. To be able to define the galaxy samples, I extensively tested GALFIT and GIM2D, two widely used 2-D fitting packages, on both simulated and real data. I find that GALFIT returns more reliable results than GIM2D; in particular due to its ability to simultaneously fit neighboring galaxies. Both codes significantly underestimate the uncertainty of the fit.

I extract a sample of BSG galaxies through automated and visual classification of rest-frame images. I find that only high-mass-density BSGs can passively evolve into typical red-sequence galaxies. I find that the high-density BSGs include $\sim 6\%$ of the massive early type galaxy population, consistent with expectations from published massive galaxy merger rates.

I find strong evolution of the galaxy magnitude–size scaling relation for disk-dominated galaxies with $M_V \lesssim -20$, corresponding to a brightening of $\sim$1 mag arcsec^{-2} in rest-frame V-band by $z \sim 1$. Only weak or no evolution was found in the mass–size relation for galaxies with $\log(\mathcal{M}/\mathcal{M}_\odot) \gtrsim 10$ in the same time, consistent with an 'inside-out' growth of galaxy disks on average. Galaxies increase in size as they grow more massive.


IST ES NICHT SELTSAM, DASS WISSENSCHAFTLER SICH ZWAR VORSTELLEN KÖNNEN, WIE ALLE MATERIE DER WELT AUS EINEM PUNKT HERAUSEXPLODIERT, DER KLEINER IST ALS EIN STECKNADELKOPF, DASS SIE DAFÜR ABER KEINEN BEZEICHNENDEREN NAMEN FINDEN ALS DEN "URKNALL"?
DAS IST ES, WORAN DIE WISSENSCHAFT KRANKT: EIN HAUFEN VON EMPIRIKERN VERSUCHT, EIN UNVORSTELLBARES WUNDER ZU BESCHREIBEN
WIE WÜRDEST DU DIE ENTSTEHUNG DES UNIVERSUMS DENN NENNEN?
"DAS HORRENDE WELTRAUMKABUMM!"

Contents

List of Figures

List of Tables

Chapter 1
Introduction

> *We are probably nearing the limit of all we can know about astronomy.*
>
> SIMON NEWCOMB (MARCH 12, 1835 – JULY 11, 1909)

1.1 Galaxies Types and the Hubble Sequence

Galaxies are the building blocks of the Universe. They contain hundreds of millions or even billions of stars, making them very complex N-body-systems, supported against gravity either by rotation or by random motions. These billions of stars evolve strongly during their lifetime (e. g. populations of young stars being bluer, populations of old stars being redder). Galaxies can also contain large amounts of interstellar gas and dust that is constantly creating new stars and is rebuilt from dying stars. Dark matter is present in galaxies as is known e. g. from rotation curves of spiral galaxies and dominates the dynamics of galaxy clusters and galaxies. Some galaxies inhabit massive black holes in their centers, seriously affecting the dynamics of at least the inner regions of these stellar systems. Magnetic fields, radiation and stellar winds can also influence the properties and evolution of galaxies. Furthermore, 'no galaxy is an island', none lives on **its** own for a (in astronomical sense) very long time, but can (and will) interact with neighboring galaxies, satellites and intergalactic gas. Putting all this together, it becomes obvious why the physics of galaxies is much more poorly understood than the physics of e.g. stars or emission line nebulae and that detailed simulations of galaxies are indeed a very complex and challenging task.

The Hubble-Sequence (see Figure 1.1) was created and first published by Edwin Hubble in his monograph 'The Realms of the Nebulae' [1936] to put some order in the diversity of galaxy morphologies. Up to today, it forms the basis of most galaxy classification schemes.

The original Hubble-sequence arranges galaxies into a continuous sequence of types, showing spiral galaxies (so-called 'late-types') on the right-hand side and elliptical (spheroidal) galaxies (so-called 'early-types') on the left-hand side of the diagram. The spirals are furthermore divided into two branches named 'normal' and 'barred' galaxies,

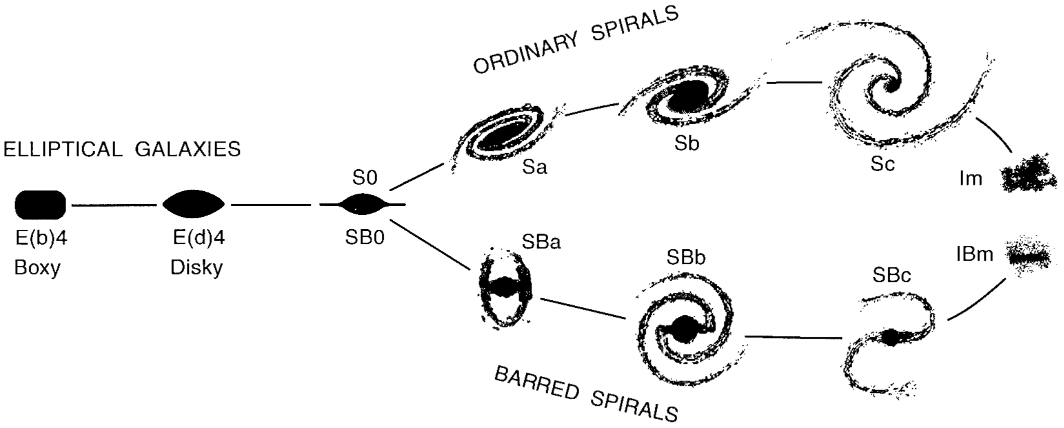


Figure 1.1 The Hubble Sequence of Galaxies as given by Kormendy & Bender (1996). Early-type galaxies are on shown on the left, late-type galaxies at the right.

which is why the Hubble-Sequence is often called the ‘tuning-fork’. The expressions ‘late-type’ and ‘early-type’ are of historical origin, when theorists thought that ‘late-type’ galaxies are ‘late’ in sense of evolution history, so ellipticals evolve into spiral galaxies by the later formation of galactic substructure. Today this picture is known to be incorrect, in fact it is now thought to be the opposite, elliptical galaxies can be created in the merging process of two spiral galaxies (Toomre & Toomre 1972), but the terminology remained from that time.

The Hubble sequence was later revised many times and described e.g. by Kormendy & Bender (1996) who (with others) added a few more types. As shown in Figure 1.1, Irregular galaxies were placed at the very right. These are galaxies that do fit in neither elliptical nor spiral galaxies. Nowadays it is also known that the old sub-classification of ellipticals (E0 to E7) is only a projection effect depending on the viewing angle of the observer, and the classification was revised to now distinguish between galaxies with more box-shape isophots and disk-shape isophots. There is evidence that these two ‘classes’ of galaxies indeed have different building scenarios and therefore form physically different populations (Kormendy & Bender 1996). Additionally, there exists a large population of *dwarf galaxies* with low luminosity and several other categories, e.g. cD galaxies, N galaxies, radio galaxies and so on. I will discuss the most important galaxy types in the following sections.

Although the Hubble sequence is a beautiful and very useful tool to put an order in the variety of billions of galaxies seen on the sky, it is still an empirical way to classify galaxies and it is not clear whether it has any physical significance. What made some collapsing gas clouds turn into elegant spiral systems like our Milky Way, whereas others became smooth, featureless ellipticals? Are there significant differences in the stellar population between the different galaxy types? Does the Hubble sequence reflect a sequence in star formation activity? The next subsections will address these questions and will try to highlight, summarize and explain the differences seen.

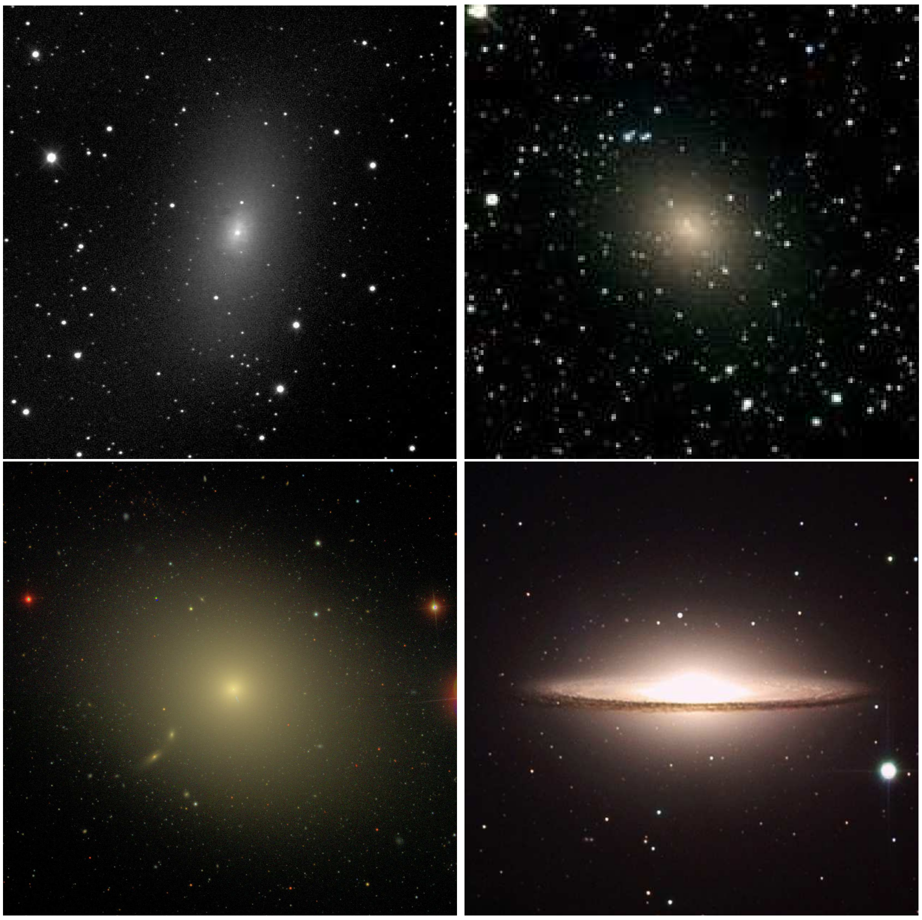


Figure 1.2 The upper left image shows M110, a typical elliptical galaxy, the upper right galaxy is NGC 185, a dwarf elliptical. In both galaxies, one can see nearly no internal structure. The lower left images shows M87, the dominant elliptical galaxy in the Virgo cluster, therefore also known as VirgoA. It contains a jet in the very center of the galaxy, which is not apparent in this image. M87 is the standard example for cD galaxies. Finally, the lower right galaxy is M104, the famous Sombrero galaxy. It represents the standard type of a S0 galaxy with both prominent bulges and disks.

1.1.1 Elliptical Galaxies E

These galaxies (see Figure 1.2), also called ‘spheroidal’ galaxies, show little internal structure. Due to the random motion of their stars, they appear elliptical on the sky as if they were spheroids or ellipsoids in space. Elliptical galaxies range from the most luminous galaxies known today, having $M_B \sim -24$ to dwarf galaxies (dE), so dim that, with today’s telescopes and instruments, they can only be observed in the local group of galaxies. In the original Hubble sequence, the observed ellipticity of the galaxy was included in the morphological designation with E0 being the most circular galaxies and E7 being the most elliptical galaxies, having an axis ratio $b/a \sim 0.3$. Although this might be thought to be relatively straightforward, **rotation** is not the reason for the flattening of the elliptical galaxies. Observations of rotation and velocity dispersions (Bertola & Capaccioli 1975) show that ellipticals rotate too slowly to explain the flattening. It can be shown that therefore the assumption of axissymmetric spatial distribution and isotropic velocity distrubution of stars at all points within the galaxy must be wrong. This means, massive elliptical galaxies must be triaxial systems, which show 3 unequal axes and consequently with anisotropic stellar velocity distributions.

Ellipticals 'flatter' than E7 show a distinct disc and bulge and are therefore classified as lenticular (S0) rather than E galaxies (see below for details about this class of galaxies).

The light profile of elliptical galaxies was thought to be well fit by a 'deVaucouleurs-profile', often also called $r^{1/4}$ law (de Vaucouleurs 1948).

$$log_{10}\left[\frac{I(r)}{I(r_e)}\right] = -3.3307\left[\left(\frac{r}{r_e}\right)^{1/4} - 1\right]$$

with r_e being the 'halflight-radius' within which half of the total flux is emitted and $I(r_e)$ being the surface-brightness at that radius. This expression is a good representation of the luminosity profile over many orders of magnitude in surface-brightness.

In fact, this is not entirely true. More generally speaking, ellipticals have light profiles corresponding to the more general Sérsic law (Sersic 1968, also see section 3.1):

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

where n, the so-called Sérsic index, defines the **shape** of the profile (with $n = 4$ this corresponds to the deVaucouleurs profile given in 1.1.1) and the parameter κ is a normalization factor closely connected to n (for details see e. g. Peng *et al.* 2002). The Sérsic index n for elliptical galaxies shows a relatively wide distribution around a value of 4, but also values of 2 and 6 are apparent. The exact value of the Sérsic index seems to be correlated to the absolute magnitude of the galaxy. The brighter the galaxy, the higher is n (the more compact is the central part of the profile, Caon *et al.* 1993). It is not entirely understood where this correlation comes from, it is subject of current research.

There are two contradicting scenarios of how elliptical galaxies are formed. One of them pictures their formation during a monolithic collapse and formation in a single burst (e.g. Eggen *et al.* 1962; Larson 1974); the other explains the formation of ellipticals as the end product of galaxy merging in a cosmological context (e.g. Toomre & Toomre 1972; Khochfar & Burkert 2005; Kauffmann & Charlot 1998) (please see §4 for a more detailed discussion). For over 30 years, theoretical N-body simulations (Naab & Burkert 2001; Toomre & Toomre 1972; Barnes 1992) support the view that elliptical galaxies form from merging galaxies. After two galaxies collided, they form an elliptically shaped galaxy with nearly no star-formation after the merger, but all their gas is transferred into stars during the merging event itself. This way, galaxy mergers themselves should show high star-formation rate and the star population in the created elliptical galaxy gets older giving

Figure 1.3 These pictures show different ‘types’ of spiral galaxies. The upper left plot shows NGC1232, a beautiful face-on spiral. One can clearly see the spiral arms, the central bulge and the star-forming regions. The upper right image shows M64, the ‘Black Eye’ galaxy. It’s a very nice and extreme example for dust in spiral galaxies. The lower left shows NGC1300, a galaxy with a prominent bar in the center of the galaxy. The spiral arms start at both ends of the bar. The lower right image finally shows NGC891, a spiral galaxy seen edge-on. One can see the dustlane blanking out the background stars.

the elliptical galaxy a red color, which indeed is seen in observations (for further details about this model see chapter 4).

1.1.2 Spiral Galaxies S,SA,SB

Spiral galaxies (see Figure 1.3) show clear internal structure like e. g. spiral arms and/or bars, star-forming regions, and/ or dustlanes. Their stars rotate around the galaxy center on almost circular orbits and reside mainly in a thin disk, so depending on the galaxy orientation in respect to the observers line-of-sight, they appear different. If one looks at them face-on, they appear round and spiral-arms and HII regions can clearly be seen. If one looks at them edge-on, you can see the dust-lane in the disk of at least more massive galaxies.

In many cases the spiral pattern is double with some degree of symmetry, but many more configurations of spiral structure are known (one spiral arms, three or more spiral arms, spiral arms that split up into several arms at some radius, ...). There are two classes of spiral galaxies. The 'normal' spiral galaxies (SA) can be decomposed into a central bulge component, similar to an elliptical galaxy in color and light profile, and a disk component which contains the spiral arms (Peletier *et al.* 1999; Jablonka *et al.* 1996). In the case of 'barred' galaxies, there are a bulge, a disk and additionally a highly elongated concentration of light in the central parts, termed a 'bar'; the spiral arms starting from the ends of the bar (see lower left panel of Figure 1.3). Although non-barred spirals are called 'normal', there are just as many barred galaxies as non-barred, and just as many spirals just between the two extreme cases. Spiral galaxies are classified as Sa, Sb, Sc according to 3 criteria, the openness of the winding of the spiral arms, the degree of resolution in the spiral arms and the size of the central bulge component and/or bar.

- Sa galaxies have tightly wound spiral arms. They are smoothed out, so clumpy structures like star clusters can not be seen. The central bulge or bar is very bright and dominant and shows no structure.
- Sb galaxies have more open spiral arms. They appear a bit clumpy. The bulge component is smaller than in Sa galaxies.
- Sc galaxies have very open spiral arms. They are patchy and can be resolved into star clusters and regions of ionized hydrogen. The bulge is very small. In barred galaxies, the bar is resolved into star clusters and HII regions, but is smaller than in Sa or Sb galaxies.
- Some galaxies with nearly chaotic structures, that were originally classified as very 'late' Sc spirals, are now classified as Sd spirals.

There are also intermediate stages between these classes. They are classified as Sab, Sbc and Scd galaxies. The light profile of the bulges have been found to be very similar to elliptical galaxies of similar luminosities; i.e., they are well-described by a Sérsic profile. The disks have a luminosity profile which may be represented by an exponential light distribution

$$I(r) = I_0 e^{(-r/h)}$$

where h is called the disc scalelength. For the MilkyWay $h \sim 3kpc$. The total luminosity of the disk is then $L = 4\pi h^2 I_0$. This profile can also be described by a Sérsic law with $n = 1$. In fact, the Sérsic profile can be used to automatically distinguish between spiral galaxies and elliptical galaxies. One fits a Sérsic profilfe to the light profile of the galaxy (see chapter 3 for extensive tests of this method) and examines the Sérsic index n. Galaxies with $n \sim 1$ are likely to be disk (spiral) galaxies, values with significantly higher indices, are more likely to be of elliptical type.

1.1.3 Lenticular Galaxies S0 or L

All galaxies with smooth light distribution and axis ratio $b/a < 0.3$ show evidence of a disk-like component. These are called lenticular (lens-like) or S0 galaxies. In some way they are similar to spiral galaxies, their light distribution can be decomposed into a very strong central bulge and an extensive disk. This way they appear intermediate between elliptical and spiral galaxies. Also, there are face-on galaxies, where clear round bulges inhabit inclined disks, that are classified as lenticular objects.

Often, the bulges have a bar-like appearance, so they can be divided into 'normal' and 'barred' lenticulars like spirals. In some galaxies, there is evidence for obscuring matter, often in the form of rings. Lenticulars that are free from obscuring matter are termed as 'early' $S0^-$. 'Later' lenticulars with increasing amount of obscuring matter are called $S0^0$ and $S0^+$.

1.1.4 Irregular and Peculiar Galaxies

They were added to the original Hubble-Sequence later. They are systems lacking both well-defined nuclei and rotational symmetry, and include everything which could not be incorporated into the standard Hubble sequence. Many of them are similar to the Magellanic Clouds, they are known as Irr I or Magellanic irregulars. Other galaxies, that could not be resolved into stars, were classified as Irr II galaxies.

Galaxies with very strange appearances that do not fit into any other class of galaxies along the Hubble sequence are referred to as peculiar galaxies. In a few galaxies, stars seem to form a ring rather than a disk or spheroid. They are known as ring-galaxies. Beautiful examples for this kind of galaxy are the 'Cartwheel' or 'Hoags object' (see Figure 1.4).

Many of the remarkable structures, like long arcs or rings of stars, may be due to gravitational interactions or collisions between galaxies. As already shown by Toomre and Toomre [1974] with computer simulations of close encounters between galaxies, merging events give rise to asymmetric structures, like e. g. tidal arms. A nice example for this is the Antennae-galaxy (see Figure 1.4), which has very long tidal tails. As already mentioned above, such structures can form if two spiral galaxies pass close to each other on prograde orbits, that means, the rotation axis of the galaxies are parallel and also parallel to the rotation axis of the galaxies rotating around the center of mass. The gravitational interaction strips of the outermost stars of both galaxies so they are thrown apart. Similar structures were found in simulations made by Naab & Burkert (1999) and (Mihos *et al.* 1995).

Ring-galaxies (see Figure 1.4) can appear when a smaller galaxy falls face-on through the center of a bigger galaxy. Then a circular shockwave in the big galaxy compresses the interstellar gas to high-enough densities so that the star formation is greatly enhanced. Therefore new born stars appear as a ring surrounding the galaxy center. But this interaction is not strictly necessary to create a ring galaxy. Some appear like rings, although there seems to be no interaction (like Hoags object).

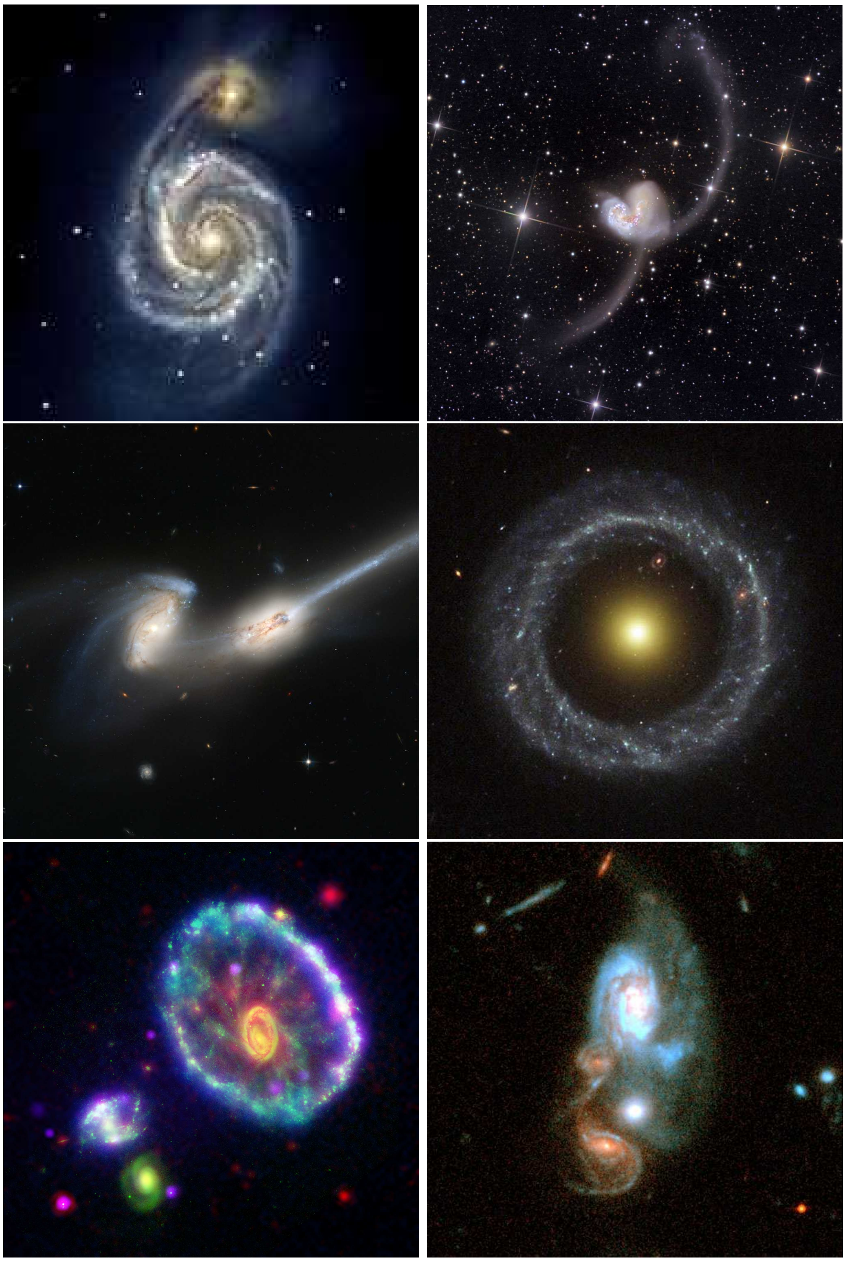


Figure 1.4 The first three pictures show famous interactions of galaxies: M51 (The ‘Whirlpool’ galaxy), NGC4038&4039 (The ‘Antennae’) and the ‘Mice’. The next two show ring galaxies: ‘Hoag’s object’ & the ‘Cartwheel’. They are created when a smaller galaxy falls face-on through a bigger galaxy creating a ring-shaped shockwave that triggers star-formation. The last picture shows a spectacular part of the GEMS survey, showing two interactions that by chance appear one behind the other.

Because of the compression of interstellar gas (ISM[1]), the magnitude of merging galaxies might change. When ISM is compressed, two things happen: The star formation rate is increased (therefore some mergers appear brighter than unmerging galaxies), and the ISM forms dust (this dims galaxies, especially in optical wavelengths, so these mergers do not appear brighter than other galaxies). Mergers are among the most luminous objects in the universe in bolometric terms, especially they are very bright in far-infrared band.

Collisions between galaxies also play a central role in models of galaxy formation. In the cold dark matter scenario, galaxies are build up by the process of hierarchical clustering in which larger galaxies are formed by the coalescence and merging of smaller galaxies. In this picture, strong gravitational encounters between galaxies are essential in forming the structures we can observe today. Among the present population of galaxies, peculiar and interacting galaxies are rare. It is subject of recent research, whether interaction was more common on earlier ages of the universe and how this rate changed (Patton *et al.* 2002; Lotz *et al.* 2006b).

1.2 Properties of Galaxies Along the Hubble Sequence

The Hubble sequence is not only useful to classify galaxies to different types but seems to have physical significance. Certain independent physical properties are correlated to the position on the Hubble sequence. A number of the more important correlations have been reviewed by Roberts & Haynes (1994). Although there are clear trends, there is also wide dispersion about these correlations:

- *Total Masses and Luminosities.* All galaxy types have huge ranges in stellar mass/luminosity. The most massive galaxies tend to be Ellipticals and S0s. At lower masses, an increasingly large fraction of spiral galaxies is observed. Irregulars tend to have reasonably low masses.

- *Neutral Hydrogen.* There is only very rarely any neutral hydrogen observed in elliptical galaxies, while all spirals and late-type galaxies have significant gaseous masses. In elliptical galaxies, the upper limit of mass of neutral hydrogen corresponds to $M_{HI}/M_{tot} \leq 10^{-4}$ while in spirals the fractional mass of neutral hydrogen ranges from about 0.01 for Sa galaxies up to 0.15 at Sm with monotonic increase between. The fractional mass of neutral hydrogen is more or less independent of the mass of the galaxy.

- *Total surface-brightness and surface density of Neutral hydrogen.* These quantities change in opposite senses along the Hubble sequence. The total surface density, as determined by the total mass of the galaxy and its characteristic radius, decreases

---

[1] Interstellar matter

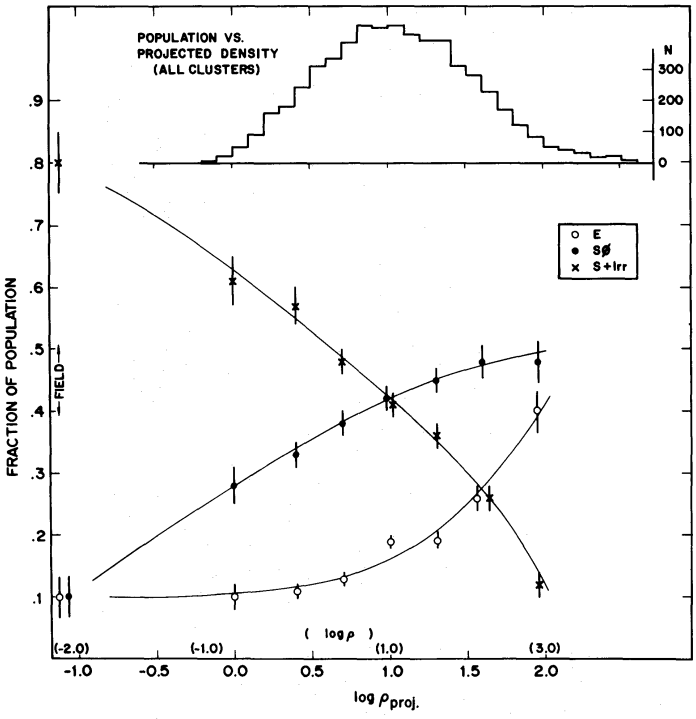


Figure 1.5 The fraction of different morphological types of galaxies found in different galaxy environments according to Dressler (1980). The local number density of galaxies is given as a projected surface density, ρ_{proj} of galaxies, which is $numbers\ Mpc^{-2}$.

monotonically along the sequence (from S0 to Im), whereas the surface density of neutral hydrogen increases.

- *Specific Angular Momentum.* The Specific angular momentum of baryons increases from ellipticals to spirals, leading to the strong flattening of disk galaxies and the shape and light-profile of spheroidals.

- *Specific Angular Momentum.* The Specific angular momentum of baryons increases from ellipticals to spirals, leading to the strong flattening of disk galaxies and the shape and light-profile of spheroidals.

- *Integrated color.* There is a strong color correlation along the sequence. Elliptical galaxies are red (old stars, no star-formation), they basically form the so-called 'red-sequence' of early type galaxies (see Fig. 1.6) in the color-magnitude diagram, late-type galaxies are blue (young stars, ongoing star-formation), they form the 'blue cloud' (The gap in between these two poulations is sometimes referred to as the 'green valley'). Despite this systematic trend there is a significant dispersion about the relation at each point along the sequence. For example, there are Sc galaxies that are red instead of blue. For this thesis, the blue, but morphologically elliptical galaxies are especially important (see chapter 4).

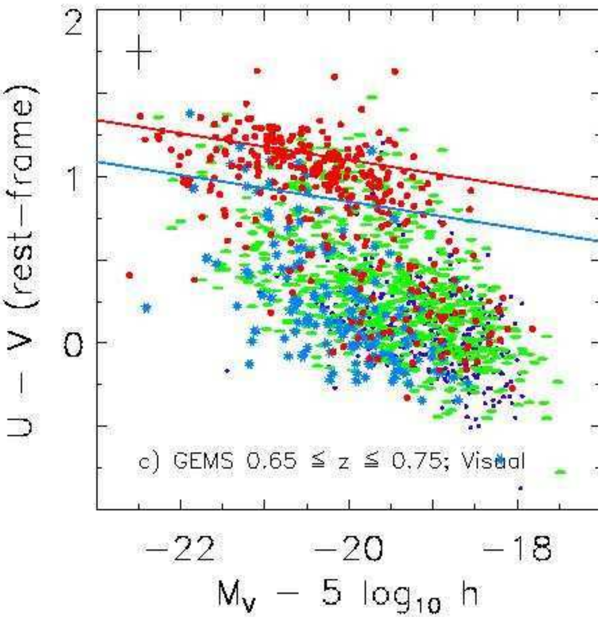


Figure 1.6 The color-magnitude diagram of galaxies as given by Bell *et al.* (2004a). Red circles show visually-classified E/S0 galaxies, green ellipses are Sa-Sm galaxies, blue stars are Peculiar/Strong Interaction galaxies, purple circles are Irregular/Weak Interaction and compact galaxies. One can see the red-sequence of red, mostly spheroidal or bulge-dominated galaxies, also highlighted as the red line, although some fraction of spheroidals with blue colors exist. Furthermore the blue cloud of star-forming spirals and mergers and the 'green valley', indicated by the blue line, are visible.

These correlations are related to different star formation rates in different types of galaxies.

One further well-observed correlation is that early-type galaxies tend to live in high-density environments whereas late-type galaxies prefer low-density environments. Dressler (1980) has plotted the fraction of different galaxy types as a function of the number density of galaxies in which they were found (Figure 1.5). Field galaxies, galaxies that are not member of a galaxy group or cluster, are located to the left of the diagram, while rich clusters of galaxies are towards the right. As one can see, in rich clusters, elliptical and S0 galaxies are much more common than spiral galaxies, while in general field, galaxies are mostly spirals and irregulars. Evidently, the environment of galaxies seems to have big influence on the morphological characteristics. One can understand this in the picture of hierarchical galaxy formation. When elliptical Galaxies form in galaxy mergers, ellipticals should exist more often in dense environments as galaxy interactions are more likely to happen here, which is exactly what is observed.

1.3 Galaxy Evolution from Theory & Observations

Our understanding of cosmology and the universe has been changing drastically in the last century. At the beginning of the 20th century, the universe was still thought to be static. This was the reason why Einstein introduced the cosmological constant in his famous equations to be able to prevent the mass in the universe from rapid collapse through gravity. When it became known that many of the nebulae seen on the sky were truly gigantic self-bound stellar systems of extragalactic origin, and that most of these systems were actually moving away from the observer (the further away, the faster they move away from us, Hubble 1936), it was noticed that the universe was not stable at all but expanding, and Einstein removed the cosmological constant from his equations and called it the 'biggest blunder' of his life. Today, the cosmological constant experiences a revival in the so-called Λ $Cold\ Dark\ Matter$ (ΛCDM) model, a cosmological model that tries to connect cosmology (e. g. Big Bang and Inflation), astrophysics (e. g. gravitational cosmic structure growth, hierarchical clustering) and particle physics (e. g. standard particle physics models) and explains the accelerated expansion of the universe as it is suggested by cosmological observations. A number of free parameters are important to describe this model. Observations of many different cosmological probes are needed to constrain the exact values of these parameters, many of which correlate among them (degeneracy). The ΛCDM model, in which galaxies grow in mass through both star formation and accretion of already-formed stars in galaxy mergers, tries to explain the origin of galaxies, when and how they form and evolve to finally show their present-day properties and the origin of their diversity and scaling laws. What is the dark component that forms the vast majority of the mass we know of today and why do galaxies cluster in space as observed, following a 'sponge-like' structure? By answering these questions we would be able to link between the present-day universe and the properties of the early universe, and to use the observations as laboratories for testing fundamental theories.

Today the cosmological constant Λ is thought to dominate the energy content of the universe by $\sim 2/3$ ('Dark energy'). Only about $1/3$ of the energy content is based on matter, but as is known 85% of this 1/3 is in form of non-baryonic 'Dark Matter' (DM, it's true nature is unclear), which is only known from indirect observations, e. g. lensing studies, see Fig. 1.7. Only 15% of all the mass in the universe is present in stars, dust, gas and baryonic matter in general. The DM interacts only gravitationally with baryons, playing an important role in the clustering of galaxies and matter in general. Baryonic matter falls to the center of the dark matter haloes that dominate the gravitational potential in large regions of the universe, and, through cooling, fragmentation and gas collapse, starts to form stars and galaxies which we can observe directly today. Disk galaxies form inside the growing CDM haloes, the properties of which influence the properties of the formed galaxy, like surface brightness, the shape of the rotation curve, the bulge-to-disk ratio of the galaxy and it's star formation rate and history. As the dark matter is not homogeneously distributed, it can explain e. g. the clustering of galaxies and the large-scale structures seen in large and deep surveys, e. g. voids, sheets and filaments in the distribution of galaxies. In higher-density environments, where DM haloes do not live as

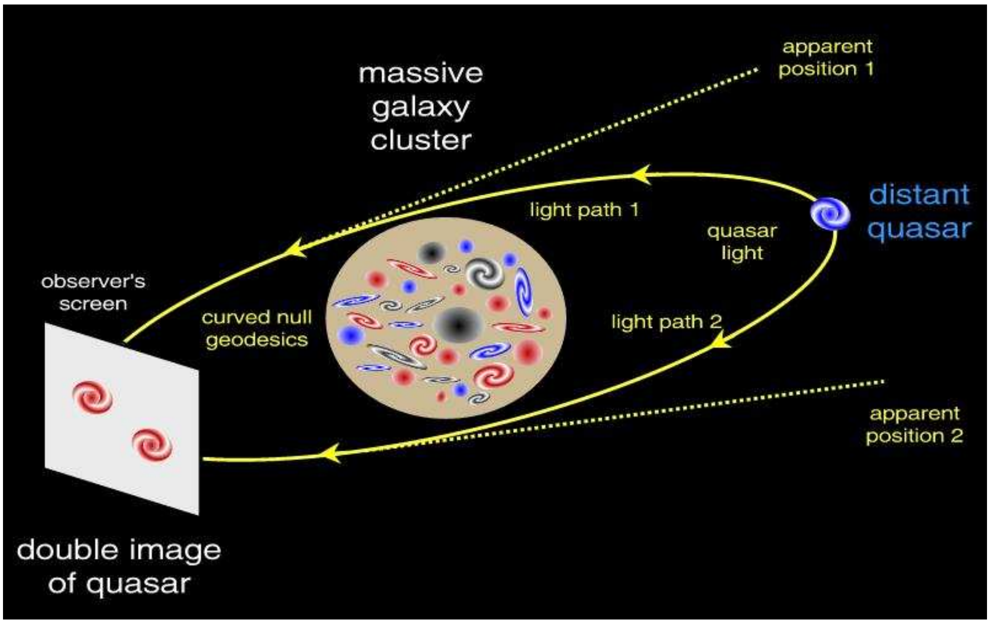


Figure 1.7 This figure explains the principle of gravitational lensing. Light from distant galaxies and QSO's is gravitationally bend around a massive galaxy cluster. This imposes different effects on the QSO image, e. g. shear and distortion. Using many galaxies and statistical methods, one can derive the distribution of the Dark Matter, which makes up most of the mass in the universe. When the lensing is strong, multiply images, or even an 'Einstein-Ring' can be seen if the alignment is perfect and geometry of the lensing object fits strict requirements, e. g. spherical symmetry.

isolated as in 'field' environments where disk galaxies form, they can merge and create massive elliptical galaxies through violent merging of the Dark matter haloes and/or the galaxies therein.

Yet, there clearly is an important contribution of baryonic matter through gas and stars that produce most of the light in the universe that we can observe, mainly through star formation (SF) and feedback mechanisms. Trying to understand the formation and evolution of galaxies and their properties, it becomes obvious that one has to understand both the evolution of Dark matter and the evolution of the content of galaxies, stars, dust and gas. Theoretical models that try to describe this evolution, need to implement both these components and their interaction to be able to reliably predict galaxy properties that can be checked by observations. This task is challenging given the fact that it is known that stars are created from collapsing gas clouds and strongly interact with their environment during their lifetime, by ionizing their surrounding by UV radiation, by winds and by supernovae at the end of their lifetime, but none of these processes is fully understood in detail.

Different hydrodynamical approaches have been made to simulate galaxy properties. In SPH ('Smoothed Particle Hydrodynamics') models, gravitational interaction of particles with both their neighbors and the dark matter particles, as well as the gas pressure are included. Complete, big simulations (e. g. Springel & Hernquist 2003) are very CPU intensive, though, a more efficient, less CPU intensive method would be appreciated. Semi-analytic models (e. g. White & Rees 1978; Somerville & Primack 1999; Kauffmann *et al.* 1993) are a good approach in which simple, yet physically and observationally motivated prescriptions for star formation and feedback replace the equations of gas-dynamics. Semi-analytic hierarchical models, describing the build-up of present-day galaxies through subsequent merging (merger-trees) of many smaller galaxies, can be coupled to population synthesis models in order to predict the luminosities and colors of galaxies as well as the luminosity function and other observed relations of nearby galaxies. One sees from these approaches, that spiral galaxies, like our galaxy, can not be explained with closed-box models (a single burst of star formation and passive evolution). Continuous infall of gas with low metallicity is required to reproduce the colors, metal abundances, star formation rates, and gas fractions of these galaxies. On the other hand, massive ellpitical galaxies can be well described by a single star burst at early times and subsequent passive evolution. To create e. g. the color-magnitude relation of very massive ellipticals observed today, some yet unknown process (feedback) has to be implemented into these models, quenching gas cooling and thus star formation. Without this process, creating the most massive elliptical galaxies would not be possible.

It becomes clearer that any kind of feedback, e. g. enriching the ISM with metals and thus changing the properties of galaxies drastically, is an important mechanism to solve many of the problems encountered in all models when trying to predict parameters of galaxies today, e. g. the overcooling problem, the shape of the luminosity function of galaxies, the metallicity of the inter-stellar medium and the properties of stellar populations of at least some types of galaxies. There are different possible feedback mechanisms to solve these problems, although it is unclear in which fractions the different possible mechanisms play a role, although all seem to be obvious, e. g. feedback from supernovae and/or active galactic nuclei (AGN Di Matteo *et al.* 2005) that inhabit central regions of some galaxies. When these supermassive black holes are fed with infalling material, accretion disks around them release, mainly through powerful jets, the largest amounts of energy known in astronomical objects and thus prevent further gas cooling, which could in fact shut off SF efficiently. This phenomenon was much more frequent in the past than in the present, the high-redshift quasars (QSO's[2]) being the most powerful examples of the AGN phenomenon.

In the past decade there has been huge progress in observational astronomy, opening a new window for the study of galaxy formation and evolution, especially at high redshift, where observations become more and more difficult. The new class of 8-10 meter telescopes with their wide-field multi-object spectrographs and imagers, the *HST* with its high resolution and high sensitivity as well as other observatories in different wavelength regimes, e. g. X-ray, UV, IR and submm allowed observers to get different and more detailed insight into galaxies and their properties at higher redshift. Deep wide-field surveys discovered new populations of galaxies and enabled observers to measure the quantities

---

[2] Quasi Stellar Objects. Very bright objects, that look like stars on astronomical images because of the small size of the light-emitting region

of high-redshift galaxies and global parameters, e. g. the dependance of different galaxy types on environment, the co-evolution of stellar bulges, the supermassive black holes and their host galaxies, the importance of galaxy merging on galaxy evolution and the build-up of the stellar mass, as well as the global cosmic star formation rate. Theoretical models need to successfully predict these quantities and relations, e. g. it is now known from different independent observations, that the global star formation rate rapidly declined by a factor of 5-10 since a redshift of $z \sim 1$, roughly $2/3$ of the stellar mass is formed in a relatively short period before this time, only $1/3$ is formed at later times. The further build-up of the massive galaxies present today, has to happen through major mergers; at $z < 1$ a significant fraction of massive galaxies undergo a major merge as is evident from studies of close-pair fractions, morphologically disturbed galaxies and the build-up of the stellar mass in morphologically early-types.

In this PhD-thesis, I will address different important aspects of galaxy evolution from the GEMS (Rix *et al.* 2004, , please also see §2) survey. In §3 I will present an important and useful test of two widely-used software packages, GALFIT (Peng *et al.* 2002) and GIM2D (Simard *et al.* 2002) to quantify parameter reliability within the data-reduction pipeline used in GEMS. One of the key predictions of hierarchical galaxy formation models, as briefly discussed above, is, that a significant fraction of elliptical galaxies form in late merging events as described above. One of the most important observations of such an assembly is the existence of blue spheroidal galaxies, which have spheroid-dominated morphologies and blue colors indicating recent star formation, as an intermediate step in the evolution of elliptical galaxies. I will address these galaxies in §4 and will try to quantify whether they are likely to be possible remnants of galaxy mergers and the progenitors of typical red-sequence galaxies as they are seen in the local universe. In §5 I will address the size evolution of disk galaxies. Finally, in §6, I will briefly introduce the GEMS 'skywalker', a little PR-program written to be able to access the huge variety of GEMS galaxies through the web with finite web-access.

Throughout this work, I will assume a cosmological model as described by $\Omega_\mathrm{m} = 0.3$, $\Omega_\mathrm{m} + \Omega_\Lambda = 1$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

Chapter 2
GEMS: Galaxy Evolution from Morphologies and SEDs[3]

Abstract

GEMS, **G**alaxy **E**volution from **M**orphologies and **S**EDs, is a large-area (800 arcmin^2) two-color (F606W and F850LP) imaging survey with the Advanced Camera for Surveys on *HST*. Centered on the Chandra Deep Field South, it covers an area of $\sim 28' \times 28'$, or about 120 Hubble Deep Field areas, to a depth of m$_{AB}$(F606W)$= 28.3(5\sigma)$ and m$_{AB}$(F850LP)$= 27.1(5\sigma)$ for compact sources. In its central $\sim 1/4$, GEMS incorporates *ACS* imaging from the GOODS project. Focusing on the redshift range $0.2 \lesssim z \lesssim 1.1$, GEMS provides morphologies and structural parameters for nearly $10,000$ galaxies where redshift estimates, luminosities and SEDs exist from COMBO-17. At the same time, GEMS contains detectable host galaxy images for several hundred faint AGN. This chapter provides an overview of the science goals, the experiment design, the data reduction and the science analysis plan for GEMS.

2.1 Introduction

The formation and the evolution of galaxies are driven by two interlinked processes. On the one hand, there is the dynamical assembly of the mass distribution in the context of dark-matter dominated, hierarchical structure formation. On the other hand, there is the star-formation history (SFH), i.e. the successive conversion of gas into stars, along with the ensuing feedback. By now, the cosmological background model and the corresponding large-scale structure growth are well constrained (e. g. Percival *et al.* 2002; Spergel *et al.* 2003, we will use $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km/s throughout), and the focus is shifting towards understanding on galaxy scales the dynamics and the physics of star formation, reflected in the structure and stellar populations of the resulting galaxies.

On the one hand, a comprehensive picture of galaxy formation must match the population statistics of integrated galaxy properties, e. g. the galaxy luminosity and mass functions or the overall spectral energy distributions (SED), and the dependence of these distributions on the larger environment. But a picture of galaxy formation should also

---

[3] This chapter has been adapted from Rix *et al.* (2004).

explain the internal structure of galaxies: e.g. their size, bulge-to-disk ratio, degree of symmetry, internal population gradients, and nuclear properties. Many of the ingredients that determine the internal structure and the SFH are qualitatively clear. For example, the size of galaxies is linked to the angular momentum of the stars and their progenitor gas, created early on through tidal torquing; spheroid stars formed before or during the last episode of violent relaxation, whereas most disk stars in large galaxies have formed after the last major merger; and major mergers are effective triggers of powerful starbursts, if the progenitor galaxies have a sizeable supply of cold gas. These same mergers are also suspected to trigger nuclear (AGN) activity by funneling gas into the vicinity of the ubiquitous central black holes.

Quantitative theoretical predictions of the resulting internal structure and SFHs of individual galaxies are neither robust nor unique, as galaxy formation involves a vast range of spatial scales, from well below 1 pc to well above 1 Mpc, along with complex geometries. Neither direct numerical simulations (e. g. Katz & Gunn 1991; Steinmetz & Navarro 2002; Springel *et al.* 2001), nor semi-analytic models (e. g. Cole *et al.* 2000; Kauffmann *et al.* 1993; Somerville & Primack 1999) can currently capture all important aspects of the problem. Turning to empiricism in light of this situation, galaxy evolution is perhaps best studied by observing directly how the properties of the galaxy population change with cosmic epoch.

Observational constraints on the galaxy population in the present-day ($z \lesssim 0.2$) universe have drastically improved over the last years, in particular through three large surveys: 2MASS imaging the sky in the near-infrared (Skrutskie 1997), and the optical surveys SDSS (York *et al.* 2000) and 2DFRS (Colless *et al.* 2001). The galaxy luminosity functions, the galaxy (stellar) mass function, the galaxy correlation function, the distribution of spectral energy distributions, the distribution of galaxy sizes, *etc.* have been (re-)determined with unprecedented accuracy (e. g. Blanton *et al.* 2003c; Norberg *et al.* 2002; Shen *et al.* 2003; Kauffmann *et al.* 2003a; Bell *et al.* 2003a; Strateva *et al.* 2001).

The observational challenge now is to come up with a correspondingly more detailed assessment of galaxy properties and galaxy population properties at earlier epochs. Over the last decade the 'look-back' approach to studying galaxy evolution has experienced a number of breakthroughs, both in obtaining distant galaxy samples with secure redshifts to study integrated galaxy properties, and in obtaining high-resolution imaging, mostly with *HST*, to study internal structure (e. g. Ellis *et al.* 1996; Lilly *et al.* 1996; Steidel *et al.* 1996; Williams *et al.* 1996; van Dokkum *et al.* 1998). Yet, for a full comparison with local samples, these data sets — typically a few hundred objects — have been much too small. This holds especially true considering only samples with redshift *and* internal structure information. For one, these samples are too small in number to allow dissecting the galaxy population by redshift, luminosity, color, size or even environment, and still be left with significant subsamples. Previous samples with redshifts and well-resolved images have also been drawn from too small an area. As a consequence, they cannot reflect the 'cosmic average' at any epoch, because luminous galaxies are clustered quite strongly at all epochs (e. g. Giavalisco *et al.* 1998; Phleps & Meisenheimer 2003).

Existing studies of morphology and internal structure have shown that to $z \sim 1$ the sizes and Hubble types of galaxies roughly resemble the nearby universe (e. g. Abraham *et al.* 1996; Lilly *et al.* 1998; Simard *et al.* 1999), whereby the significance of possible differences from $z \sim 0$ (e. g. the higher incidence of distorted morphologies) is weakened by discrepant sample definitions, small sample sizes and survey volumes and by the observational effects of $(1+z)^4$ surface brightness dimming and of bandpass shifting. At $z \gtrsim 2$, galaxy images lose their *prima facie* resemblance to the nearby universe and appear more compact, but there, too, the band pass shifting may give an exaggerated impression of true evolutionary effects of the galaxy population (e. g. Labbé *et al.* 2003).

Recently, the COMBO-17 project (e.g. Wolf *et al.* 2001, 2003a) has afforded a thirty-fold increase by number over the earlier redshift surveys (See also Fried *et al.* 2001; Im *et al.* 2002; Le Fèvre *et al.* 2003; Davis *et al.* 2003, for other recent or ongoing surveys). COMBO-17 incorporates deep ($R \lesssim 23.5$) photometry in 17 optical filter bands, providing redshifts good to $\delta \mathrm{z}/(1{+}\mathrm{z}) \lesssim 0.02$ for both galaxies and AGNs. From a sample of $\gtrsim 25\,000$ galaxies with $z \lesssim 1.2$, the survey has explored the population and integrated properties of galaxies since these redshifts. Building on earlier results, COMBO-17 has detailed and quantified the increasing prominence of massive galaxies without young stars (Bell *et al.* 2004b), the shift of high specific star formation activity to low mass systems (Wolf *et al.* 2003a), and the SED-differential evolution of galaxy clustering (Phleps & Meisenheimer 2003). Furthermore, COMBO-17 has provided a deep insight into the population and evolution of low-luminosity AGNs (Wolf *et al.* 2003b).

Yet, as any other ground-based imaging survey, COMBO-17 could add little to elucidating the evolution of internal structure over this redshift interval. The goal of the present project, GEMS (Galaxy Evolution from Morphologies and SEDs), is to provide an order of magnitude improvement in assessing the evolution of the *internal structure and morphology* of galaxies over the 'last half' of cosmic history (actually the last 8.5 Gyrs to $z \sim 1.2$) through wide-area, high-resolution imaging with *HST*.

Foremost, GEMS should address a) how the galaxy merger and tidal interaction rate evolved since $z \sim 1$; b) which portion of the global star formation rate at any given epoch is externally triggered, through tidal interaction or mergers; c) whether stellar disks grew "inside out"; d) whether the formation of bulges entirely preceded the formation of their surrounding disks; e) whether stellar bars are a recent ($z \leq 1$) phenomenon; f) whether the drastic decay in nuclear accretion activity is reflected in any drastic change of the host galaxy population.

Obviously, we would be most interested in tracing the evolution of individual objects. Yet, only the evolution of population properties is observable. In practice, one tries to bridge this gap and answer the above questions by assessing separately the redshift evolution of various structural parameter relations and of space densities for different galaxy sub-samples: e.g. the relation of disk- or bulge-size *vs.* their luminosity or stellar mass; the space density of large disks; the ratio of disk and bulge stars at different epochs, the fraction of young stars in disks, etc..

GEMS and a number of other cosmological imaging surveys (in particular the narrower, but deeper GOODS survey; Giavalisco & GOODS Team 2003; Ferguson *et al.* 2004; Moustakas *et al.* 2004) with *HST*, have been enabled by the advent of the Advanced

Camera for Surveys (*ACS*, Ford *et al.* 2003), which vastly improves the survey speed of *HST*. The GEMS observations were planned at a time when much research effort in observational cosmology is centered around a number of selected fields, such as the HDF's North and South, or the Chandra Deep Fields, where multiwavelength observations from the X-rays to the far-IR and radio are creating synergies; the GEMS mosaic encompasses such a field, the Chandra Deep Field South.

In the remainder of the chapter we outline the experiment design (§2.2), the initial data reduction (§2.3), the image analysis and initial galaxy catalog (§2.4), and the planned science analysis.

2.2 Experimental Design

The immediate goal of the GEMS survey is to provide high resolution images from which to extract an empirical data base of 'structural parameters' that describe the stellar bodies for a large sample of distant galaxies. Here we outline the rationale for the particular survey implementation.

To resolve the internal structure of galaxies at $z \sim 1$, with expected typical scale-lengths of ~ 2 kpc, one needs imaging at a spatial resolution considerably finer than their apparent size: e.g. 2 kpc project to $0.''26$ at $z = 0.75$. To date *HST* is still far more efficient to deliver this over wide fields over wide fields than AO on large ground-based telescopes, such as afforded by CONICA on the VLT (Lenzen *et al.* 2003).

Sample Size: The *desideratum* is the distributions of galaxy size, light concentration, bulge-to-disk-ratio, and morphology as a function of redshift, luminosity, SED, and perhaps environment. Even considering only one number to characterize the internal structure of galaxies, one needs to estimate the frequency distribution of galaxies in a four-dimensional parameter space , (z,L,SED,structure). For a handful of bins per axis and $\sim$10 galaxies per bin (or S/N ≥ 3), one needs samples of $\sim 10^4$ galaxies.

Choice of Survey Area and Field: To approach representative sampling of environments, the field size must be well in excess of the correlation length of typical (L_*) galaxies (which is ~ 5 Mpc comoving for $0.3 \lesssim z \lesssim 1$ Phleps & Meisenheimer 2003; Coil *et al.* 2004), and even twice as large for red, early type galaxies (e.g Daddi *et al.* 2001). This scale corresponds to $7' - 11'$ at $z = 0.75$, or three times *HST*'s field-of-view in a single pointing ($\sim 3'$) with the *ACS* (Ford *et al.* 2003). The need for large samples with redshifts, faint limiting magnitudes and imaging with *HST*'s restricted field of view lead to a densely sampled, contiguous field. To date, the COMBO-17 survey provides redshifts and SEDs in three disjoint fields Wolf *et al.* (2003a) of $\sim 30' \times 30'$ each, one including the Chandra Deep Field South (CDFS Giacconi *et al.* 2001), which we refer to as the 'extended CDFS' (E-CDFS) area. Note that the results of (Wolf *et al.* 2003a) show, that even for such large field sizes the galaxy population variations due to large scale structure are still very significant, e.g. $> 50\%$ *rms* for luminous red galaxies over redshift intervals of δz~ 0.2.

In part, we chose the E-CDFS because it appeared *a priori* representative with respect to its galaxy population, as opposed to e. g. the Abell 901 cluster field in COMBO-17. But foremost, the field is preferable because of the intense focus of research at other, complementary wavelengths, in particular in X-rays with Chandra and XMM observations (Rosati *et al.* 2002) and in the thermal infrared with upcoming SIRTF observations (GOODS Dickinson *et al.* 2003). As we will detail in §3, GEMS imaging is coordinated with the multi-epoch GOODS imaging over the central $\sim$ 25% of the total GEMS area. The GEMS survey area and its spatial relation to the GOODS and COMBO-17 field is illustrated in Figure 2.1. The central co-ordinates of the COMBO-17, and hence GEMS field are α=03h 32m 25s, $\delta = -27\ 48'\ 50''$ (2000).

Flux Limit and Filters: To reach 'typical' galaxies (L $\sim L_*$) to redshifts of z$\sim$1, one needs samples with redshifts to a magnitude limit of $m_R \sim$ 23.5 (Wolf *et al.* 2003a). The GEMS imaging depth was designed to permit robust galaxy model fits for most objects that are in the COMBO-17 redshift sample, $m_R \leq$ 23.6. To get S/N $\geq$ 20 on extended objects near this magnitude requires about one orbit of exposure time with the *ACS* in F850LP.

The structural parameters (de Jong 1996; Kranz *et al.* 2003) and morphology (see Rix & Rieke 1993), especially of late types galaxies, depend on the observed wavelength. Therefore, one must study morphology evolution at comparable rest-frame wavelengths across the explored redshift. The *ACS* filters chosen were F606W (between the Johnson V and R bands, hereafter "V") and F850LP (corresponding approximately to, and hereafter referred to as, the z-band). For some redshift ranges these observations provide immediately galaxy images in the rest-frame B-band ($\sim$ 4500 Å). For most redshifts one can reconstruct such a rest-frame image through pixel-by-pixel interpolation across the two bands, or through modest extrapolation in the other redshift ranges. The lowest redshifts of interest, z$\sim$0.2, require a blueward extrapolation of the observed V-band flux by 10% in λ and the highest redshifts, z=1.2, a 10% redward extrapolation of the z-band flux. For redshifts in between one can interpolate between the two observed filters; at z$\sim$0.33 our V-band corresponds directly to rest-frame B, as does the z-band at z$\sim$1. This choice of filters also provides consistency with the GOODS data at the field center.

Observations in two filters are crucial not only for the reconstruction of the rest-frame B-band, but also for color information, especially radial gradients, within one galaxy. Given limited observing time, area and imaging depth were deemed more important than a third filter. In cycle 11, 125 orbits of *HST* time (G0-9500, PI: Rix) were awarded to carry out this program. All these data have no proprietary period and are freely accessible.

2.3 Data

The full details of tile lay-out (Figure 2.1), the observations, the data reduction, and the data quality assessment will be given in (Caldwell *et al.* 2007) here we provide a brief overview.

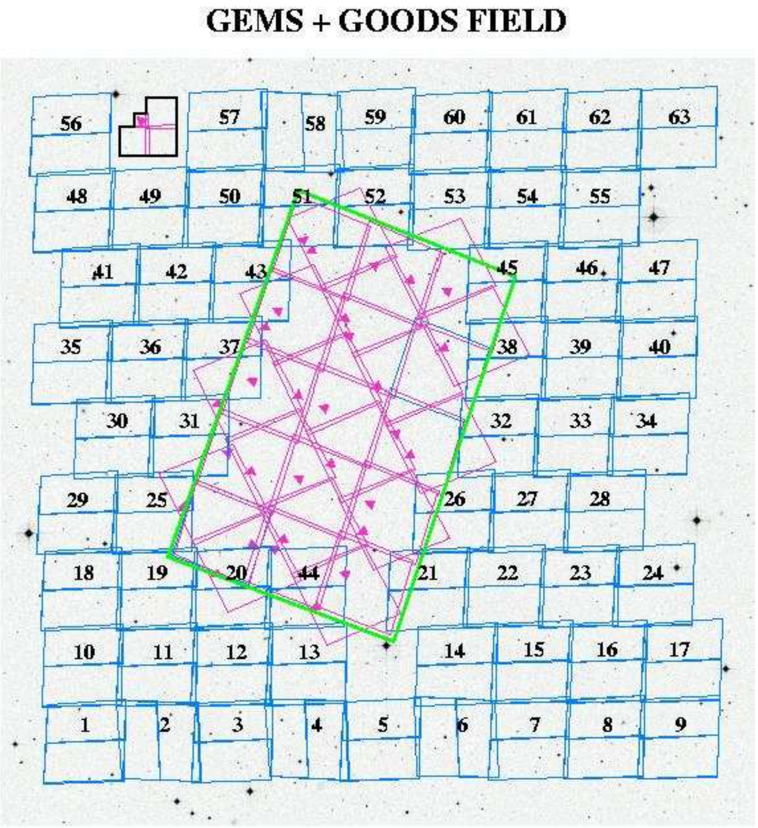


Figure 2.1 Layout of the GEMS image mosaic. With 800 square arcminutes, GEMS nearly covers the extended Chandra Deep Field South from COMBO-17 (underlying r-band image, see Wolf *et al.* 2003a), which measures $\sim 30' \times 30'$; the orientation is North up and East left. The individual GEMS tiles, labelled by their *HST* visit number are shown as pairs of rectangles (*ACS* chips). The mosaic tiles indicated in pink at the center and not aligned with the overall field are the first epoch observations of GOODS which have been incorporated into the overall GEMS analysis. The tilted large rectangle (solid green line) indicates the area of planned SIRTF observations for GOODSÅ few tiles have been omitted from the overall mosaic to avoid the brightest stars in the field. Observations for four tiles (2,4,6,58) had to be at different roll angles to assure guide stars. The area of the HDFs is indicated at the top left.

2.3.1 Observations

The bulk of the GEMS observations (59 visits, or 117 orbits) were carried out with the *ACS*'s WFC (Ford *et al.* 2003) between Nov 4 and Nov 24, 2002. Two visits were executed on Sep 14, 2002 and one each on Feb 24 and Feb 25, 2003. The first epoch observations of the GOODS survey that cover the central position of the GEMS field were taken in July and August 2002 (Giavalisco *et al.* 2003). The tile pattern of the overall mosaic (Figure 2.1) was laid out to a) encompass the GOODS epoch 1 data; b) create a large contiguous imaging field; and c) avoid excessively bright stars that would lead to excessive charge bleeding and scattered light on the CCD. Of the 63 tiles, 59 are oriented North-South. For four tiles, the availability of guide stars forced an orientation at right angles to the remaining ones (see Figure 1).

Each *HST* orbit visit (see labels in Fig. 2.1) consisted of three separate 12 min to 13 min exposures each for both V-band (F606W) and z-band (F850LP), dithered by 3" between exposures. The exposures of each tile in each filter required one full orbit with overhead. The dithering was chosen both to close the inter-chip gap and to provide sub-pixel sampling for drizzling of the final image. In each visit the first orbit was spent on V and the second on the z-band, where the rapid re-acquisition allowed a slightly longer (by 3 min) total exposure time.

2.3.2 Data Reduction

For the first version of the GEMS data the underlying approach was to reduce each tile in each filter separately, i.e. each set of 3 dithered exposures taken within an orbit was first treated as a completely independent data set. To assure data homogeneity, we re-reduced the first epoch GOODS data at the center of the overall GEMS area in the exactly same way as the GEMS data.

Each frame was processed using *CALACS* (www.stsci.edu/hst/acs/analysis) to take care of bias and dark current subtraction, flatfielding and to include the photometric calibration information. Frame combination and cosmic-ray rejection were accomplished with multidrizzle (Mutchler, Koekemoer, & Hack 2003), resulting in a combined image and a variance array on a 0."03/pixel grid (as opposed to the original 0."05/pix of the individual frames). We opted for a relatively fine 0."03 scale, to avoid resolution degradation in subsequent operations, even though it implies more strongly correlated pixel noise (Caldwell *et al.* 2007).

Cosmic ray rejection with three dithered frames worked excellently. As GEMS does not address time variable phenomena, any faint and rare residual cosmic rays are not of concern for its immediate science goals.

The astrometry of each image tile was tied to the overall catalog from the ground-based COMBO-17 *r*-band image (Wolf *et al.* 2001, W03), with an *rms* of $\sim 0.''14$ per source (see also §4.1). Both filters of each GEMS tile are tied to the COMBO-17 frame independently, but the V-band frames were subsequently *micro-registered* to a fraction of a pixel with respect to the z-band frame, for the color distribution analysis of individual sources.

Flux calibration was done using the best available zero points as of Feb. 2003, $V_{AB} = 26.49$ and $z_{AB} = 24.84$. The resulting point source sensitivities (5 σ) are: m_{lim} (V) = 28.25 and m_{lim} (z)= 27.10, in AB magnitudes. The angular resolution of the images, λ/D$\sim$ $0.055''$ and $0.077''$ in V-band (F606W) and z-band (F850LP), respectively, corresponds to physical resolutions of 500 pc and 700 pc for galaxies at $z \sim 0.75$, comparable to galaxies in the Coma Cluster imaged with $1''$ seeing. Figures 2.2 and 2.3 give a visual impression of how the GEMS images compare to the two most immediately related

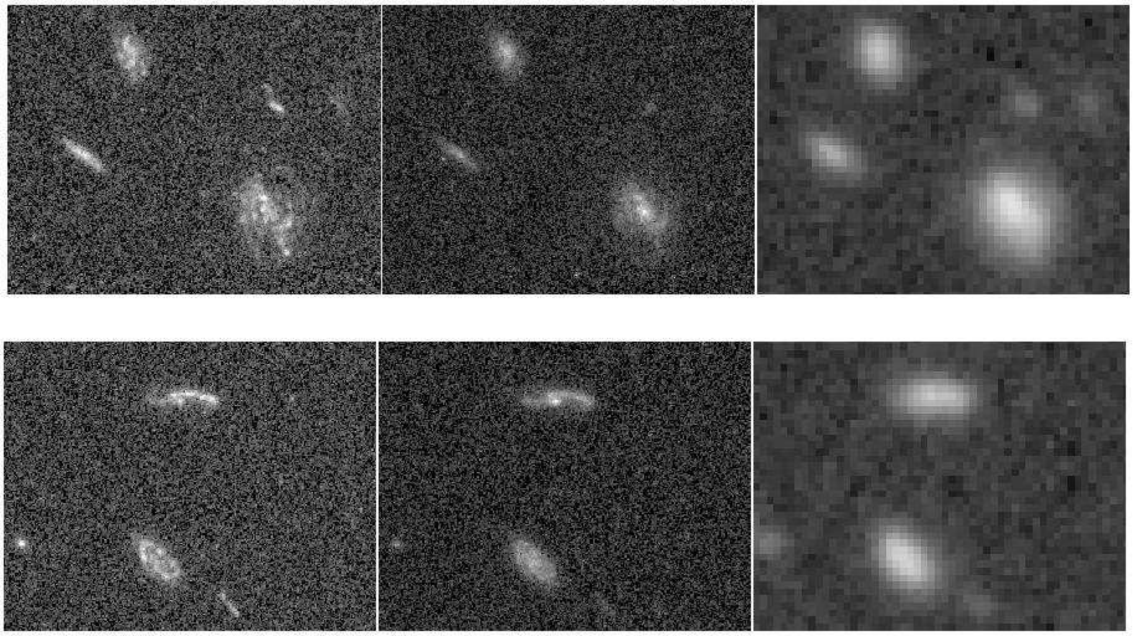


Figure 2.2 Comparison of the GEMS V-band (F606W) data (left panel in each row), and z-band (F850LP) data (center) with the deep, good seeing ($0.7''$ resolution) COMBO-17 R-band data (right half of each sub-panel). While comparable point-source and surface brightness sensitivity can be reached from the ground, the advantage in source parsing and in the assessment of morphological and structural information is manifest. Each panel is $14'' \times 10''$ on a side.

data sets: the deep *r*-band image from COMBO-17 and the deep, 5-orbit GOODS images. The total affective area of the GEMS mosaic is 796 arcmin^2.

We have not found any significant tile-to-tile variations in the relevant data properties (noise, sensitivity, *etc.*) and it appears that the intra-tile variations in sensitivity are also negligible. Further details will be given in (Caldwell *et al.* 2007).

2.4 Data Analysis

As the largest, multi-color image taken with *HST* to date, GEMS can be applied to wide range of scientific problems. Yet, the immediate focus of GEMS is to study the internal structure of galaxies for which redshifts and SEDs exist from COMBO-17. To accommodate the narrow and broader goals, the data analysis is broken down into three steps: 1) a catalog of "all" objects well detected in the GEMS z-band (F850LP) image, 2) a match-up with the COMBO-17 catalog, 3) the fitting of parameterized image models to selected source postage stamps. At a later stage this will be followed by the creation and analysis of color images. As for many applications of immediate interest, the longest ac-

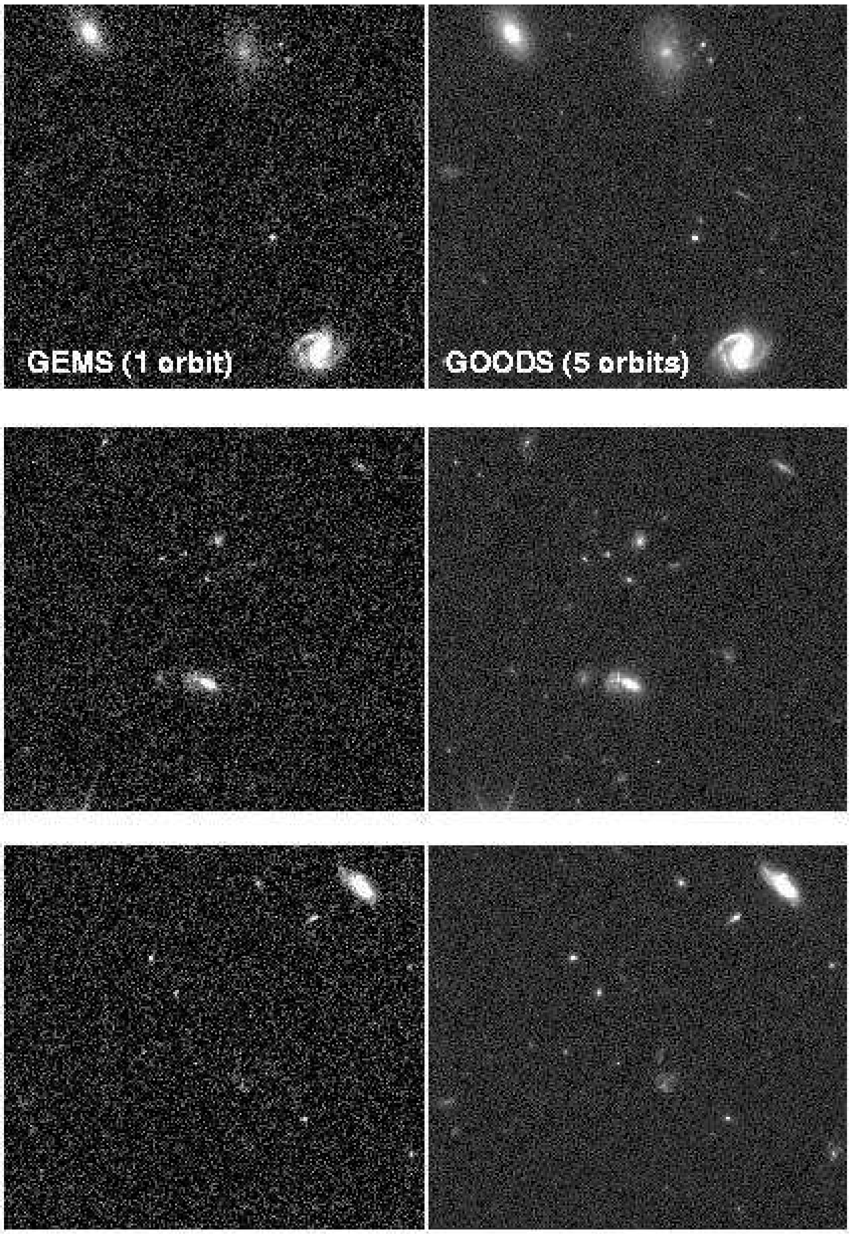


Figure 2.3 Comparison of a single-orbit exposure in z-band, as used throughout GEMS and of a 5-orbit exposure in the same band, reflecting the full exposure time of the GOODS deep imaging. Each panel is $33.6''x31''$ on a side, or 0.00036 times the total GEMS mosaic area.

cessible rest-frame wavelength is most relevant, thus the first version of the GEMS catalog is "driven" by the z-band image, with the V-band image providing color information.

2.4.1 Object Detection and Deblending

For parsing an image into an object catalog, the most widely used image software at present is SEXTRACTOR (Bertin & Arnouts 1996). As for the COMBO-17 catalogs, we apply SEXTRACTOR to the GEMS mosaic to obtain positions and a variety of photometric parameters for each detected source. We configure SEXTRACTOR to produce a GEMS source catalog that a) contains nearly all objects from the statistical COMBO-17 sample ($m_r \leq 23.6$); b) avoids spurious deblending of the largest galaxies, which show ample substructure in the *HST* images, reflecting spiral arms, OB associations, *etc.* ; c) provides a homogeneous, flux- and surface brightness-limited catalog of all sources in the z-band GEMS mosaic, regardless of COMBO-17 or other external information.

Even the first two requirements cannot be achieved with a single SEXTRACTOR parameter setting. The point-source flux limit of the GEMS z-band image is more than two magnitudes fainter than the COMBO-17 catalog limit (for typical SEDs). But the ground-based data, drawing on long exposures and large pixels, have at least as high a surface brightness sensitivity as the *ACS* data (see Fig. 2.2). To pick up all diffuse, low-surface brightness objects from COMBO-17, the SEXTRACTOR program requires an object detection threshold that is so sensitive that inevitably weak features in the outer parts of large galaxies get deblended spuriously, as illustrated in Figure 2.4. The judgment of "over-deblending" was made by visual inspection, independently by two of the co-authors (D.H.McIntosh & M. Barden).

To meet our catalog requirements, we then employ two different source detection configurations with SEXTRACTOR: a) a conservative, *cold* setting that avoids spurious deblending of large objects with strong substructure, but does not pick up all faint, low-surface-brightness objects in the COMBO-17 catalog; and b) a *hot* version, assured to detect all faint objects at the expense of an occasional over-deblending of a large source (see Fig. 2.2). The following SEXTRACTOR configuration parameters define cold and hot source detections: (i) the detection threshold above background DETECT_THRESH = $2.3\sigma_{\rm bkg}$ (cold), $1.65\sigma_{\rm bkg}$ (hot); (ii) the minimum number of connected pixels above threshold DEBLEND_MINAREA = 100 (cold), 45 (hot); (iii) the minimum contrast between flux peaks for deblending multiple sources DEBLEND_MINCONT = 0.065 (cold), 0.06 (hot); and (iv) the number of sub-thresholds considered during deblending DEBLEND_NTHRESH = 64 (cold), 32 (hot). For both cases we employ a weight map ($\propto$ var^{-1}) and a 3 pixel (FWHM) tophat filtering kernel. The use of weight maps reduces the number of spurious detections in low signal-to-noise (S/N) areas of each image (e.g. near image edges).

We apply SEXTRACTOR to our z-band mosaic only; galaxies appear to have more substructure in the bluer V-band imaging, which increases the number of spurious over-deblendings and hampers meeting our detection criteria. Our final catalog consists of all (18,528) objects detected with our cold configuration, augmented by an additional

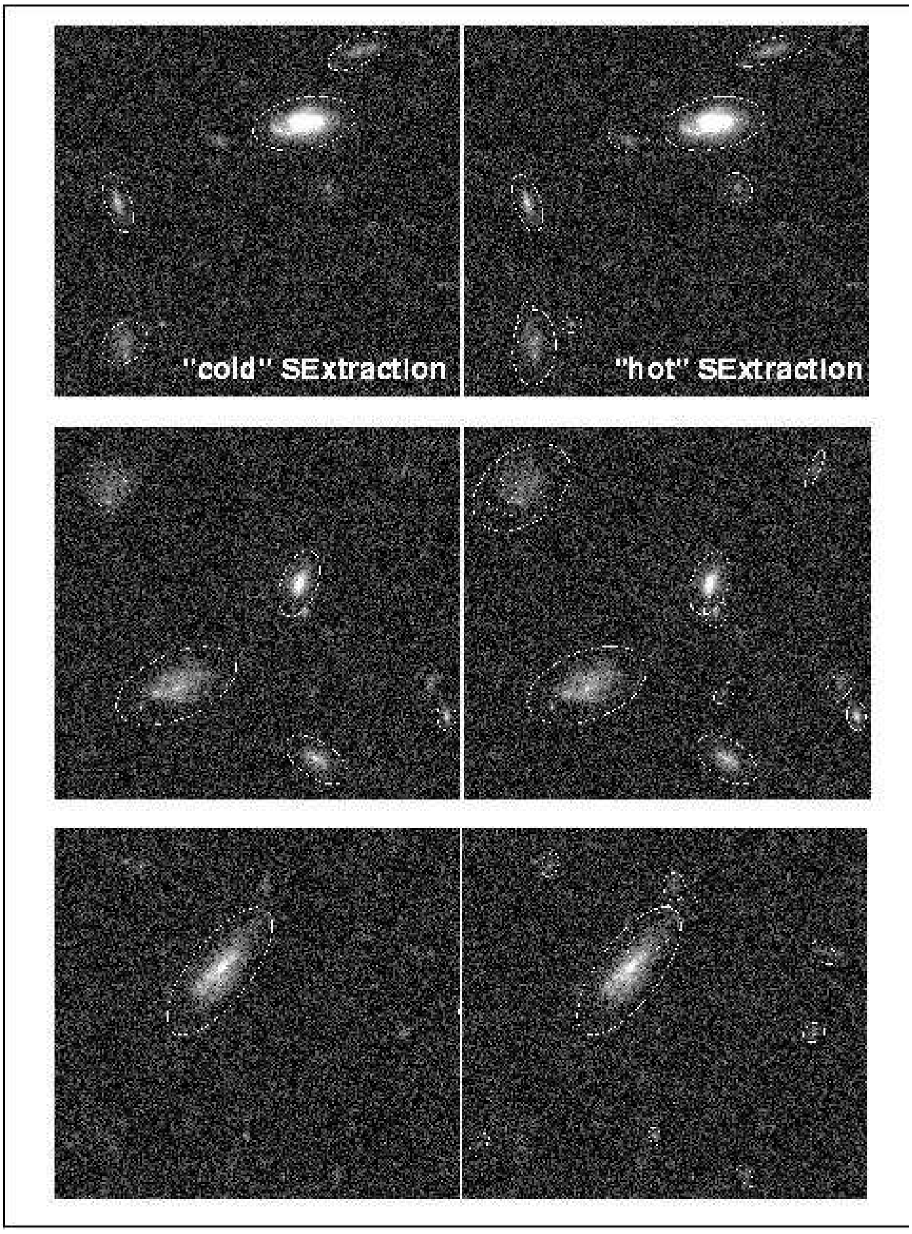


Figure 2.4 Two-pass strategy for object detection and deblending: the left-subpanels show the source identification with a conservative ("cold") setting of the SExtractor parameters that avoid over-deblending of large objects and galaxies with lumpy structure. The right panels shows the result of SExtractor with a parameter setting that picks up objects closer to the noise threshold, at the expense of occasionally breaking up objects erroneously. Our final object catalog consists of the left, cold objects augmented by the missing right, "hot" objects, but only those that do not overlap with a cold object. Each panel is $16.8'' \times 15.4''$ on a side.

(23,153) hot detections, but only those found *outside* of the isophotal area of sources from the "cold" catalog. Note, that while the SEXTRACTOR parameters for the hot and cold configurations were fine tuned interactively, the final GEMS catalog (of 41,681 unique sources) is produced strictly algorithmically. This GEMS catalog will be published and described in complete detail in (Caldwell *et al.* 2007).

We cross-correlate the final GEMS source catalog with the ground-based COMBO-17 redshift catalog solely on the basis of the object coordinates. There are 9,833 objects with redshifts from COMBO-17 (m$_R \lesssim 24$) in the E-CDFS field of COMBO-17$\dot{\text{W}}$e consider as a match the nearest redshift coordinate of a GEMS z-band position if it is within $0.75''$. The average RMS angular separation between matches is $0.13''$ and the fraction of unclear or blended detections at COMBO-17 coordinates is $\sim 1\%$. We find 8,312 unique GEMS sources with redshift matches resulting in a 84,5% success rate; the roughly 14.5% COMBO-17 objects without GEMS detections are due to the larger coverage of the E-CDFS region by COMBO-17 compared to our *ACS* imaging (Figure 2.1).

2.4.2 Image Simulations

We explore the detection limits of the GEMS mosaic, with the above described SEXTRACTOR cold+hot object detection configurations, by extensive Monte-Carlo simulations (Häußler *et al.* , 2007, see chapter 3). Simulated galaxy images were added to the actual data frames and processed as above. The detectability — and the subsequent ability to extract structural parameters — depends mostly on the effective surface brightness of the object (Figure 2.5), with much weaker dependences on the overall size and the axis ratio (see § 3). When defining the mean surface brightness of a galaxy image as:

$$\langle\mu_z\rangle \equiv \frac{I_{tot}}{2} \times \frac{1}{\pi r_{eff}^2 \times q},$$

where I_{tot} is the total flux, r_{eff} is the effective, or half-light radius and q is the axis ratio, we find the characteristic (80%) completeness limit of the GEMS *galaxy* sample to be $\langle\mu_z\rangle = 24$ for exponential profiles and $\langle\mu_z\rangle = 25$ mag arcsec^{-2} for R$^{1/4}$ profiles.

2.4.3 Point Spread Function

While the majority of galaxies in the combined COMBO-17/GEMS sample are resolved in the sense that the intrinsic half-light radius, R$_e$, is larger than the diffraction limit of *HST* ($\lambda/D \sim 0.077''$ at z-band), virtually all objects have central flux gradients on angular scales much smaller than the point spread function's (PSF) FWHM. This is particularly true in the cases of AGNs, where the unresolved central source often dominates. Any image modeling of the galaxies in GEMS requires therefore an accurate knowledge of the PSF.

There are two basic ways to construct a model PSF for the subsequent image interpretation: a calculation of the theoretically predicted PSF, using e. g. TinyTim (Krist 1993)

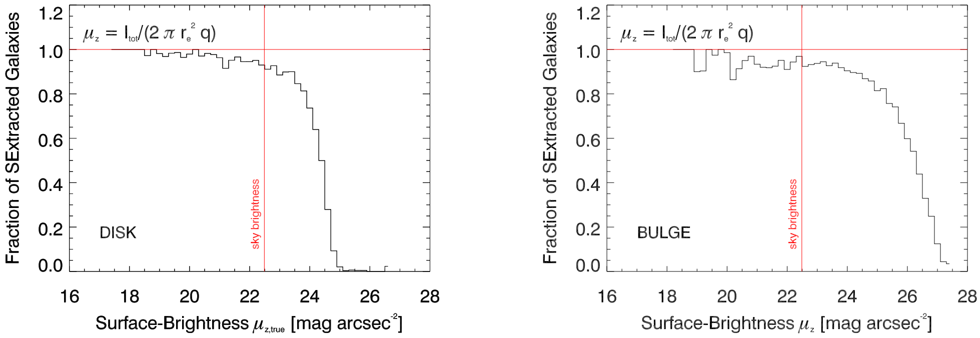


Figure 2.5 Completeness in surface brightness of the GEMS z-band imaging for detecting exponential disks (left plot) and de Vaucouleurs bulges (right plot) with SEXTRACTOR and subsequently fitting them. The definition of the mean surface brightness is given in §2.4.2. Even for bright galaxies, object overlap causes a small fraction of them not to pass detection and fitting. The vertical line indicates the background flux of the *ACS* data. The higher central concentration of de Vaucouleurs bulges permits the detection of $r^{1/4}$ objects with lower effective surface brightness than exponential disks.

or an empirical construction from the point sources within the data (e. g. *daophot* Stetson 1987). By necessity, the synthetic PSF is based on model assumptions, some of which are not sufficiently understood. On the other hand, the empirical PSF depends on a finite number of bright but unsaturated stars in the data. For strong PSF variations within the field, they may be hard to derive from fields at high Galactic latitude.

The large number of *ACS* images obtained within GEMS homogeneously acquired and reduced, has enabled us to study the inter- and intra-tile variations of the PSF in some detail. We found that variations among different tiles are negligible, while the PSF dependence on position within an *ACS* frame is noticeable but still small. We have performed extensive simulations assuring us that for fitting galaxy images without prominent AGN components, one universal, empirical high S/N PSF per filter is fully sufficient for all tiles and for all positions within each tile (see § 3).

Active Galaxies with a strong nuclear point source require a more elaborate treatment, due to the spatial PSF variations within the tiles. For such cases we use appropriate sub-tile PSF representations, jointly derived from all pointings. For each AGN the sub-tile PSF is constructed from at least ~ 30 stellar images near its pixel position (Jahnke *et al.* 2004b).

2.5 Summary

The GEMS project, which stands for “Galaxy Evolution from Morphologies and SEDs”, has produced the largest color image taken with *HST* to date, providing structural and morphological information for over 10,000 distant galaxies. We have presented here an overview of the science goals, the experiment layout, the observations, and the main data reduction steps.

Chapter 3
GEMS: Galaxy fitting catalogues and testing parametric galaxy fitting codes: GALFIT, GIM2D[4]

Abstract

In the context of recent *HST/ACS* surveys of intermediate redshift galaxies, we explore how their structure and morphology can be best described by parameterized models. We present a study of the inherent limitations of this technique at S/N of typical *HST* survey data (GEMS), we compare two widely used fitting codes, GALFIT and GIM2D, and we present the catalog of fit results for all 41,495 objects detected in the GEMS survey (Rix *et al.* 2004). We show a detailed comparison of GALFIT and GIM2D that (in the setup that we used here) describe galaxy images by fitting single-component Sérsic models to the two-dimensional light profiles. We test the reliability of both codes using a suite of simulated images, constructed to represent the characteristics of GEMS. For the sake of testing, the input galaxy parameters span a wider range than observed in GEMS. Our results are therefore a reasonable reflection of the uncertainties of the fitting analysis used in recent GEMS publications. We find that fitting accuracy depends sensitively on the input galaxy structure. Exponential profiles with Sérsic indices n of 1 — characteristic of galaxy disks — are on the whole fit well, whereas de Vaucouleurs profiles with $n = 4$ — characteristic of galaxy spheroids — are harder to fit owing to the large amount of galaxy light at large radii in such profiles. We find that both codes provide reasonably reliable fits whenever the average surface brightness $< \mu >_{eff}$ within the half-light radius is above the background surface brightness: for the bulk of $n = 4$ galaxies with $< \mu >_{eff}$ $(F850LP) <$ 22.5 magnitudes arcsec^{-2} a random scatter [in half-light radius] of $< 25\%$ for GALFIT and $< 50\%$ for GIM2D are readily achievable with no systematic offsets. Structural parameters for $n = 1$ galaxies of similar surface brightness are recoverable with much greater accuracy by both codes. We furthermore find that the errorbars returned by both codes are significantly smaller than the scatter of parameters in our tests and do not reflect the true

---

[4] This chapter has been submitted to ApJ.

'uncertainty' of the fit. We use the results of these tests to derive parameter uncertainties for individual objects that are published in this chapter. We highlight the importance of a correct sky estimate as an input to the parametric fitting analysis. We find that GALFIT results are robust to the presence of neighboring images, as GALFIT can simultaneously fit nearby neighbors while fitting the galaxy of interest. In contrast, GIM2D suffers significant systematic errors for $n = 4$ galaxies with bright nearby neighbors owing to the difficulty of fully masking out neighboring galaxies; there seems to be no work-around to this source of systematic error in GIM2D's current implementation. While this crowding error affects only a small fraction of galaxies in typical field *HST* surveys, GIM2D analysis of deeper cosmological images or of more crowded galaxy fields must account for this source of systematic error. We further explore the impact of image depth on fitting results, using a comparison between fit results from real galaxies in GEMS and the deeper GOODS data. We find that, given that both codes return consistent results for both deep and shallow images, real galaxies seem to be well described by a general Sérsic light profile. On the balance, GALFIT offers a number of important advantages over GIM2D for galaxy fitting on large moderate depth *HST*/*ACS* datasets, foremost its much higher speed and its robustness to nearby neighbors.

3.1 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 is large look-back surveys, where 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.* 2006) 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 billion years of cosmic time, since $z = 3$. In this chapter, we exhaustively test and tune two parametric galaxy fitting codes, GALFIT (Peng *et al.* 2002) and GIM2D (Simard *et al.* 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 presents parametric fits for 41,495 objects in the *HST* GEMS (Rix *et al.* 2004) dataset.

There are two main approaches towards describing galaxy structure from the two-dimensional information contained in image data. Non-parametric 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.* 2006b). The main disadvantages of non-parametric methods are 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.* 2003c). 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 non-parametric fitting codes, but have the advantage that light at large radii can be accounted for 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 Sersic (1968) model, which describes the radial surface brightness profile of a galaxy by the Sérsic function given by

$$\Sigma(r) = \Sigma_e \cdot exp\left[-\kappa((r/r_e)^{1/n} - 1)\right], \tag{3.1}$$

where r_e is the radius of the galaxy (Note that for a Sérsic fit $r_{\rm e}$ 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 7 free parameters. The Sérsic profile represents a more general form of the exponential light-profiles seen in galactic disks ($n = 1$), the $R^{1/4}$-law (de Vaucouleurs law) profiles typical of luminous early-type galaxy ($n = 4$) (e.g., de Vaucouleurs 1948; Freeman 1970); fitting with this profile has been explored in detail in a number of works (e.g., Simard 1998; Simard *et al.* 2002; Graham *et al.* 2005; Trujillo *et al.* 2001). Figure 3.1 shows some examples of Sérsic profiles with different n. Many authors have used a constant value of $n = 2.5$ or $n = 2.0$ 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.* 2004b; 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 chapter 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[5]. We compare the performance of the GALFIT and GIM2D automated galaxy fitting codes, which are suitable for fitting large datasets such as GEMS, STAGES (Gray & STAGES collaboration 2006, in prep.), 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 chapter in §3.2. In §3.3 we explore different set-ups, converging on 'best-fitting'

---

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

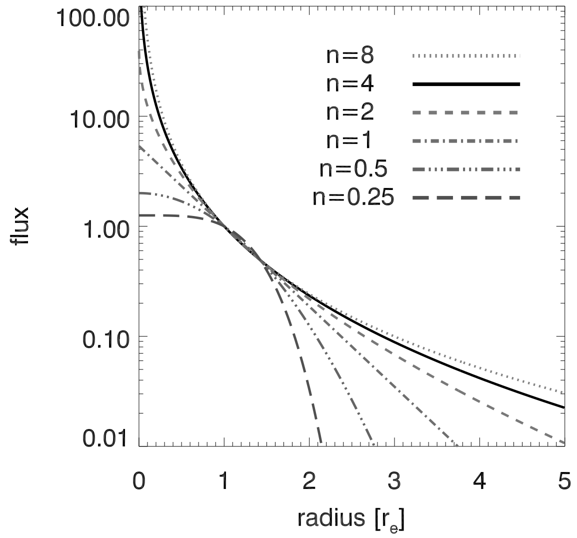


Figure 3.1 This plot shows Sérsic profiles for different values of the Sérsic index n, normalized to have the same flux at r_e. One can see that profiles with high Sérsic indices $n \gtrsim 2$ have more flux at larger radii; thus, a good estimate of the background sky level is particularly important for precise fitting of high-n galaxies.

set-ups for each fitting code. §3.4 summarizes the results from our testing of these 'best-fitting' set-ups using both simulated galaxies (§3.4.1 and §3.4.2) and real galaxies (§3.4.3). 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 §3.5, and publish a catalog of GALFIT fitting results for all 41,495 detected galaxies from the GEMS survey in §3.6. All results from this chapter and a number of other catalogs and images useful for testing galaxy fitting codes are presented on the GEMS webpage[6].

3.2 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 work; the results obtained from fitting these simulations are discussed in §3.4.

In this chapter, 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 thesis), and Sérsic profiles with a Sérsic index of $n = 4$ representing a de Vaucouleurs luminosity profile of a galactic bulge/elliptical

---

[6] see http://www.mpia.de/GEMS/gems.htm

galaxy (we will call them 'spheroids'), respectively. 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 fitting results for this sample can also be downloaded from the GEMS webpage.

This section is arranged as follows. The simulation of individual galaxies is described in §3.2.1. Section 3.2.2 describes the construction of simulated GEMS frames from the individual galaxy simulations, including the addition of realistic noise.

3.2.1 Simulation of Individual Noise-free Galaxies & 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 (3.1) to compute 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 (100x100 pixels up to 200x200 pixels depending on object size) have been oversampled by a factor of 10, and the very inner parts (10x10 up to 20x20 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.

3.2.2 Simulation of Crowded Images

To realistically test galaxy extraction and fitting codes requires creating 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.

Table 3.1 Simulation parameters for the sample of $n = 1$ galaxies

Simulation Parameter	Min	Max	Distribution
mag [mag]	20	26.5	uniform
r_e [pixel]	2	316	uniform in logarithmic space $r_e < 10^{7.36-0.233\cdot mag}$, mag being chosen magnitude for object
b/a	0.18	1	uniform in $\cos(i)$, i being inclination angle corrected for intrinsic thickness: $b/a = \sqrt{\cos^2(i) + (\sin(i)\cdot 0.18)^2}$ 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

Table 3.2 Simulation parameters for the sample of $n = 4$ galaxies

Simulation Parameter	Min	Max	Distribution
mag [mag]	20	27	uniform
r_e [pixel]	2	630	uniform in logarithmic space $r_e < 10^{4.79-0.1\cdot mag}$ $r_e < 10^{11.49-0.392\cdot mag}$
b/a	0.45	1	uniform in $\cos(i)$, i being inclination angle corrected for intrinsic thickness: $b/a = \sqrt{\cos^2(i) + (\sin(i)\cdot 0.45)^2}$
PA [deg to image]	0	180	uniform
Sérsic index n	4	4	fixed

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, completeness, and to allow pushing both parametric fitting codes to their limits. The simulations have a random distribution in size between 2 and 316 (uniformly distributed in logarithmic space) pixels 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 §3.4.2. The exact distributions of simulation parameters are given in Tables 3.1 and 3.2.

After simulating the galaxy profiles and putting them in an empty image, this final image was convolved with a real F850LP-band PSF derived from the GEMS dataset (Jahnke *et al.* 2004b). 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, 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 images depth and to make sure that objects appear in neither of the two bands) and visually checking those images to identify patches of 500x500 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 disk-like $n = 1$ galaxies only (for results see §3.4.1.1) and the other contains spheroidal $n = 4$ profiles only (see §3.4.1.2). Not all 800 galaxies were recovered by SEXTRACTOR. Roughly 80% of the objects where recovered, depending on the simulated profile shape and the distribution of galaxy parameters in the particular image (see Figure 3.2). In particular, very large and low surface-brightness galaxies were not detected (see Figure 3.2 and Rix *et al.* 2004). Due to the fact that spheroids are easier to detect due to their centrally concentrated light profiles, less galaxies were recovered in the disk sample.

By design, these simulated tiles are artificial in two ways. Firstly, the galaxy input parameters span a wider range in parameter space than real galaxies. Secondly, 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 *ACS* image from GEMS (see Figure 3.3). They contain many more LSB 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 §3.4.2) to estimate the effects of deblending given 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, non-smoothness, lumpiness and/or spiral features of real galaxies.

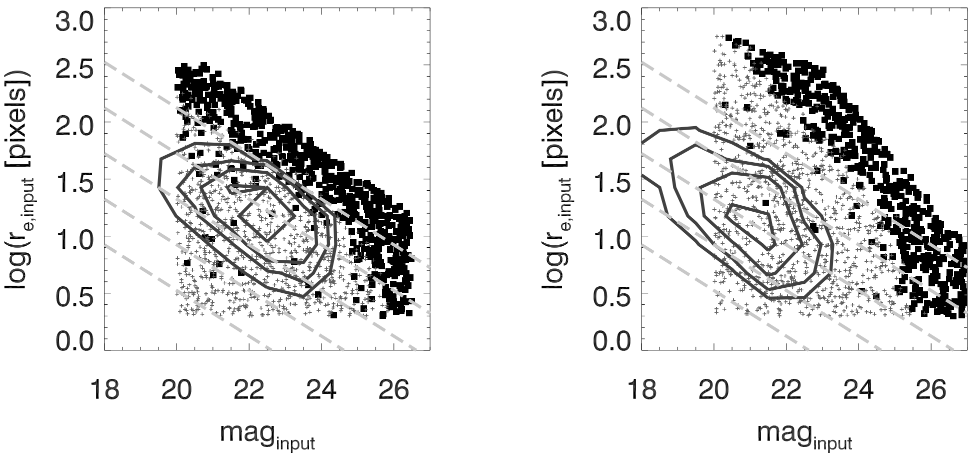


Figure 3.2 F850LP magnitudes and sizes for the full set (all symbols) of 1600 simulated n=1 galaxies (left) and n=4 galaxies (right), and the subsets that were detected by SExtractor (grey crosses). The black squares indicate galaxies that were missed during object detection. The indicated contours show the magnitude-size space populated by actual GEMS galaxies used by Barden *et al.* (for disk galaxies 2005, left) and McIntosh *et al.* (for spheroidal galaxies 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 non-detected galaxies in both samples reflect 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 overplot long-dashed lines of constant surface brightness of 17, 19, 21, 23 and 25 magnitudes arcsec^{-2} from bottom to top.

3.3 Galaxy Fitting: Description, Basic Considerations, 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 the previous section. 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

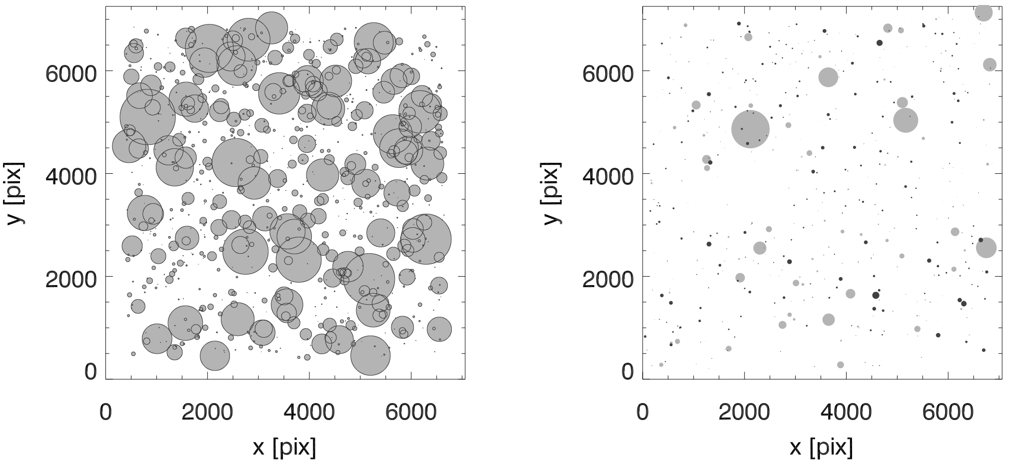


Figure 3.3 Density of sources in simulation images. *Left:* One of the two simulated images with galaxies having Sérsic index of 4. We plot a circle with radius r_e (simulated) at the correct x, y location of each simulated galaxy. In total there are 800 simulated $n = 4$ galaxies placed in a single *ACS* tile, ~ 250 of which are too LSB 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:* the 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 (in dark grey) and 148 with n=4 (light grey). Stars and the few objects (in total 46 objects) that ran into a fitting constraint where 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 §3.4.1.1 and §3.4.1.2

tests we carried out in order to optimize these setups in §3.3.1 for GALFIT and §3.3.2 for GIM2D.

The initial conditions and setups for both GALFIT and GIM2D are determined using SEXTRACTOR output images and catalogues. We use SEXTRACTOR (version 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 the following sections 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.

3.3.1 GALFIT

GALFIT is a 2D galaxy fitting software package written by Peng *et al.* (2002). We used GALFIT Version 2.0.3b from Feb. 2, 2005 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 4 threads simultaneously to efficiently use all 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 multi-object, fits this happens relatively rarely, if at all, both through the simulations (§3.4.1) and through comparison of fitting results for real galaxies from GALFIT and GIM2D (§3.4.3).

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 work we use this capability for testing purposes only (see §3.3.1.2).

In the following section, 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 §3.3.1.2.

3.3.1.1 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 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 Barden et al., in prep). For every object in the SEXTRACTOR catalog GALAPAGOS did the following.

1. First, GALAPAGOS determined the size of the required postage stamp for each ob-

ject. This was done using different object sizes and angles given by SEXTRACTOR:

$$Xsize = 2.5 * a * kron * (|\sin(\theta)| + (1 - ellip) * |\cos(\theta)|) \tag{3.2}$$

$$Ysize = 2.5 * a * kron * (|\cos(\theta)| + (1 - ellip) * |\sin(\theta)|) \tag{3.3}$$

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 to have enough sky pixels present in the postage stamp to give a robust fit of the object, while at the same time 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 where 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 ellipse as the mask[7]. 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%/46% of the fits, at least *one* secondary 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 9/7/12 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 6 consecutive elliptical annuli, each with a width of 60 pixels (measured along the semi-major axis; corresponding to 1.8" using the GEMS data with 0.03"/pixel). These 6 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 6 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 6 annuli. This made the area where the sky

---

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

is determined to be an ellipse between 35 and 215 pixels in semi-major axis for the smallest objects (between 15 and 30 r_e) and an ellipse of width of 180 pixels at around 4-6 r_e for the bigger objects (for details see Barden et al, 2007, in prep.). 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 §3.3.1.2).

4. In the same step, by dividing the elliptical individual annuli into 8 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 although a strong gradient was present (i.e. the object lay outside of the original GEMS tile), we fit 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%, 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 where less galaxies are recovered.

Table 3.3 Starting guesses for GALFIT when using GALAPAGOS

Parameter	Starting guess from SEXTRACTOR
mag	`MAG_BEST`
r_e	$0.162 \cdot$ `FLUX_RADIUS`$^{1.87}$
b/a	1 - `ELLIPTICITY`
PA	`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

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 3.3). 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 $r_e = 0.162 \cdot R_{flux}^{1.87}$, 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 by a factor of 1.5 enlarged SEXTRACTOR apertures that tells GALFIT which pixels it should use and which pixels it should ignore during the fit). Finally, the GEMS PSF (see Jahnke et al., in prep.) 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 $0.3 < n < 8$, $0.3 < r_e < 500$ [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.

3.3.1.2 GALFIT Sky Test

The estimate of the sky background is of critical importance in determining parametric or non-parametric 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 in §3.3.2.3). We test three setups: *i)* the isophotal sky, *ii)* the sky value determined by SEXTRACTOR, and *iii)* allowing sky to be a free parameter, to be estimated by GALFIT.

In Figure 3.4 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 was known to be 18.14 ± 0.03, indicated by the vertical dashed line in both plots. SEXTRACTOR recovers a mean value of 18.29 (σ=0.10) for disk galaxies and 18.40 (σ=0.11) for spheroidal galaxies. The isophotal estimator in GALAPAGOS gives a mean value of 18.13 (σ=0.10) for disk galaxies and 18.26 (σ=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

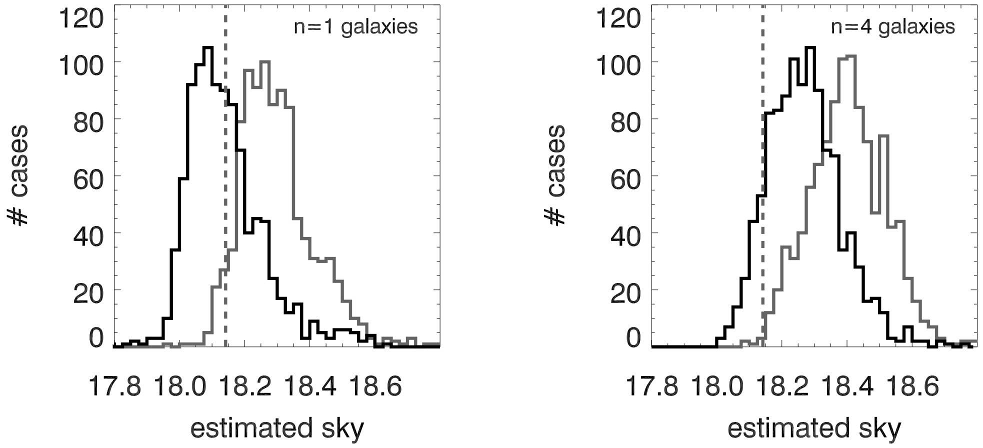


Figure 3.4 This plot shows the recovered sky values for both used estimators, SEXTRACTOR (grey) and isophotal (determined within GALAPAGOS, black) for each of the simulated samples. The vertical dashed line indicates the true value. One can clearly see that both methods tended to overestimate the sky value for the spheroid sample, mainly due to the large wings of galaxies in this sample contributing to the sky level.

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[8] 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 3.5. 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 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

---

[8] see $http : //www.stsci.edu/science/goods/catalogs/r1.0z\_readme/$, chapter 5.1 Local sky background

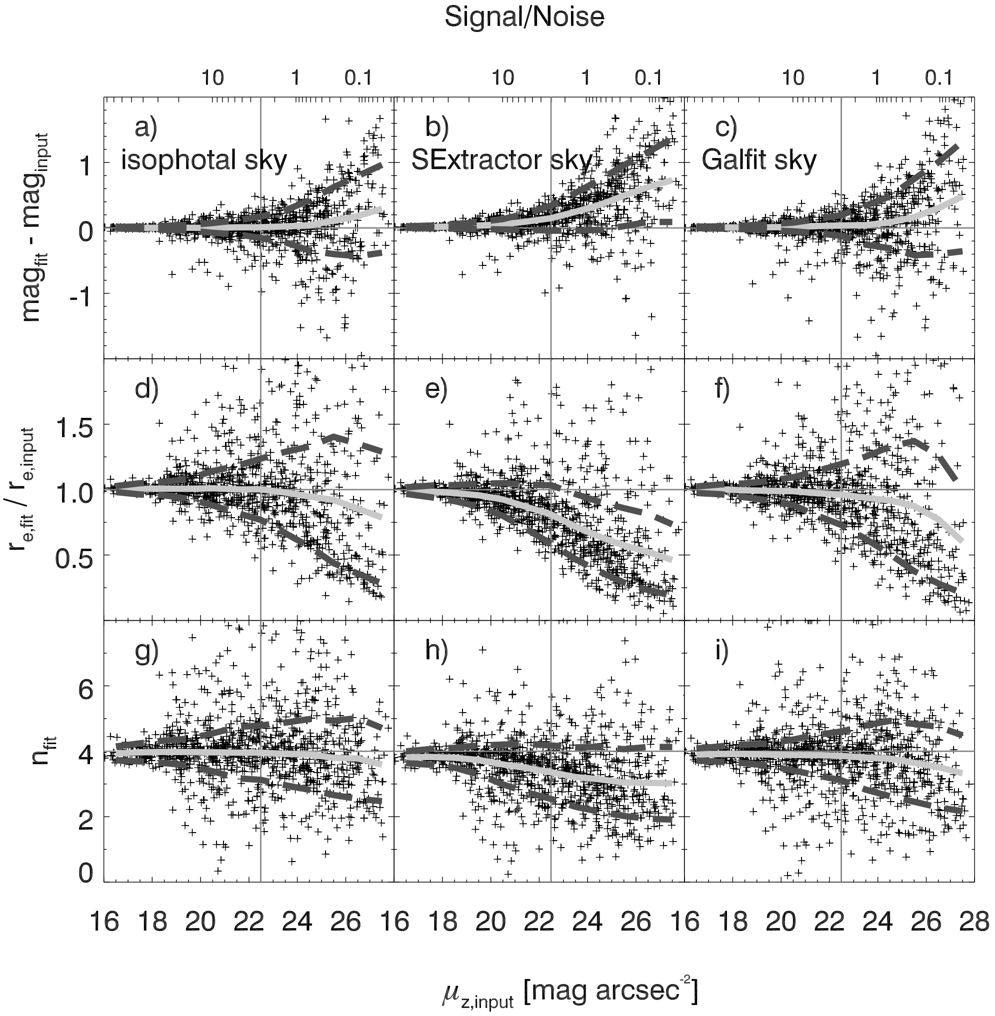


Figure 3.5 This plot shows the fitting results of the spheroid galaxy sample when fitted with GALFIT using three different sky estimates (see §3.3.1.2): isophotal sky from GALAPAGOS (left), SEXTRACTOR sky (middle) and the sky returned by GALFIT when allowed to fit it as a free parameter. The X-axis shows the simulated mean surface brightness within r_e defined by $\mu = mag + 2.5 \cdot \log(2 \cdot b/a \cdot \pi \cdot r_e^2)$, where mag is the magnitude, b/a the axis ratio and r_e the half-light radius of the object in arcsec. The thin vertical grey line in the plot indicates the brightness of the sky background. The upper X-axis shows the mean signal-to-noise ratio per pixel within r_e calculated by: $S/N = \langle\varrho\rangle \cdot [\langle\varrho\rangle + \langle\varrho_{sky}\rangle + \sigma_{sky}]^{-1/2}$, where $\langle\varrho\rangle$ is the average counts in a galaxy pixels within r_e (basically μ), $\langle\varrho_{sky}\rangle$ is the background flux within a pixel and σ_{sky} is the uncertainty of the background sky (...caption continues on next page...)

[To figure 3.5:] 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-grey line. The thick light-grey line and the thick dashed dark-grey line indicate the mean value and 1σ values for different surface brightness bins. The small crosses show the galaxies that where fitted 'successfully', meaning that the fit returned a result and that it did not run into fitting constraints.

shows the simulated mean surface brightness μ_{input} of the galaxies within an ellipse with semi major axis r_e and the axis ratio b/a:

$$\mu = mag + 2.5 \cdot \log(2 \cdot b/a \cdot \pi \cdot r_e^2) \tag{3.4}$$

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

$$S/N = \langle \varrho \rangle \cdot [\langle \varrho \rangle + \langle \varrho_{sky} \rangle + \sigma_{sky}]^{-1/2} \tag{3.5}$$

where $\langle \varrho \rangle$ is the average countrate [in e^-] for galaxy pixels within r_e, $\langle \varrho_{sky} \rangle$ is the background flux [in e^-] within a pixel, and σ_{sky} is the uncertainty of the background sky estimation, estimated using 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 (Figure 3.5). 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 brightness (approximately 2 magnitudes arcsec^{-2} above the sky level) for SEXTRACTOR-derived sky values. In all cases, the effects are primarily systematic in nature, and could in principle be corrected for, if required.

According to these test results, we choose to use the isophotal sky estimate for GALFIT analysis. Even though, for simulated galaxies, allowing GALFIT to determine the sky appears as reliable as our isophotal sky method, for real galaxies, complex structures that deviate from profile assumptions may affect sky estimation. We therefore choose to hold the sky values fixed to our best estimate for GALFIT fitting of both simulated and real galaxies.

3.3.2 GIM2D

GIM2D (*G*alaxy *Im*age *2D*) was written by Luc Simard (Simard 1998; Simard *et al.* 2002) as an IRAF package for the quantitative morphological analysis of galaxies. We use version 3.1 for the analysis in this thesis. For a single Sérsic fit we work in 7-dimensions with the bulge fraction parameter set to $B/T = 1$; thus, we find the best-fit model described by $f_{\rm tot}, r_{\rm e}, e, \phi_B, dx, dy$, 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 datasets, GIM2D ought to be run on many machines in parallel.

3.3.2.1 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 catalogues, a square postage stamp was cut from the large image, centered on each galaxy with size given by $4a_{\rm iso} \times 4a_{\rm iso}$, where $a_{\rm iso}$ 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 are presented later in this section, and in §3.4.1.2. 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 ones — 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 non-object (sky) pixels as specified by the SEXTRACTOR segmentation map. As such, this method is closely dependent upon 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 with non-idealized profiles. We test the effect of different sky estimates in detail in §3.3.2.3.

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 $N_{\rm ICF}$ models throughout the allowed parameter space, selects the best one, and then reduces the search volume by a factor equal to $N_{\rm ICF}$. The final result from the ICF is used as the starting point by the Metropolis algorithm. The GIM2D website gives a default value of $N_{\rm ICF} = 100$.

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 3.5 (for bright galaxies) and 3.6 (for faint galaxies) and will be discussed in detail in the following sections starting with the recommended GIM2D setup (§3.3.2.2), sky tests (§3.3.2.3), other tests (§3.3.2.4), concluding with the final adopted best-fitting setup (§3.3.2.5).

3.3.2.2 GIM2D Recommended Setup

In Figure 3.6, in the leftmost panels, we show fitting results for the setup that is recommended on the GIM2D webpage[9]. This recommended setup, in particular, has `'dobkg'`=yes and `'initparams'`=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 $\sim 50\%$ in r_e near the sky level. Fitting results are strongly systematically biased towards 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 `'initparams'` and `'dobkg'`.

Through extensive testing, we find that the best results are obtained when *both* `'initparams'`=no and `'dobkg'`=no (see rightmost panels in Figure 3.6), and when the SEXTRACTOR local background is used (see §3.3.2.3). Setting `'initparams'`=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[10]. We note that setting both dobkg and initparam to no is substantially better at all surface brightness, although the scatter at

---

[9] The GIMFIT2D description is http://www.hia-iha.nrc-cnrc.gc.ca/STAFF/lsd/gim2d/, and specifies the last program update of March 19, 2001

[10] 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. 3.6 and setup G in Table 4 shows that at these limits the GIM2D fitting results suffer from $\lesssim 10\%$ biases.

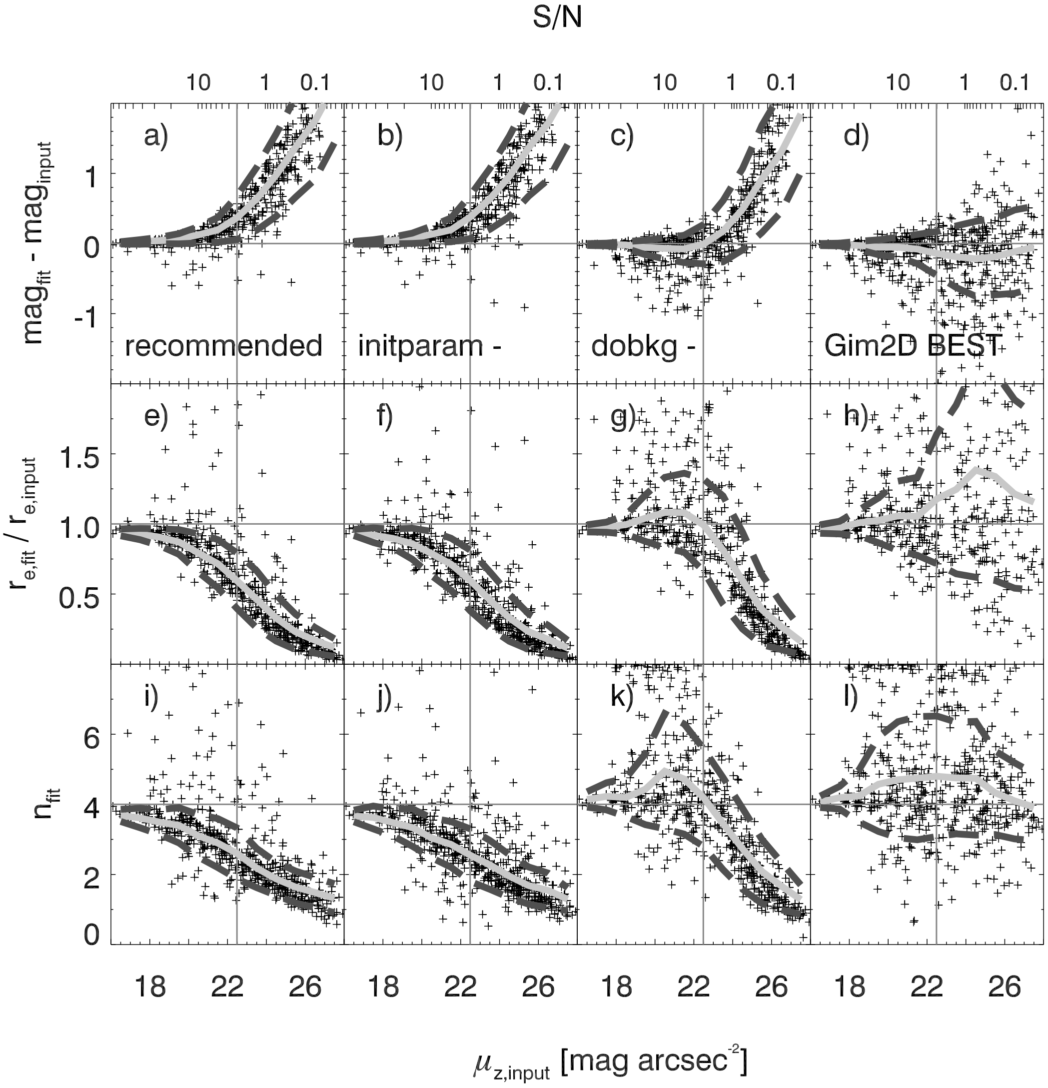


Figure 3.6 This plot compares the recommended GIM2D setup to setups where we used different settings of the GIM2D parameters ‘initparam’ and ‘dobkg’ [from left to right: setup I (initparams=+, dobkg=+, recommended),G (-, +),H (+, -)and E (-, -; best), see Table 3.6]. The X-axis shows the simulated surface brightness of the galaxies. The Y-axes again show magnitude difference, size ratio and Sérsic index fitting results. The thick light-grey line and the thick dashed dark-grey line indicate the mean value and the 1σ line for different surface brightness bins. One can see easily that our *best-fitting* GIM2D setup (see §3.3.2.5) fits galaxies with much less systematic bias than the initial setup recommended on the GIM2D webpage. This is especially true for galaxies fainter than the surface brightness of the sky.

fainter surface brightness is large. It is worth noting that compared to these two parameter choices, other effects such as the sky value used (see §3.3.1.2) and the image size (§3.3.2.4) appear to produce only low level improvements.

Table 3.4 Parameter limits used for GIM2D when using initparam=no

Parameter	min	max
mag	20	27
r_e	0.3	500
ellipticity	0.0	1.0
PA	all	all
n	0.2	8.0
centering x	0	3.0
centering y	0	3.0

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 3.4). Fits that ran into any of the given fitting constraints were removed from the sample for the analysis in this chapter.

From our findings, we *strongly discourage* GIM2D users to use `'initparam'`=yes and `'dobkg'`=yes.

3.3.2.3 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 §3.3.1.2 for GALFIT. We tested GIM2D using SEXTRACTOR local sky (setups D, E and H in Table 3.5) and the isophotal sky that we used for GALFIT (setups A, B, C, J and K). We carried out one test — setup F — where we used the 'real' background of 18.14 counts, determined on the 'sky' image. 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 G and I we tried fixing the sky background to the value determined directly by GIM2D (`'dobkg'`=yes). We show results from these different tests in Figure 3.7.

It is interesting that GIM2D performs somewhat better using SEXTRACTOR local sky, whereas we showed in §3.3.1.2 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

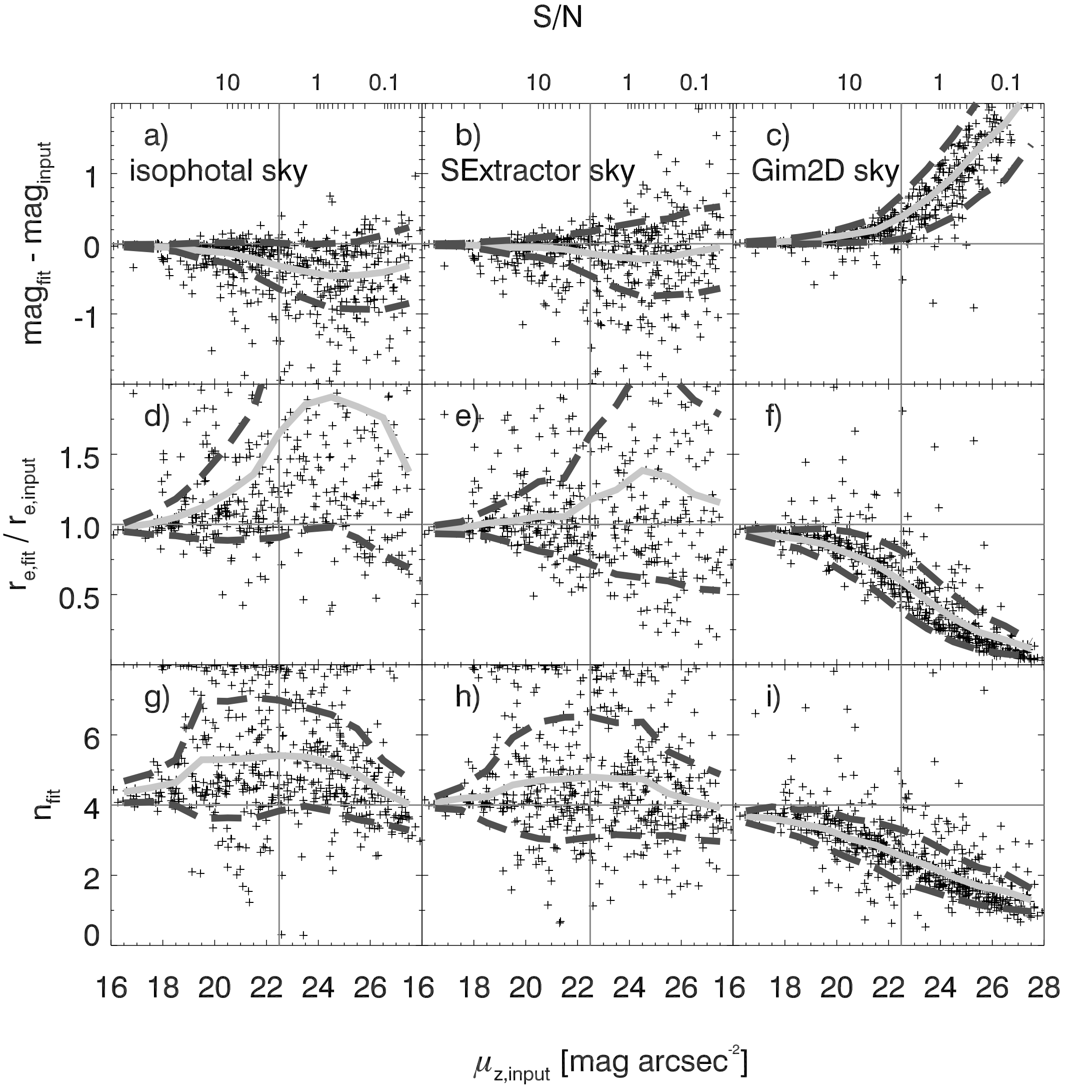


Figure 3.7 This plot shows the same as Figure 3.5 but for GIM2D results (setups B,E and G in Table 3.5). As is immediately obvious from these plots, GIM2D performs best when using the SEXTRACTOR background held fixed during the fit.

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.

3.3.2.4 GIM2D Other Tests

Table 3.5 GIM2D: Fitting of $n = 4$ Simulations: Bright subsample with $\mu_{\mathrm{in}} < 22.5$ and $\mathrm{mag}_{\mathrm{in}} < 22.5$

Setup	N/N_{tot}	Sérsic n		r_{50} ratio		$\Delta\mathrm{mag}$		e ratio		ΔPA		Quality
		mean	σ	mean	σ	mean	σ	mean	σ	mean	σ	
A	161/545	4.81	0.76	1.12	0.16	-0.09	0.09	1.01	0.06	-0.03	1.6	62.26
B	165/539	4.75	0.67	1.10	0.17	-0.08	0.08	1.01	0.06	0.05	2.0	53.60
C	161/533	5.57	1.49	1.14	0.22	-0.14	0.13	1.02	0.07	-0.20	2.0	193.85
D	168/533	4.37	0.43	1.02	0.09	-0.03	0.05	1.01	0.06	-0.16	1.5	9.85
E	164/540	4.34	0.61	1.02	0.13	-0.04	0.07	1.01	0.05	0.00	2.2	8.85
F	163/533	5.68	1.40	1.27	0.30	-0.16	0.14	1.01	0.06	-0.03	1.7	277.71
G	167/551	3.30	0.49	0.82	0.14	0.09	0.09	1.00	0.05	-0.18	1.9	72.15
H	165/549	4.32	0.56	1.01	0.10	-0.03	0.07	1.02	0.06	-0.10	1.8	7.71
I	168/546	3.33	0.51	0.83	0.13	0.09	0.09	1.01	0.05	-0.01	1.6	64.76
J	164/531	4.84	0.76	1.11	0.17	-0.09	0.08	1.01	0.06	-0.17	2.0	65.09
K	162/539	4.82	0.75	1.10	0.17	-0.08	0.07	1.01	0.06	0.34	2.2	58.59

Note. — This Table summarizes the results from all GIM2D testing for bright galaxies; see Table 3.6 for results for faint galaxies.

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 3σ) from the simulated value and scatter for the 5 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:

$$Quality = 1000 * [(\Delta n/4 - 1)^2 + (\Delta re - 1)^2 + (\Delta mag)^2 + (\Delta(b/a) - 1)^2 + (\Delta PA/180)^2] \quad (3.6)$$

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).

Explanation of the setups:

(A) isoph. bkg, *initparam=yes*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes
(B) isoph. bkg, *initparam=no*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes
(C) isoph. bkg, *initparam=no*, $N_{\mathrm{ICF}} = 100$, $6a_{\mathrm{iso}}$ image sizes
(D) SExtr. local bkg, *initparam=no*, $N_{\mathrm{ICF}} = 100$, $2a_{\mathrm{iso}}$ image sizes
(E) SExtr. local bkg, *initparam=no*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes, *best* setup
(F) bkg $= 18.14$, *initparam=no*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes
(G) *dobkg=yes*, *initparam=no*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes
(H) SExtr. local bkg, *initparam=yes*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes
(I) *dobkg=yes*, *initparam=yes*, $N_{\mathrm{ICF}} = 100$, $4a_{\mathrm{iso}}$ image sizes, *recommended* setup
(J) isoph. bkg, *initparam=no*, $N_{\mathrm{ICF}} = 25$, and $4a_{\mathrm{iso}}$ image sizes
(K) isoph. bkg, *initparam=no*, $N_{\mathrm{ICF}} = 400$, and $4a_{\mathrm{iso}}$ image sizes

To determine the best-fitting setup to use with GIM2D, we performed 11 different tests (rows A-K) as shown in Tables 3.5 (bright galaxies) and 3.6 (faint galaxies). For

Table 3.6 GIM2D: Fitting of $n = 4$ Simulations: Faint subsample with $22.5 < \mu_{\rm in} < 26.0$

Setup	$N/N_{\rm tot}$	Sérsic n		r_{50} ratio		Δmag		e ratio		ΔPA		Quality
		mean	σ	mean	σ	mean	σ	mean	σ	mean	σ	
A	153/545	2.45	0.92	0.53	0.27	0.53	0.50	0.96	0.16	-0.88	8.1	652.75
B	149/539	5.13	1.26	1.94	0.93	-0.47	0.46	1.03	0.20	-1.00	6.3	1178.66
C	147/533	5.52	1.36	3.22	1.93	-0.76	0.57	1.00	0.20	-1.22	10.6	5635.45
D	149/533	4.52	0.92	1.44	0.67	-0.23	0.43	1.05	0.21	-0.13	7.0	264.95
E	151/540	4.63	1.45	1.38	0.78	-0.20	0.55	1.05	0.23	-1.10	7.4	212.71
F	146/533	5.38	1.19	2.84	1.51	-0.73	0.53	0.99	0.25	-0.88	7.0	4026.70
G	155/551	1.79	0.44	0.28	0.11	1.08	0.49	1.02	0.19	-0.57	9.2	1992.40
H	154/549	2.46	0.86	0.50	0.23	0.57	0.48	1.01	0.19	-0.64	5.5	730.11
I	154/546	1.78	0.40	0.28	0.12	1.10	0.47	1.04	0.15	-1.86	7.5	2027.37
J	147/531	5.05	1.27	1.89	0.95	-0.47	0.49	1.02	0.16	-1.17	7.6	1083.47
K	148/539	5.24	1.30	2.36	1.45	-0.52	0.51	1.02	0.16	-1.43	6.5	2221.91

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 $N_{\rm tot}$ galaxies in the sample for which GIM2D returned a result. The ‘faint’ sample included galaxies with $23.5 < \mu_{sim} < 26$ magnitudes arcsec^{-2}. We 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 §3.4.1.3, we illustrate this behavior by running GIM2D on 3 different postage stamp widths of $2a_{\rm iso}$ (setup D), $4a_{\rm iso}$ (setup E), and $6a_{\rm iso}$ (setup C). 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 3.5 and 3.6). The best-fit stamp cutouts we adopt here have sides equal to $4a_{\rm iso}$. 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 $4a_{\rm iso} \times 4a_{\rm iso}$, due to the requirement of simultaneously fitting neighboring galaxies. We will further quantify the effect of neighboring galaxies in §3.4.1.2 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 $N_{\rm ICF} = 100$ (setup B in Table 5), to results for $N_{\rm ICF} = 25$ (setup J) and 400 (setup K). We found that the results are independent of the number of ICF models.

3.3.2.5 GIM2D Best-fitting Setup

As is apparent especially from Table 3.6 and Figure 3.6, the best combination of parameters for our simulations was given by setup D, using SEXTRACTOR background (*dobkg=no*), *initparam=no* and $4a_{\rm iso}$ as image sizes. We choose this setup to be our *best* and use it throughout this thesis to compare GIM2D results with GALFIT results.

3.4 GALFIT/GIM2D Comparison Using Optimized Setups

In this section, we discuss the results of testing our best setups of GIM2D and GALFIT. Section 3.4.1 describes the results obtained using the simulated images with artificial distributions of galaxy parameters as explained in §3.2. Section 3.4.2 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 3.4.3 sums up the results of tests where real images of different depths were fit and the results intercompared.

3.4.1 Results of Fitting Simulated Galaxy Images

3.4.1.1 Results of Pure Disk Simulations

Figure 3.8 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, 6 (0.6%) fits crashed. GIM2D fitted 870 (87%) successfully, 46 (5%) ran into fitting constraints, 81 (8%) of the fits crashed. There are 4 (0.4%) galaxies for which both codes failed.

Crosses in Figure 3.8 represent galaxies that were fitted by both codes. Grey squares show galaxies that were fitted by that code only; the other code failed to return a useful result either through running into one of the fitting constraints, or the fit crashed. The thick light-grey line and the thick dashed dark-grey line indicate the mean value and the 3σ 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 6 plots. The 3 rows show the results for magnitude (plots a and b), size (c and d) and Sérsic index (e and f), respectively. The thin vertical line indicates the brightness of the sky background within the GEMS survey. This is also roughly the limit up to which galaxies are used for science within GEMS. The Y-axes show deviations of the fitting values to the true parameter values, the horizontal

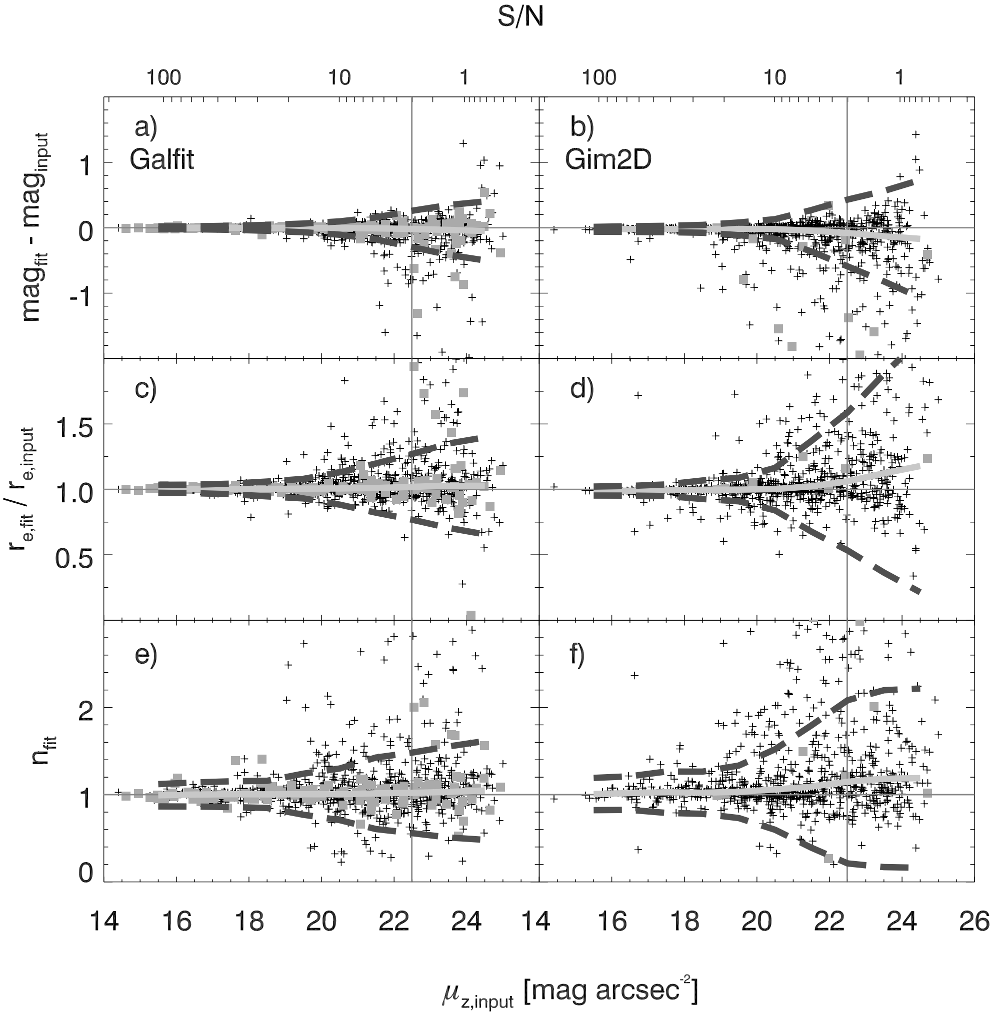


Figure 3.8 Fitting results for GALFIT (left) and GIM2D (right) for the set of simulated $n = 1$ galaxies. The X-axis again shows the input surface brightness, the thin vertical grey line indicates the brightness of the sky background. The Y-axes are the same as in Figure 3.5. The input value is indicated as the horizontal thin dark-grey line, in this case representing a Sérsic index of 1 for the sample of disk galaxies. The mean value of the deviations and the 3σ line are indicated for different surface brightness bins. The small crosses show the galaxies that where fitted ‘successfully’ by both codes. Grey squares indicate galaxies that where fitted by one code only and the other code did not return a meaningful result.

thin line indicating the ideal value, which, in 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 3 different surface brightness bins: first, the galaxies of highest surface brightness which one clearly would want to be fitted well with any code; secondly, galaxies within a surface brightness bin of 1 magnitude around the surface brightness of the sky; and thirdly, the faintest galaxies, much fainter than the sky surface brightness. The third sample 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 offset between the input and recovered values in Figure 3.8; 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 ($r_{fitted}/r_{simulated}$, $\sigma \sim 0.18$) for GIM2D and a ratio of 1.02 for GALFIT ($\sigma \approx 0.08$). This trend continues towards fainter surface brightness, although at no point does the systematic size offset exceed 20%.

From Figure 3.8 (grey 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 fit by GALFIT (those fit only by GIM2D) are fainter than average, the parameters are discrepant even using GIM2D, whereas galaxies not fit by GIM2D (those fit only by GALFIT) are fit almost as well by GALFIT as other galaxies with the same surface brightness.

3.4.1.2 Results of Pure Spheroidal Simulations

Figure 3.9 shows the same plots as Figure 3.8 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 2 (0.2%) crashed in GALFIT, 56 (5.1%) ran into constraints; GIM2D crashed on 31 galaxies (2.8%), 36 (3.3%) additional fits ran into fitting constraints. 54 (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, both codes crashed on 4 (0.4%) galaxies in common.

When comparing Figure 3.9 to Figure 3.8, one should be aware that the X-axis is shifted by 2 magnitudes arcsec^{-2} towards fainter surface brightness, and that the dashed lines for clarity indicate 1σ instead of 3σ as in Figure 3.8. 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. Firstly, spheroidal profiles are in principle harder to fit due to the

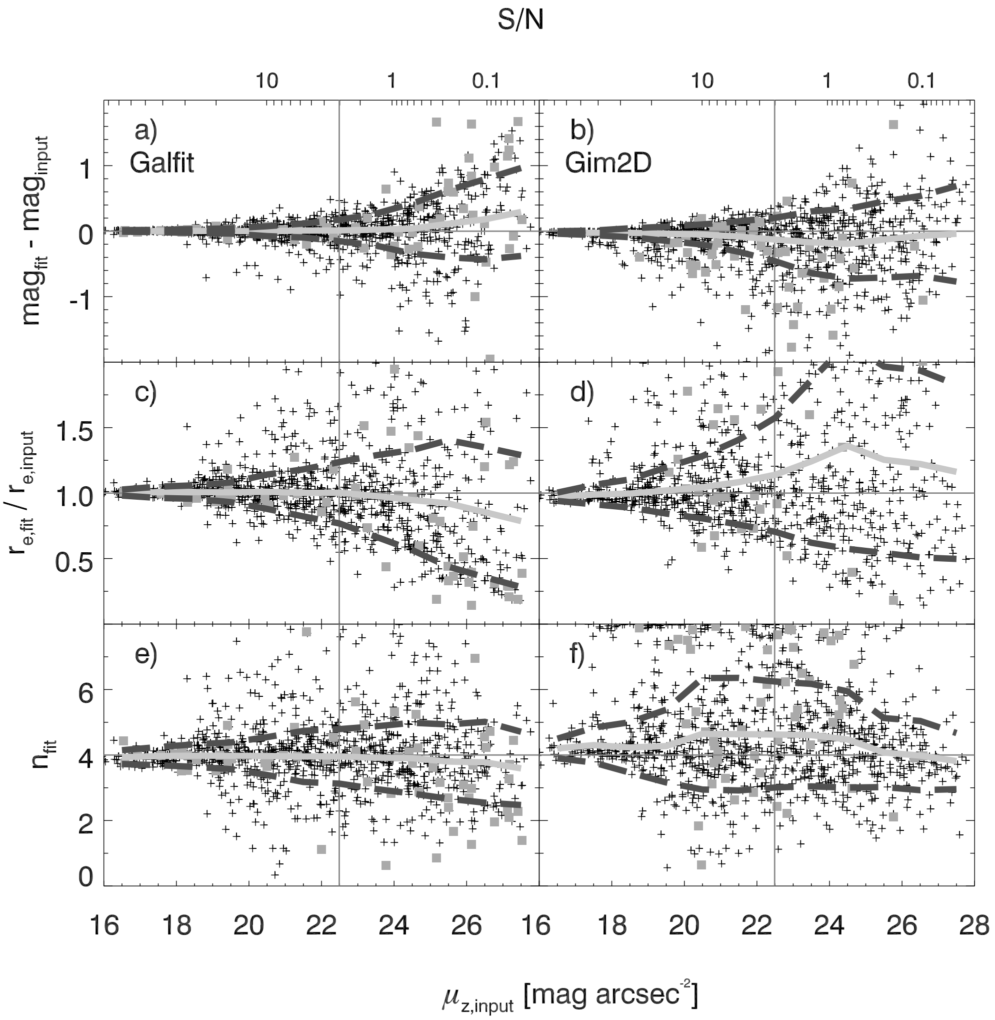


Figure 3.9 Code comparison for $n = 4$ galaxies. This figure is formatted in a similar way to Figure 3.8, but for the sample of simulated $n = 4$ galaxies. The X-axis is shifted by 2 magnitudes arcsec^{-2} compared to Fig 3.8; furthermore, in this plot the grey dashed line represent the 1σ limits. The 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.

importance of the outskirts of the light profile — this makes using an appropriate sky estimate much more important for a successful fit. Secondly, 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, $\sigma \sim 0.23$ 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 towards 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.

3.4.1.3 Deblending Effects

Given the significant differences in philosophy when it comes to the deblending techniques between GALFIT (multiobject fitting & masking) and GIM2D (masking only), we explored 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 where 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 3.10, 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 (panels a and b) 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 (panels c and d), it is clear that 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 is the magnitude deviation. In an attempt to disentangle the influence of distance and brightness, we try to correct for the systematics observed in panels c and d by removing the offsets and the slope, showing the results in panels e and f. It is clear that distance and brightness effects of nearest neighbor cannot be easily corrected, 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.

3.4.2 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 where $n = 1$ and $n = 4$ galaxies were intermixed with realistic clustering, sizes and magnitudes. Towards this

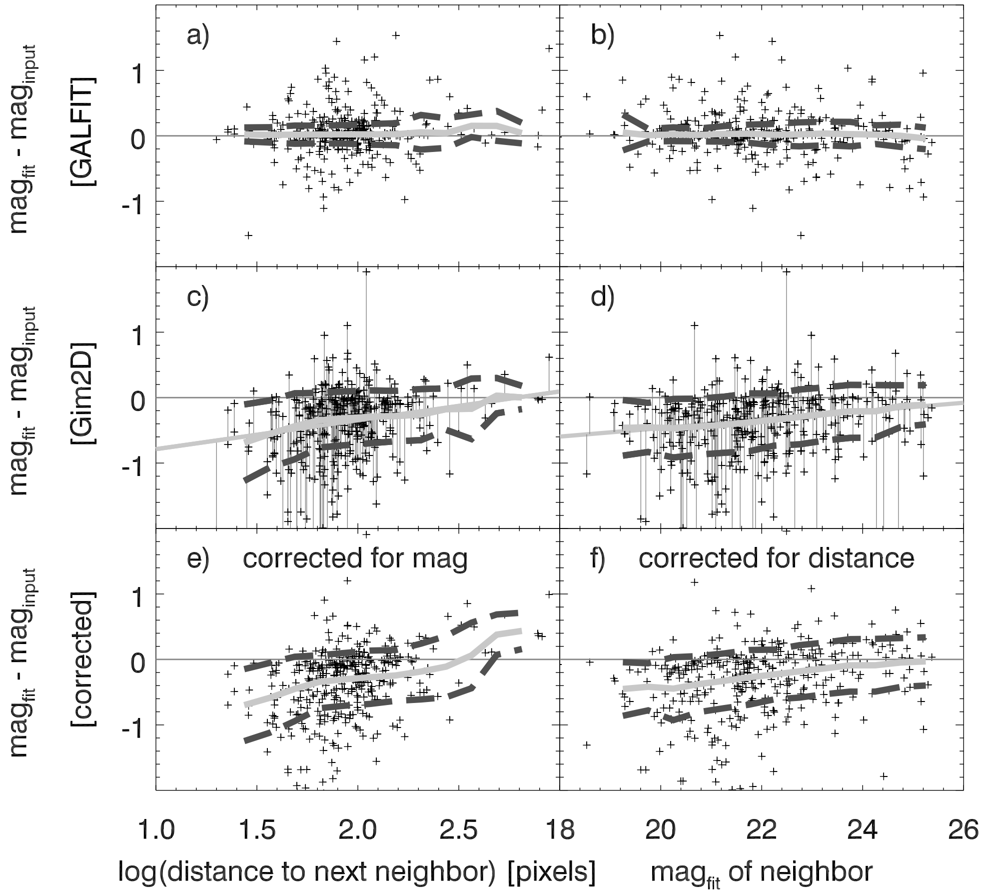


Figure 3.10 The impact of neighboring galaxies on fit results with GALFIT (upper row) and GIM2D (middle and lower row). The left column shows magnitude deviations from the simulated values as a function of the distance to the next neighbor; the right column shows the difference between the recovered 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 panels c and d by showing the distance dependence of the offset-magnitude relation residuals (panel e), and the magnitude dependence of the offset-distance relation residuals (panel f).

goal, 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; all others 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 3.11 and 3.12. 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, left) and McIntosh et al. (2005; spheroidal galaxies, right), to show which areas of parameter space are especially important for scientific analysis of data.

Inspecting Figs. 3.11 and 3.12, it becomes clear that GIM2D and GALFIT perform *equally well* for galaxy populations with clustering and properties typical of medium-depth cosmological *HST* surveys. GALFIT shows increased scatter and mild systematic offsets compared to the earlier 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 only slightly larger scatter and systematic offsets than GALFIT.

3.4.3 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 1-orbit depth GEMS survey and the 5-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 3.13 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 1-orbit depth *HST/ACS* data.

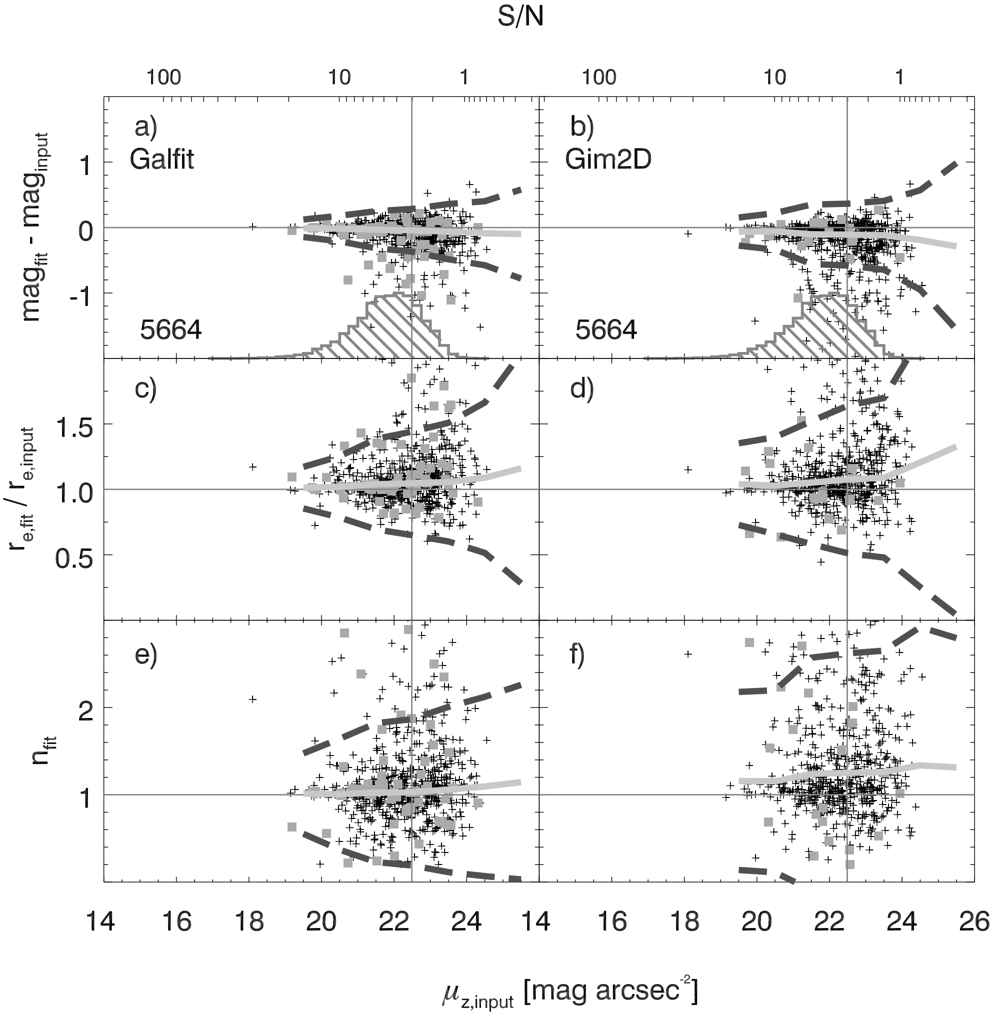


Figure 3.11 Fit results from re-simulated GEMS tiles (see §3.4.2), disk galaxies. Same as Fig. 3.8, but for a sample of 2 GEMS tiles that were re-simulated in order to create a more realistic distribution of galaxy parameters and object density (the results for the $n = 4$ galaxies in this sample are shown in Figure 3.12). In the upper panel we overplot the surface brightness histogram of the 5664 disk galaxies that where selected for analysis by Barden *et al.* (2005) showing where fitting accuracy is especially important.

3.4.4 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 realistic uncertainties given by how well the codes recover input parameters for simulated galaxies. In Fig. 3.14, 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

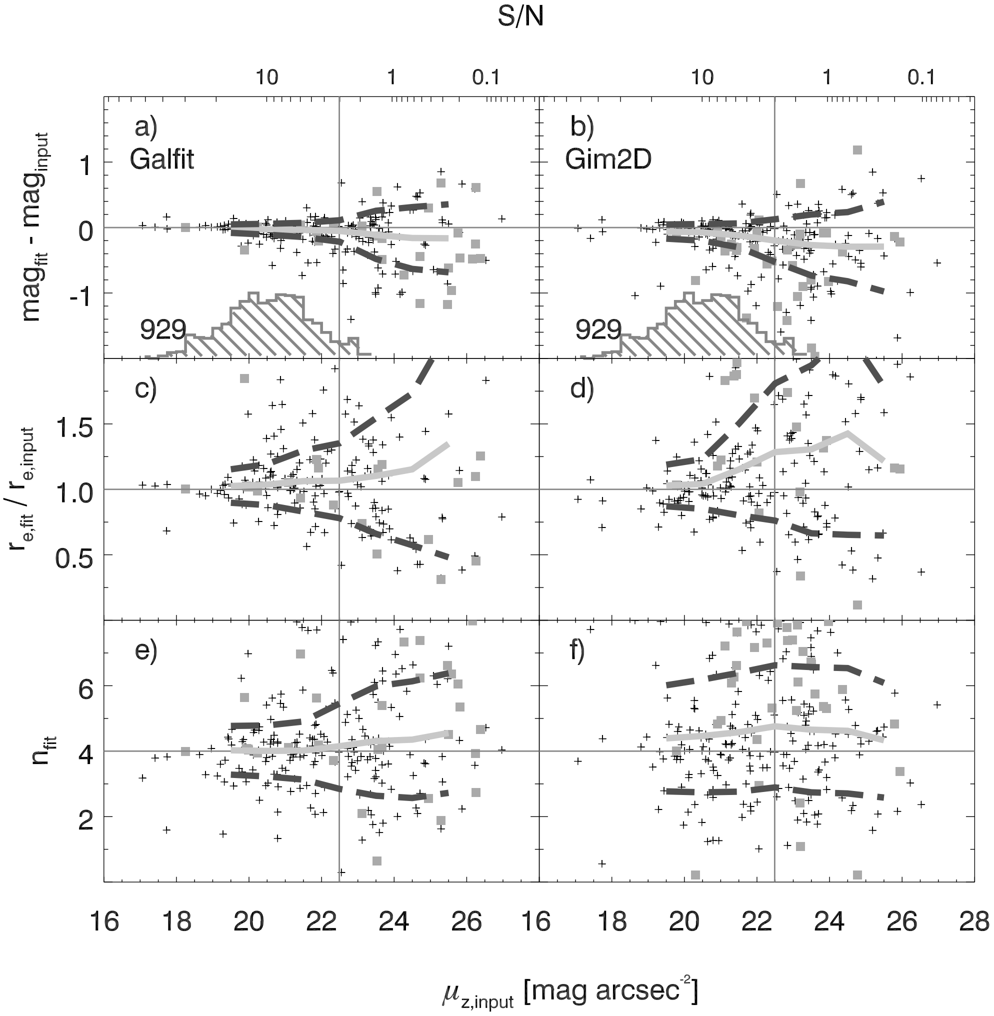


Figure 3.12 Results from re-simulated GEMS tiles, $n = 4$ galaxies. Same as Fig. 3.9, but for the $n = 4$ galaxies in the re-simulated GEMS tiles. The histogram in the upper panel shows the surface brightness distribution of the 929 red-sequence galaxies that were selected for analysis on McIntosh *et al.* (2005).

galaxies the deviation Δ is much larger than the error estimate σ[11]. Under the assumption that the error estimates are correct, σ/Δ should be >1 for 68% of the galaxies. Fig. 3.14 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. Fig. 3.14 shows no difference between the histograms of σ/Δ for GALFIT

---

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

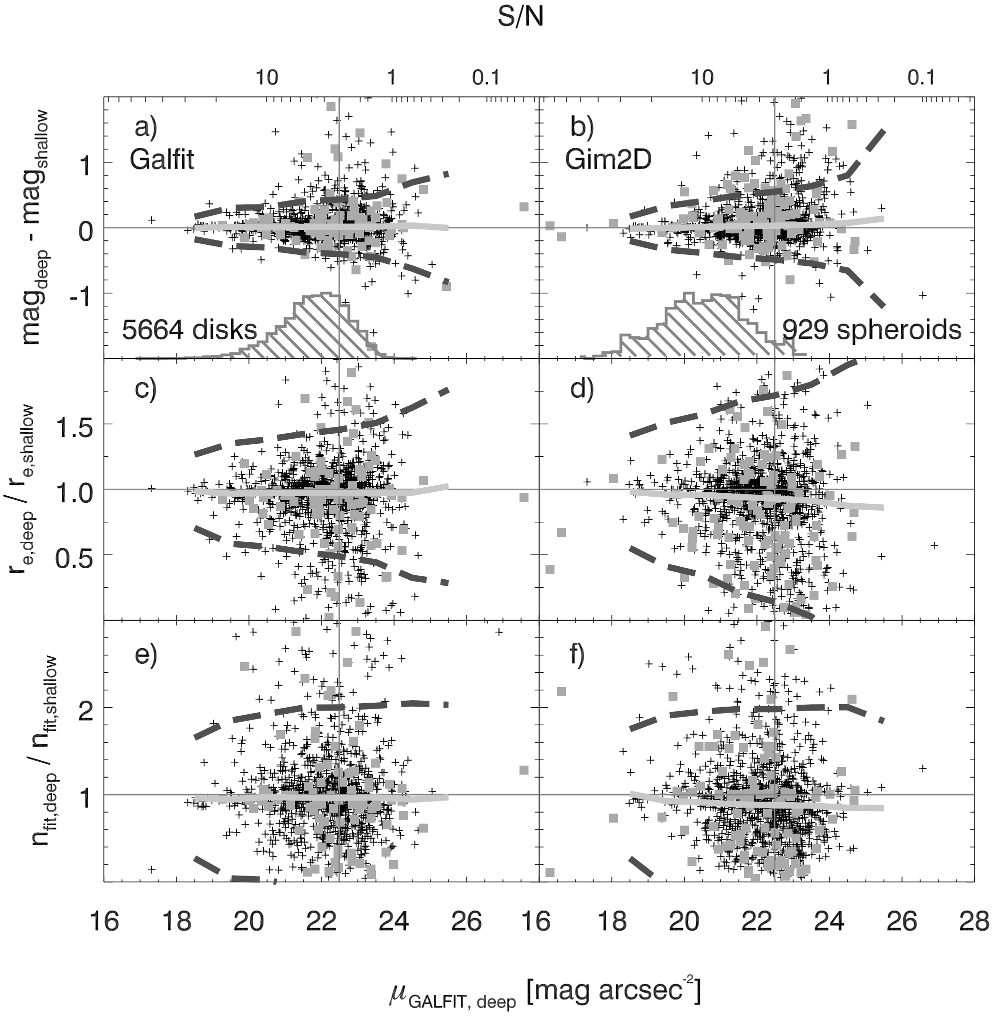


Figure 3.13 Comparison of fits to deep vs. shallow images. The left column shows results using GALFIT and the right column shows results for GIM2D. Both codes were run on the same sample of galaxies. Note that in this plot the X-axis shows the GALFIT surface brightness derived from the deeper GOODS data. 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 overplotted 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), right histogram) to highlight out the area of parameter space where fitting and independence of the image depth is particularly important.

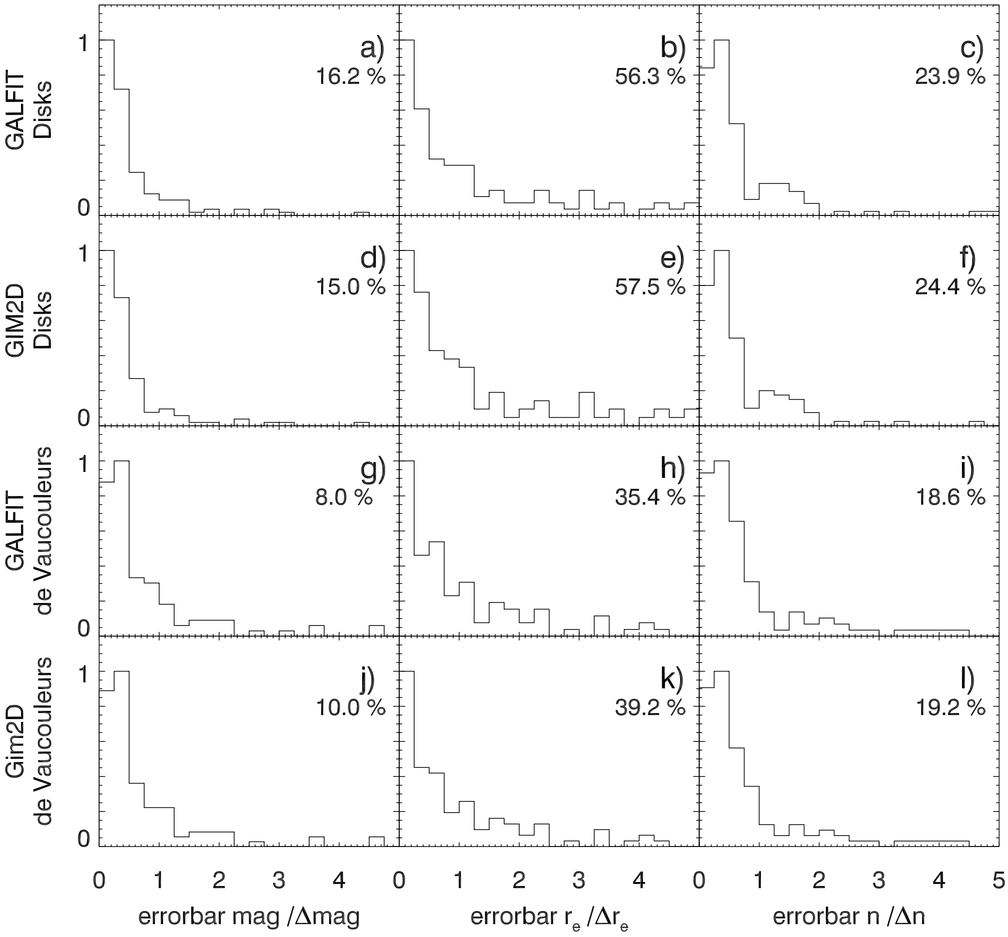


Figure 3.14 A comparison of estimated and real errors for two simulated samples of disk ($n = 1$) galaxies and spheroidal ($n = 4$) galaxies. Shown is the histogram of the errorbars σ devided by the deviation from the input value Δ. We show σ/Δ instead of the more intuitive quantity Δ/σ, which would show a very wide distribution and effects are 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 errors are dramatically underestimated by both GALFIT and GIM2D; the r_e uncertainties are significantly better estimated.

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, relying instead on the mean and width of the parameter distributions from Figs. 3.11 and 3.9 at the surface brightness of the galaxy in question.

The uncertainties given in Table 3.6 are calculated and interpolated by using the surface brightness μ and the Sérsic index n and the results from the simulated data (see §3.6 for details about this procedure).

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

- GALFIT is substantially less CPU intensive than GIM2D, reducing the cost and time of fitting large datasets.

- 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 datasets, 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.

3.5 Comparison with Pignatelli et al. (2006)

In this chapter we present an extensive and thorough test of the two different 2-D galaxy fitting codes GALFIT and GIM2D. In this section, we compare our results with Pignatelli *et al.* (2006), who compared results from testing those two codes with their own 1-dimensional profile fitter, GASPHOT.

Pignatelli et al. conclude 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 the Pignatelli *et al.* (2006), which we felt may significantly affect their conclusions.

- For the simulations examined in this chapter, they used the IRAF task `mkobject`, which, as we described in §3.2, is inexact for the inner pixels of a simulated galaxy light profile. The differences in the profiles can lead to systematic errors of 10-20% in the fitting parameters.

- In their paper, Pignatelli et al. allow the sky value to be fitted as a free parameter for all three codes. We argue in this chapter that this is a non-optimal 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 chapter 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 returns more accurate galaxy parameter values.

- Pignatelli et al. 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. 3.10). 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. Their argument that GALFIT does not deal well with blended galaxies is an artefact of the mode in which they chose to use GALFIT.
- Pignatelli et al. 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 false impression of the frequency and severity of GIM2D's difficulties with nearby neighbors. Also, it seems that Pignatelli *et al.* (2006) 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.

3.6 GEMS GALFIT Results

In this section, we present the GALFIT F850LP-band fitting results of all 41,495 GEMS objects that were found by SEXTRACTORẆe include fit results fr 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 image edge was taken. Table 3.6 shows the 10 first objects in the catalog and gives the format of the catalog published in the online version of this work. It includes the following values:

1. RA (1), Dec (2): RA and DEC, given by SEXTRACTOR(J2000).
2. tile (3): the GEMS tile in which the galaxies appears
3. Snum (4): the SEXTRACTOR catalog number of this object
4. GEMSID (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, halflight-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 WCS, defined north-to-east, PA_{WCS}) as well as their 'uncertainties'. These uncertainties are *not* the errorbars returned by GALFIT; as shown in §3.4.4 these errorbars 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 deg for PA). In the table published online and on the GEMS webpage, the uncertainties are saved 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).

9. f_{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 Fig. 3.15 and 3.2, 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 work to quantify the effects of completeness. Accordingly, we have made extensive suites of simulation catalogs available to interested users on the GEMS webpage 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 3.16 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 (plotted in grey) 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 3.6, one should be very careful when using their fitting results. All these galaxies are indicated by $f_{sci} = 0$.

Figure 3.15 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).

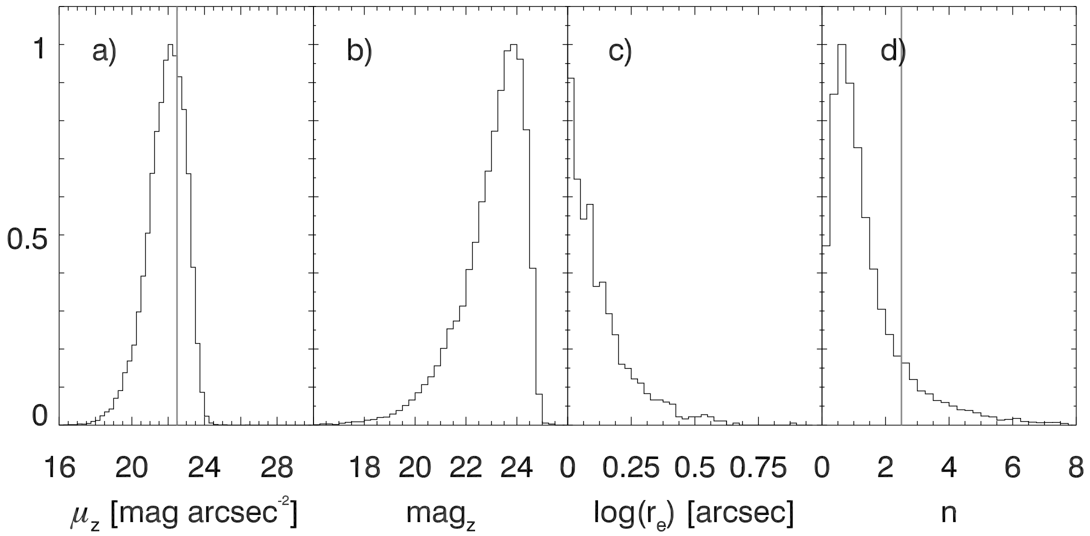


Figure 3.15 This figure shows histograms for 23,187 objects from the GEMS survey that were fitted without running into constraints when fitted by GALFIT and selected according to the selection criteria in Barden *et al.* (2005). From left to right we show surface brightness μ, apparent magnitudes, apparent sizes r_e (logarithmic scale) and Sérsic index n. For comparison we overplotted the surface brightness of the sky as a vertical line in the leftmost plot and the cut of n=2.5 which is frequently used to distinguish between disk- and bulge-dominated galaxies in an automated fashion.

Table 3.7 GALFIT fitting results for all GEMS galaxies

RA	DEC	tile	Snum	GEMS ID	PosX	PosY	sky	GALFIT fitting results					f_{con}	f_{sci}
								mag	r_e	n	b/a	PA (image/WCS)		
deg	deg				pix	pix	cnt	mag	pix			deg		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	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

This table shows the first 10 objects of the GEMS fitting results published in the online version of this work. For every object that was found by SEXTRACTOR we give RA, DEC, the GEMS tile name, the SEXTRACTOR number of that object, the GEMSID (containing RA and DEC) 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 5 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 PA in respect to the image (counted counterclockwise from vertical line according to the convention within GALFIT) and the world coordinate system (in brackets, north through east). Uncertainty estimates are derived from the simulations, following §3.4.4. 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) ($f_{sci} = 1$). The complete version of this table is in the electronic edition of the Journal. The printed edition contains only a sample.

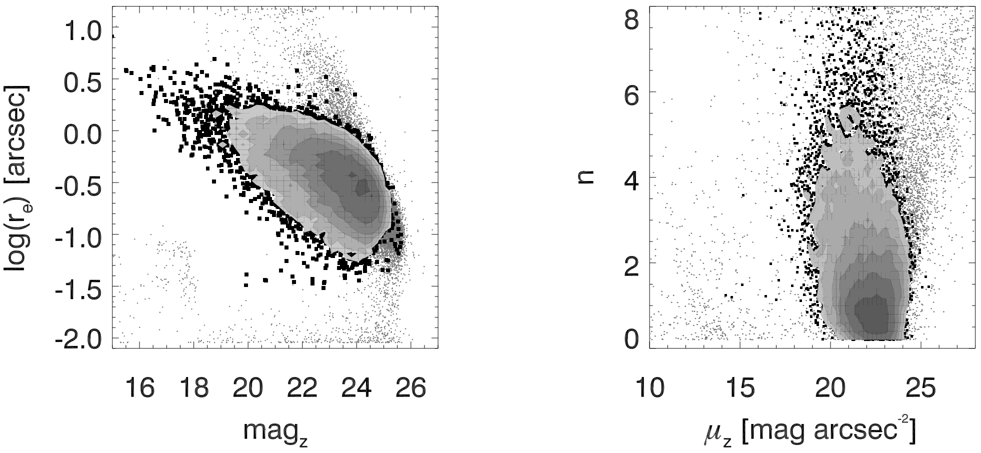


Figure 3.16 Here we show parameter distributions for all galaxies in the catalog published in Table 3.6, excluding the ones where the fit ran into any of the fitting constraints. Galaxies plotted in grey do not end up in the science sample according to the selection criteria used in Barden *et al.* (2005). 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.

3.7 Conclusions

In this chapter, 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:

- The performance of both GALFIT and GIM2D are 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.

- The setup of GIM2D, recommended by the author, using `'dobkg'`=‘yes’ and `'initparams`=‘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 of 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.

- Both codes are able to fit $n = 1$ galaxies reasonably well with relatively little bias (Figures 3.11 and 3.12). 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 galaxy populations and clustering typical of medium-depth cosmological *HST* surveys, there are no large differences between results obtained using GIM2D and GALFIT.

- The errorbars 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 errorbars.

- 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.

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

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

- GALFIT works best using an isophotal sky value given by GALAPAGOS instead of using the local values given by SEXTRACTOR.

- On the balance, 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.

Chapter 4
GEMS: Exploring the Evolution of Blue Spheroidal Galaxies

After a collision, a car is a wreck, not a new type of car.
DE VAUCOULEURS, 1974

4.1 Introduction

Trying to explain the formation of spheroidal galaxies, one can in principle picture two contrasting scenarios: the formation of spheroids (ellipticals) in a single-burst, monolithic collapse (e.g. Eggen *et al.* 1962; Larson 1974); or formation through hierarchical galaxy assembly, where spheroids are the result of merging pre-existing galaxies (e.g. Toomre & Toomre 1972; Khochfar & Burkert 2005; Kauffmann & Charlot 1998).

The Monolithic Collapse and Formation in a Single Burst. In this picture (e.g. Eggen *et al.* 1962; Larson 1974), spheroidal galaxies form at very early times, with the rapid collapse and conversion into stars of a large gas cloud. Gas cools efficiently during this process and star formation happens in a big, single burst. After this process, some process (possibly AGN-feedback) removes remaining gas from the galaxy, leaving behind a gasless spheroidal galaxy, although Naab *et al.* (2005) have shown that this is not strictly necessary when heating of shocks of further infalling material prevents gas cooling. Through later accretion of gas galactic disks can be built up, slowly turning these galaxies into more spiral-type galaxies with prominent disks.

If this picture is correct, spiral galaxies are formed from spheroids (or/and the bulges of spiral galaxies), indeed explaining the color distributions of these galaxies; spheroids contain mainly old stars, spirals contain many young stars dominating their light and making them look rather blue.

Two of the important observationally testable outcomes of such a scenario are that spheroidal galaxies should already be in place at high redshifts and that spiral galaxies should be more common today than they were in former times. Also, the red-sequence, formed by the red and dead spheroidal galaxies in the color-magnitude-diagram of galaxies (see Fig. 1.6), should already be present at high redshifts and should only evolve passively (becoming fainter & redder) when stars in the galaxies age. A late build-up of the red-sequence (e.g. the total mass of the red sequence) cannot be explained in this picture.

Hierarchical Galaxy Assembly. In the hierarchical galaxy assembly (e.g. Toomre & Toomre 1972; Khochfar & Burkert 2005; Kauffmann & Charlot 1998), spiral galaxies are created in collapses of gas in dark-matter haloes. These spiral galaxies can interact with neighboring galaxies and merge on a timescale of ~ 0.5 Gyr. During this 'merging' event, most of the available gas is drawn to the center and converted into stars (Sanders & Mirabel 1996). Yet, some inevitable fraction of the gas falls in at later times and rebuilds a star-forming disk which persists for many Gyrs — in clear conflict with the observations (Springel *et al.* 2005). Therefore, a feedback mechanism has to blow out the remaining gas, ceasing star formation and leading to the eventual formation of a red and 'dead' spheroidal galaxy that just passively fades and reddens with time, now that there is no gas left to form new stars. One of the main problems of the theoretical models following this ideas is, that the star-formation (SF) within galaxies has to be shut of by a (yet unknown) process in order to be able to create the most massive red-sequence galaxies observed today (AGN-feedback might be a possible solution).

This model would lead to a continuous build-up of the red-sequence since $z \sim 1$ through merging events during this period, turning blue spiral galaxies into massive red-sequence galaxies, leading to an increase in the total mass in red-sequence galaxies.

Put together, one of the observable key discriminants between these two scenarios would be whether the red-sequence does build up at late times or not, a late build-up would prefer the scenario of hierarchical galaxy assembly through galaxy mergers.

The build-up of the red-sequence. Today it is in fact known from observations that the red-sequence builds up at z<1 (e.g. Borch *et al.* 2006; Bell *et al.* 2004b; Brown *et al.* 2006, see figure 4.1). Red sequence galaxies from little new mass *in situ* by star formation, therefore, new mass must arrive on the red sequence through the cessation or truncation of star formation in typically blue star-forming galaxies. Some pissible physical mechanisms include:

- galaxy merging in the picture of hierarchical galaxy formation, both gas-rich ('wet') mergers of spiral galaxies and gas-less ('dry') mergers of already existing spheroidal galaxies (Barnes & Hernquist 1992);
- disk fading in bulge-disk-galaxies (Couch *et al.* 1998), making spheroidal galaxies more common with time;
- environmental effect, e. g. , gas-stripping from galaxies falling into galaxy clusters (Gunn & Gott 1972). This would lead to spheroids being more common in galaxy clusters, which indeed is observed (Dressler 1980).

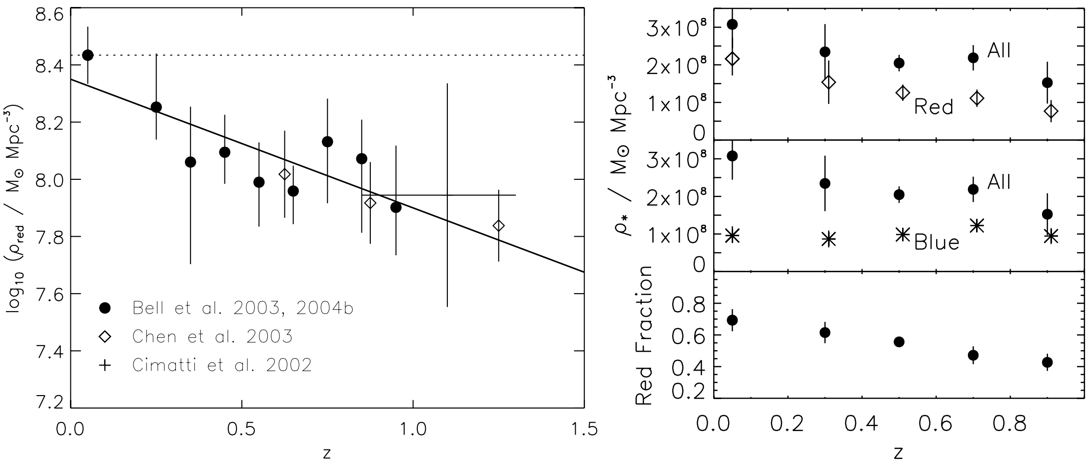


Figure 4.1 Build-up of the red-sequence since $z \sim 1$: The left plot (Bell 2004), that puts together results from different surveys (Bell 2003; Bell *et al.* 2004b; Chen *et al.* 2003; Cimatti *et al.* 2002) shows the evolution of the stellar mass in red-sequence galaxies over the last 8 Gyrs. The solid line shows a rough fit to the total stellar mass density in red sequence galaxies in the Cole *et al.* (2000) semi-analytic galaxy formation model. In agreement with the galaxy formation model, a clear increase in the stellar mass on the red sequence is demonstrated since $z \sim 1$. The right plot (Borch *et al.* 2006) shows the integrated stellar mass density as a function of redshift from the COMBO-17 (Wolf *et al.* 2003a) survey. The total for all galaxies is shown by filled circles, for red-sequence galaxies by diamonds, and for blue cloud galaxies by asterisks. In the lower panel the fraction of mass in red sequence galaxies is shown as a function of redshift. As one can clearly see, the total mass in the red-sequence rises with time whereas the total stellar mass in blue galaxies does not, running counter to the naive expectation that the bulk of the mass increase should be in the blue star-forming galaxies.

To gain further insight into the manner of the growth of the red sequence, it is important to study the population of galaxies which are plausibly transforming from blue and disk-dominated to red and spheroid-dominated.

A natural prediction of the hierarchical galaxy assembly would be a population of blue spheroidal galaxies, representing an intermediate stage of evolution as the young stars, created during the merging process, fade and redden and the timescale τ_{redden} for this process is longer than the timescale τ_{merger} for the dynamical assembly into a spheroid-like galaxy (see § 4.5 for a detailed discussion of these timescales). Such systems are very rare in the local Universe. However, as shown by Menanteau *et al.* (2004), such systems were much more common at early times. To examine these systems in detail, especially at relatively high redshift to be able to observe their evolution, one needs a big redshift survey that provides restframe colors for all galaxies, and high-resolution (e. g. *HST*) data, enabling the observer to examine the morphologies and sizes of galaxies. This combination of datasets is available through e. g. the COMBO-17 (Wolf *et al.* 2003a) and GEMS (Rix *et al.* 2004) surveys. As one of the science goals in the GEMS survey, we have studied the properties of these blue spheroids (BSGs).

The following sections will describe the estimation and analysis of restframe images (§4.2), technical issues, e. g. , PSF-matching and smoothing, identification of BSGs (§4.3) before we finally show our results from this analysis in §4.4 and discuss our findings in §4.5.

4.2 Restframe Images

In order to estimate galaxy evolution over a large range of redshifts, it is useful to analyze galaxies at a fixed restframe (intrinsic) wavelength. Morphological k-correction is important (Abraham 1998) as galaxies show different properties at different wavelength, e. g. star forming regions and spiral arms. e. g. it is known that galaxies show different sizes in different wavelength bands, making restframe very important for this work, as much of the analysis is based on galaxy sizes. The easiest 'correction' possible can be done by measuring galaxy properties in the observed band closest to a given restframe band as e. g. done by the GOODS team in Ravindranath *et al.* (2004) or in GEMS by Jogee *et al.* (2004). In this work, we take a different approach: we use the high-resolution GEMS *HST/ACS* data with its 1-color information and the redshifts derived from COMBO-17 (Wolf *et al.* 2003a) to estimate restframe images on a pixel-to-pixel basis. This is possible, because optical color-color diagrams are degenerate, which is known as the classical age-metallicity-dust degeneracy. All these effects can redden galaxy light, but luckily for us they all redden the light by roughly the same amount. We can use this to our advantage and estimate the restframe colors for each galaxy on a pixel-to-pixel basis from only **one** color that GEMS provides.

There are two key steps toward creating useful images of a galaxy in the restframe: i) construct and characterize a set of stellar-population models that we used to 'translate' the observed color into restframe colors (§4.2.1); and ii) construct and optimize the process used to estimate colors on a pixel-to-pixel basis, and to create the final restframe image (§ 4.2.2).

4.2.1 The PÉGASE stellar-population models

In order to be able to convert between observed and rest-frame, and between different observed frames, a grid of stellar populations with a variety of properties and redshifts is required. We have constructed such a grid using the PÉGASE stellar population model (see Fioc & Rocca-Volmerange 1997, for an earlier version of the model). PÉGASE creates a stellar population with a star formation history according to $\psi = ce^{-t/\tau}$, for all $t < A$, where A is the age, τ is the e-folding time of the star formation history, and c is an arbitrary normalisation constant related to the stellar mass of the system. For all $t > A$, the star-formation-rate (SFR) is $\psi = 0$. In total, we constructed 45 different templates

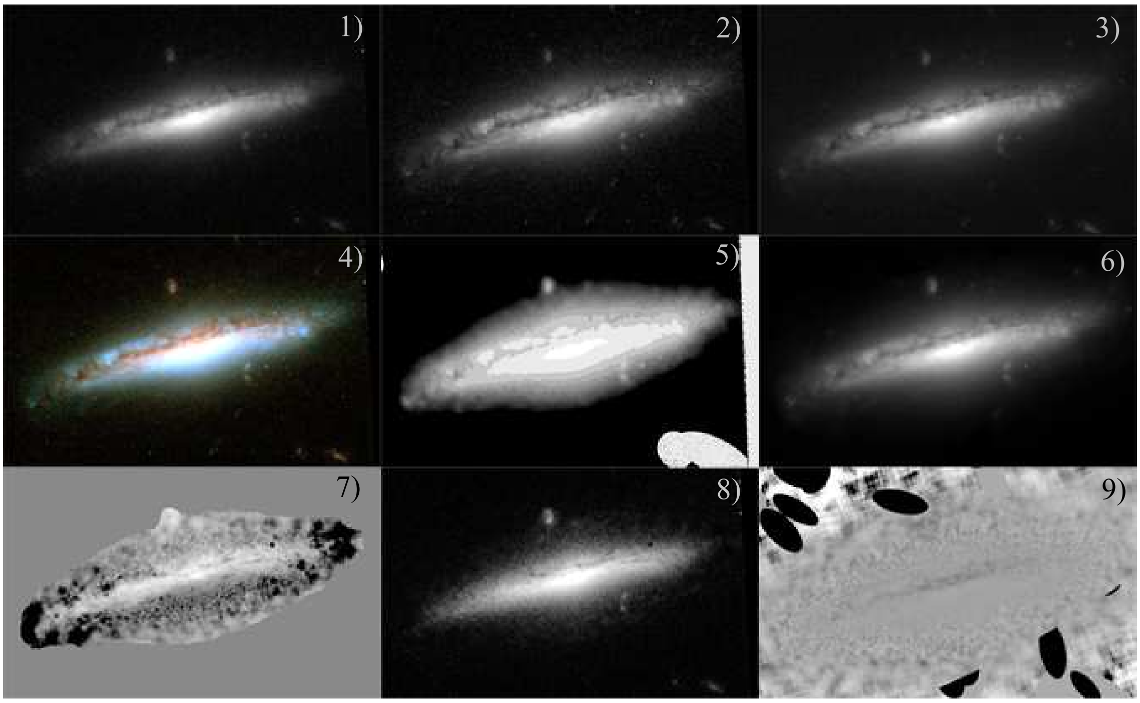


Figure 4.2 These images explain how the restframe g-band images were derived and tested. the rightmost column shows images from testing the pipeline on deeper GOODS data (Giavalisco *et al.* 2004), the left and middle columns show images used for this work directly. Details are explained in the following:
1) z-band image: The z-band image for this galaxy from GEMS.
2) V-band image: The V-band image for this galaxy from GEMS.
3) i-band image: The i-band image from the deeper GOODS survey that was used to test the restframe estimation.
4) A 2-band composite color-image from the color-images used for the GEMS-skywalker (see also chapter 6).
5) Image to control the smoothing of the observed GEMS images. From bright to dark this image shows in which step the S/N of a certain pixel was high enough in order to be used to derive the colormap (frame 7) that was used for the estimation of the restframe image. 14 steps were taken in total. For all pixels with still too low S/N, the mean color of the most smoothed 20% of the pixels was taken as color. In this image, one can also see the masking that was used during the whole process in the lower right corner.
6) estimated i-band image: This is the observed-i-band image as it was estimated from using GOODS z- and V-band. Images like this were used to test the pipeline as they were compared to the observed i-band images from GOODS.
(...caption continues on next page...)

[To figure 4.2:]
7) This is the colormap (V-z) of the galaxy as derived from z-band and V-band images after the smoothing. The scale of this image is arbitrary.
8) The final restframe image derived from the pipeline (restframe-g-band). This image was used in the further analysis of this work, especially for deriving galaxy parameters like Sérsic index and r_e from profile fitting using GALFIT (see §4.3 and especially §3for details). As is obvious from this example, this image is noisier than the original images. This is due to the fact that a) it is derived by adding a noisy color map to one of these images and b) we used smoothing to increase the S/N of each pixel. This smoothing introduces some structures, especially in low-S/N regions. One can also see that the dustlane in this particular galaxy can not bee seen very well in the restframe image. This is mostly an effect of smoothing, small structures as dust are smeared out by this step.
9) The relative deviation between estimated and observed GOODS i-band images. The scale of this image is arbitrary, the fluffy structure especially at the outskirts of the galaxy are created by the smoothing steps within the pipeline. As one can see, the created image represents well the true observed i-band image. There is no dependance on galaxy brightness and only slight effects of pixel color can be seen, especially along the (red) dustlane. The masking in this image was different from what can be seen in 5), this is because these tests were carried out on GOODS where the masking was different due to different image depth.

with different values of τ and A, selected to give the best match to the run of colors as a function of redshift in COMBO-17 (see Fig. 4.3).

It must be noted that these grids were tuned to reproduce the global colors of galaxies in COMBO-17, yet are applied pixel-by-pixel to GEMS galaxies in this work, where bursts of star formation and large variations in dust amount/geometry are expected. Only being able to use 1 color in GEMS to probe stellar population/dust variations, we cannot have any more than a 1-dimensional set of models (i.e., our 1D sequence of 45 models). Given that we needed to define a 1D set of models, we then chose to tune the models to get COMBO-17's global colors as accurate as possible, instead of choosing a perhaps more physically-motivated set of models which a priori already severely violated the observed global colors of galaxies in COMBO-17.

Three additional ingredients were added to provide a better match to galaxy colors as a function of redshift in COMBO-17: a template-dependent metallicity, selective dust extinction (where the amount of extinction is template-dependent), and a 'gradual shutoff' of star formation towards recent times. The run of metallicity with template number is shown in Figure 4.3. Galaxies dominated by older populations are more metal-rich in this scheme. This was designed to maximize the strength of the 4000 break for the reddest galaxies.

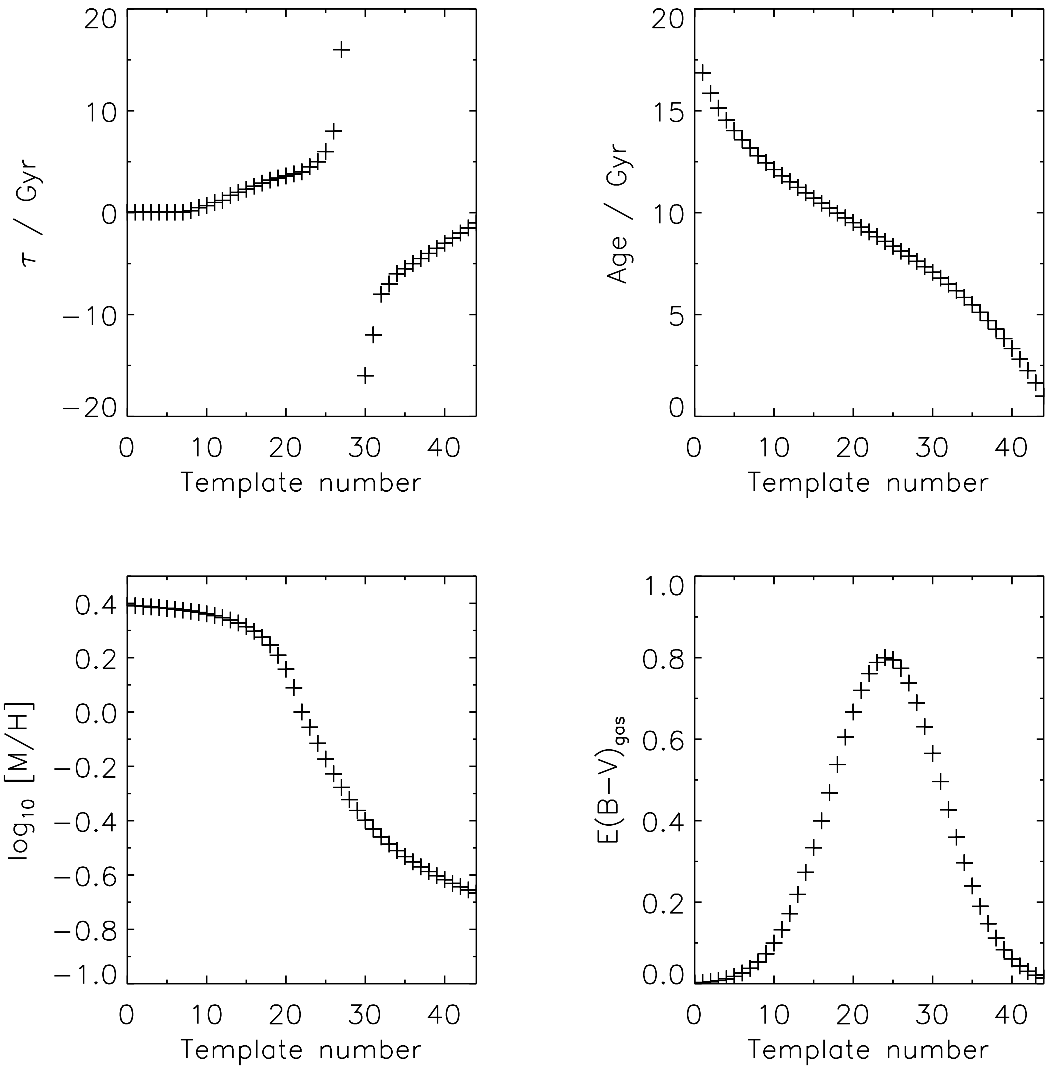


Figure 4.3 The properties of the stellar population grid used to convert between different observed and rest-frame passbands as a function of redshift. The grid is a 1-dimensional sequence of a variety of star formation histories, ages, metallicities and dust contents. See text for details.

Selective extinction is added according to Calzetti *et al.* (1994).

$$\begin{aligned} A_\lambda/E(B-V)_{\rm gas} &= 1.17(-1.857 + 1.040/\lambda[\mu m]) + 1.78 \quad [\lambda > 0.63\mu m] \quad (4.1) \\ &= 1.17(-2.156 + 1.509/\lambda[\mu m] - 0.198/\lambda^2[\mu m] + \\ &\quad 0.011/\lambda^3[\mu m]) + 1.78 \quad [\lambda < 0.63\mu m] \quad (4.2) \end{aligned}$$

where $E(B-V)_{\rm gas}$ is the reddening in the gas phase. The reddening in the stars is roughly 1/2 of the gas phase reddening in this model to simulate the clustering of young stars inside and near the region in which they were formed. The amount of gas extinction as a function of template number is shown in Fig. 4.3.

Finally, in order to reduce the amount of blue flux (required to match the behavior of V-I vs. R-I observed frame colors in COMBO-17) it was necessary to introduce further suppression of the light from very young stars. This was done in a somewhat artificial fashion: the SFR was artificially reduced from 100% of the value given by the τ model 200 Myr **before** the epoch of observation to only 20% of the value given by the τ model **at** the epoch of observation (where the fraction is a linear function of the time before epoch of observation, i.e. age). A physical interpretation for this type of model would be the obscuration of star-forming regions in optically-thick dust clouds which cannot be detected at optical wavelengths, and the successive emergence of young stars from these dust clouds with time (e. g. Charlot & Fall 2000).

4.2.1.1 Testing the PÉGASE models on COMBO-17 and GEMS data

The colors as estimated from the grid of models were tested on different sets of data, both comparing to galaxy colors from COMBO-17 and GEMS as a function of redshift. A full analysis if these tests is far beyond the scope of this work, we will discuss and show the most basic results here.

From tests on COMBO-17 colors (see Fig. 4.4 for part of these results), it is clear that the grid estimates colors in a pretty good fashion when estimating $R-I$ vs. $V-I$. This is by construction: for GEMS F606W-F850LP we wanted to get the color-color plots in this part of wavelength space right, so the grids were created in a way that they match these colors best.

When comparing the estimated colors from the grid to F606W-F850LP colors from GEMS (see Fig. 4.5 for part of these results) as a function of $U-V$ rest frame and $U-B$ vs. GEMS observed frame, the agreement is quite encouraging.

When we use GEMS magnitudes from both z-band and V-band (photometry within an ellipse of $2r_e$ semi-major axis) to predict COMBO-17 colors in restframe $U-V$ (see Fig. 4.6, blue galaxies left plot, red galaxies right plot), we get good results with only small offset and a 0.1 mag scatter. This is close to the limit of what we suggested to be acceptable, but given all the uncertainties, this should be good enough to estimate restframe images on a pixel-to-pixel basis. The M_B magnitudes can be predicted from GEMS magnitudes with 0.1 mag offset such than GEMS is brighter than COMBO-17 and $\lesssim 0.2$ mag scatter.

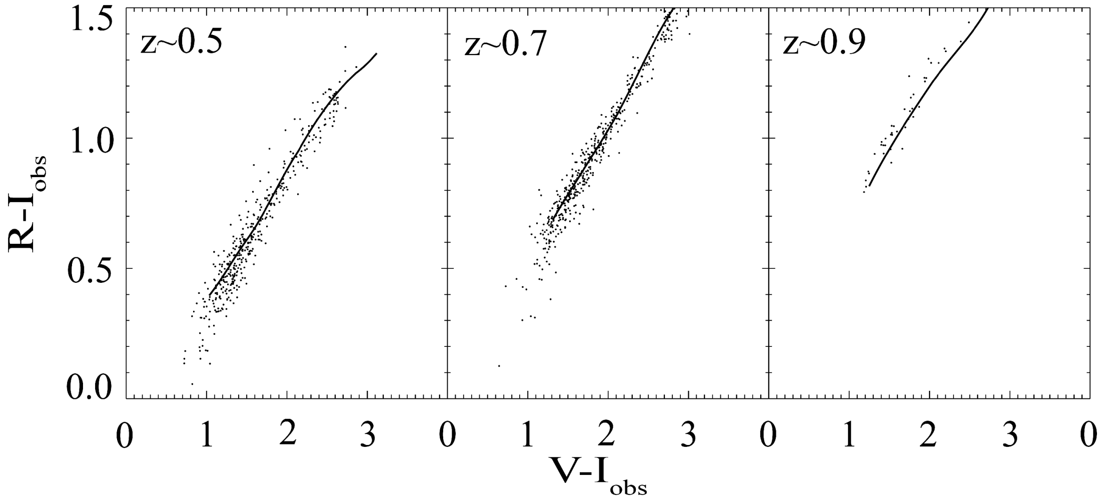


Figure 4.4 Testing the PÉGASE models on COMBO-17 data: These plots show color-color diagrams $R - I$ vs. $V - I$ for real galaxies fromCOMBO-17(data points) and the colors predicted by the models (as solid line). One can see that models and predictions agree quite well at all redshifts.

The most basic test that we carried out was based on the exact pipeline that we used to estimate the restframe images, including masking and smoothing (see §4.2.2 for details about this procedure). Using this pipeline and GEMS z- and V-band images, we estimated i-band observed frame images on a pixel-to-pixel basis for all galaxies in the shared data with the GOODS project using the same set of population models that we used for restframe estimation (see images 1 and 2 in Fig. 4.2 for z- and V-band images and image 6 for the i-band image as estimated using our pipeline). For the GOODS project, there is also i-band data (F775W) available, which allowed us to directly compare our estimated images with 'reality' as shown in image 3 in Figure 4.2; image 9 shows the relative deviation between estimated and observed image. One can see that in principle the estimation works, but has some problems e. g. at the ~ 0.1 mag level in the dustlane of this galaxy and in the low S/N outskirts of the galaxy where it introduces smoothing features. Figure 4.7 shows results from this test as a function of redshift. The y-axis shows relative deviations of the estimated image when compared to the GOODS image. Each data point represents the mean deviation of all pixels within an ellipse of semi-major axis $2r_e$ (as derived from the initial z-band fit using the procedure discussed in §3) for one galaxy in our test sample. The bars on the right side show typical errorbars for redshift estimation of COMBO-17 ($\delta z \sim 0.03$) and the standard-deviation of the distribution of the deviation pixel-by-pixel, although individual distributions can be much wider (shown by the bigger errorbar representing one of the largest standard deviations in our sample). As we can see from this plot, the estimation of i-band observed frame images works with 10% systematic offset and 20% scatter. While testing different grids of stellar population models, it

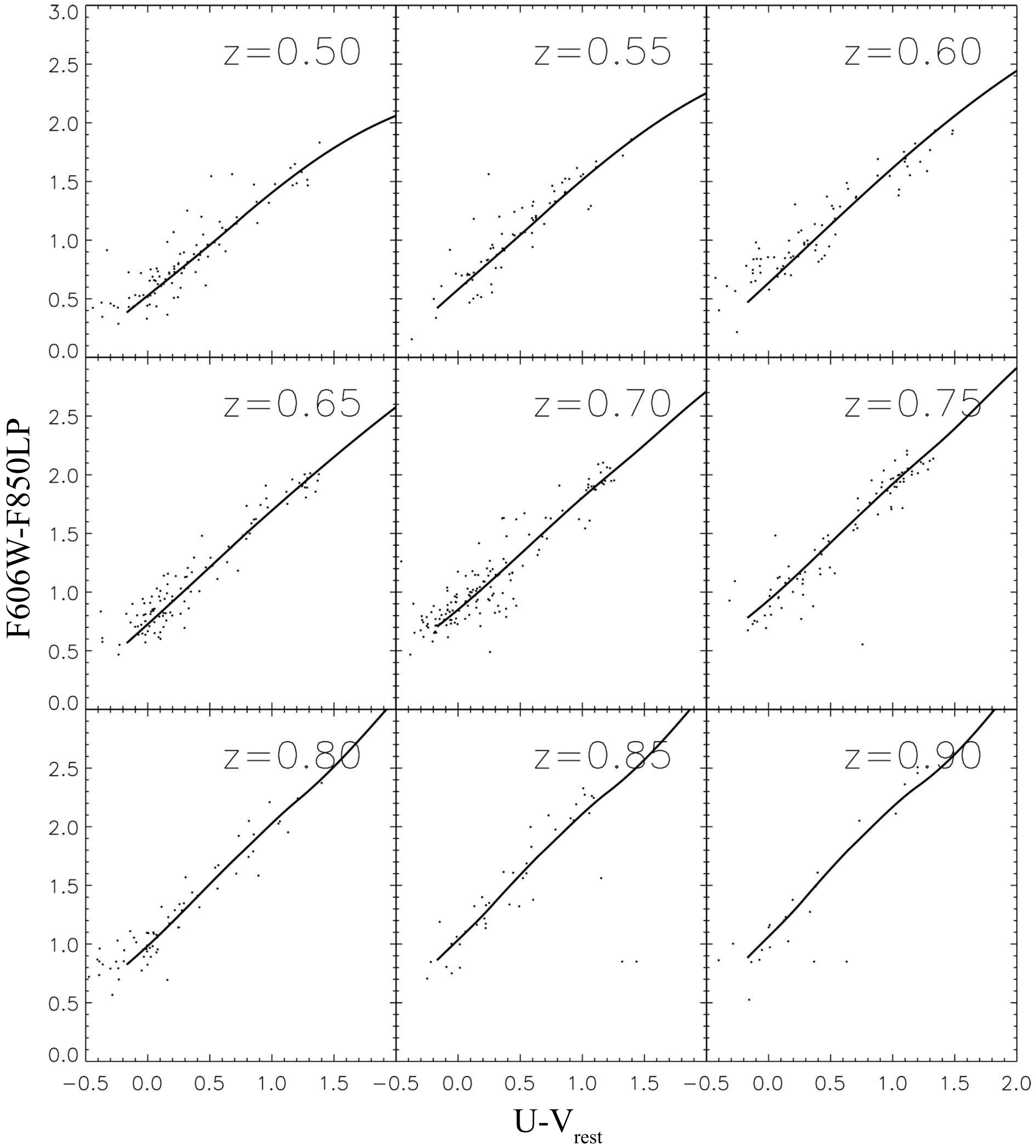


Figure 4.5 Testing the PÉGASE models on GEMS data: We compare F606W.F850LP colors from GEMS with $U - V_{rest}$ from COMBO-17. Also here, the models agree well with the observations, although scatter is obvious.

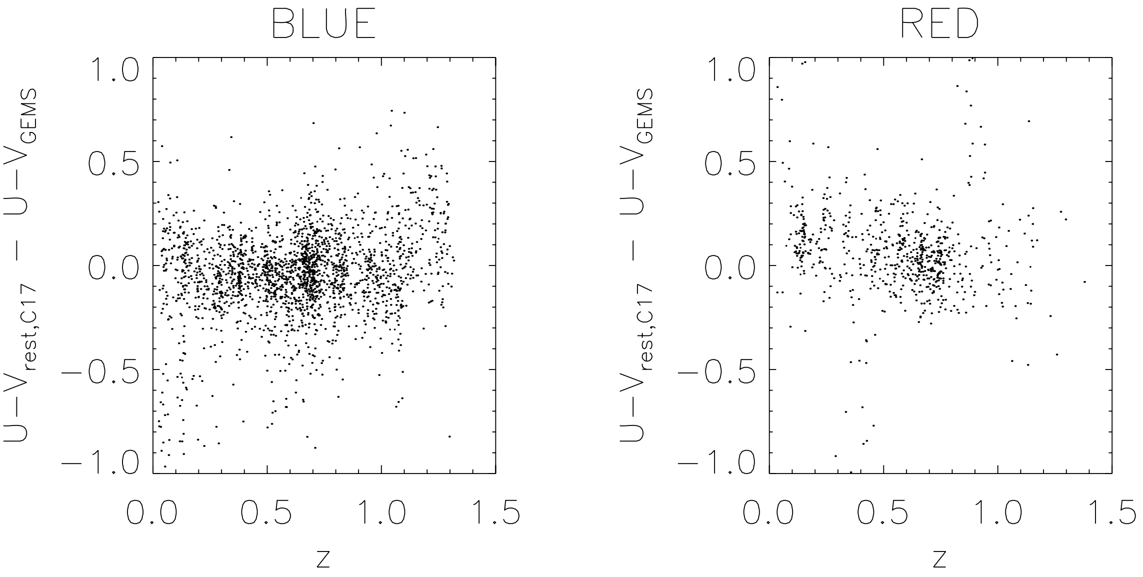


Figure 4.6 Testing the PÉGASE models on GEMS vs. COMBO-17 data: We use GEMS colors to predict COMBO-17 $U-V_{rest}$ colors and compare these with the colors actually observed by COMBO-17. Results are shown as a function of redshift. As one can see, there are small systematic trends with redshift, but all in all, the esimated and observed colors agree quite well.

was not possible to reduce these effects, e. g. it was not possible to remove the systematic offset seen in this plot.

4.2.2 Estimating the restframe g images

A second key ingredient to restframe image creation is the challenge of deriving accurate colormaps from noisy data, which can then be translated to restframe color images using the stellar population grids described in § 4.2.1. It is clear that some kind of smoothing is required to extract color information for low S/N areas. It is also clear that any smoothing will introduce unwanted artefacts in pixels with strong surface brightness or color gradients. We have developed a complex, well-tested pipeline which will be described in detail step-by-step in this section.

- We try to estimate restframe images for all galaxies for which we successfully found both COMBO-17 information (e.g. redshift which is obviously needed for this work) and GEMS z-band fitting results in order to be able to mask out neighbors successfully (this will be explained in further detail below). In total, we start with a sample of 9315 galaxies. We are able to create restframe images for 9163 of these (98.3%).

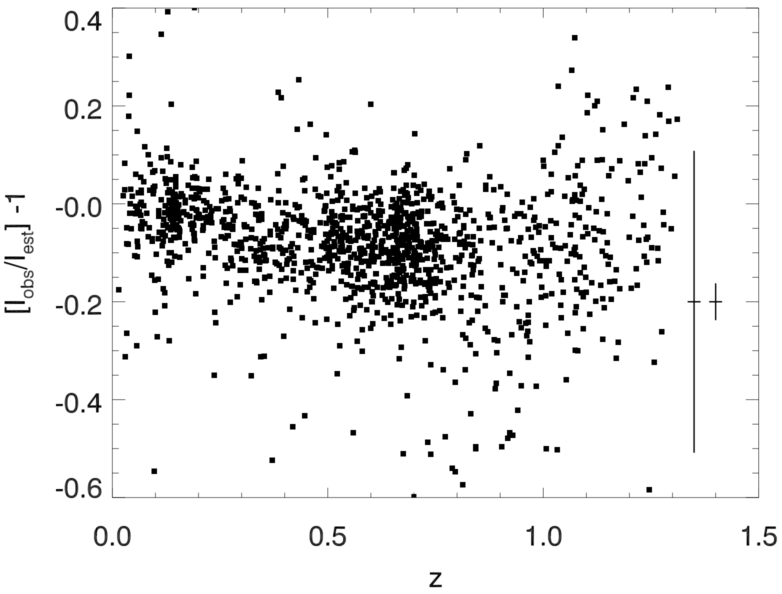


Figure 4.7 Testing the PÉGASE models on estimating GOODS i-band: We show mean deviation of all pixels within $2r_e$ of a galaxy when comparing the estimated i-band observed frame image to the real i-band image from the GOODS survey as a function of redshift. The ideal deviation of 0 is indicated as the long dashed line. Median errorbars are indicated on the right side, the bigger errorbars shows one of the extreme examples to guide the eye for reliability of these colors. As one can see from this plot, there are small systematic trends over redshift that we could not get rid off when testing different types of stellar population models, but in general, the true and the estimated color agree on a 10-20% level.

- In the initial step, the images are sky-subtracted. We subtract the isophotal sky values as explained in detail in §3.3.1.2, although here, this routine was also run on V-band data in order to be able to sky-subtract these images. The typical uncertainties of this sky estimation, as shown in § 3, is of order 1%.

- We fit 2-D gaussian functions to both images and derive the object's center from this fit in each band. This has to be done in order to remove small image shifts between V- and z-band due to poor registration of the images, which would severely influence the colormaps derived. We shift the V-band image to match the z-band image on a $1/10$ pixel level. As this is especially important for galaxies with high concentration and less important for galaxies with low concentration, we carry out this step only for galaxies with $n > 2$ from the initial z-band fits. This prevents problems when deriving the center for the galaxies with low n, where the center from fitting a gaussian is not derived very securely. If the offset was bigger than

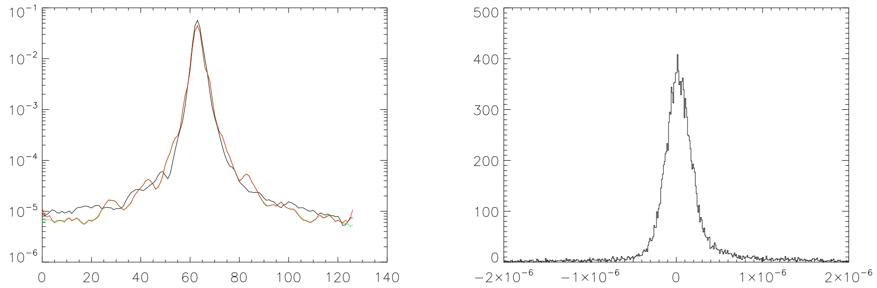


Figure 4.8 This figure shows the the quality of the PSF correction. In the left figure, the black line represents a cut through the original V-band PSF, the red line represents the V-band PSF after the correction which has to be compared to the original z-band PSF which is shown in green and can hardly be seen because the correction is indeed very good. The right plot shows a histogram of the relative deviations between the corrected V-band PSF and the z-band PSF. Numbers being on a level of 10^{-6} clearly show that the PSF correction applied is very good.

3 pixel, though, the images were **not** shifted (345 galaxies). In these cases, it was clear from visual inspection that the gaussian was not a good fit to the data, we expected image offsets to be much smaller than this. In total, 2437 galaxies (26%) where shifted with typical offsets of 0.1-0.2 pixels (and an additional bump around 1 pixel as some GEMS images show 1 pixel offset between V and z band).

- We convolve the V-band image with a correction PSF in order to show the same PSF as the z-band image. This has to be done in order to prevent artificial color-gradients of galaxies as one of the images is ‘smeared out’ more severely than the other. Figure 4.8 shows results from these correction. As immediately becomes obvious from these plots, the correction is very well understood and is good at a 10^{-6} level.

- We create a mask image for each galaxy (see image 5 in Fig. 4.2), masking out cosmics and neighboring objects (as known from the GALAPAGOS pipeline as blown-up SEXTRACTOR ellipses). This is important as neighboring objects (especially ‘false’ neighbors at a different redshift) may influence the colormap.

- In order to increase the signal-to-noise for each pixel, we iteratively smooth both V- and z-band image, holding the total flux fixed. As we want the S/N to be $S/N > 20$ in both bands for each pixel, we initially only keep the pixels where this is true. The rest of the pixels are smoothed in $n = 14$ steps (over a box of $[2 * n + 1]^2$ pixels), until their S/N was high enough. The noise in each step is calculated as

$$\sigma_{image} = \sqrt{image + sky + readnoise^2 + \sigma_{sky}^2 + \sigma_{smooth}^2 - \sigma_{theory}^2} \tag{4.3}$$

where 'image' is the image after the smoothing, 'sky' is the background sky that was subtracted in the initial step, 'readnoise' is the readnoise of the GEMS frame, σ_{sky} is the uncertainty from the sky level estimation (0.05 count at a sky level of 18.3 counts, $\sim 0.2 - 0.3\%$), σ_{smooth} quantifies the smoothing structures as introduced in the smoothing itself (estimated by running the same smoothing scheme on empty patches of sky, see images 7-9 in Figure 4.2 for an impression of this structure) and σ_{theory} is theoretical noise that was suggested from the smoothing of empty sky defined as $\sqrt{\frac{(4.98)^2+18.34}{pixels}}$, where 4.98 is the readnoise and 18.34 the background of our data and $pixels$ is the number of pixels over which we smoothed the image in that step. σ_{theory} has to be subtracted here, because it is included in both sky and σ_{smooth} and counting it double would bias the S/N map.

Image 5 in Figure 4.2 shows a greyscale map which represents in which smoothing step the S/N was high enough so the values were accepted for this pixel, from white being high-S/N initially without smoothing to dark grey being high-S/N in the last smoothing step. Black pixels where too-low S/N even in this last step. If no high-S/N pixels were apparent, the estimation of the restframe image was skipped (20 galaxies).

- We have to set pixels with $flux < 0$ to a very small value in order to be able to convert the image into a magnitude map. The conversion is done by using

  $$mag = zeropoint - 2.5 \cdot \log(flux/0.03^2) \tag{4.4}$$

  where 'zeropoint' is the zeropoint of each image and 0.03 is the pixelscale of the GEMS images so that the created image shows magnitudes arcsec^{-2}. By subtracting the images from V- and z-band, a colormap $(V - z)_{obs}$ is created, low-S/N images (see step before) are assumed to have the mean color of the 20% most-smoothed pixels of the object (see image 7 in 4.2), taking care of possible color-gradients within galaxies.

- This colormap is translated using the grid of stellar population models from PÉGASE (see §4.2.1 for details) to a $g_{rest} - z_{obs}$ image. Besides the 45 models, the used grid contains information about 17 filters and 33 redshifts which are used to interpolate between the two closest redshifts in the grid to the redshift of the object. During the estimation process, we check whether the model is strictly monotonic in the color-color-diagram. If is is not, the interpolation might fail for those pixels that happen to fall into this region of the color-color-diagram (231 objects had warnings due to this reason, we estimated the images non the less).

- The new $g_{rest} - z_{obs}$ colormap is added to the initial (non-smoothed) z_{obs} image in order to get a restframe g_{rest} (Sloan g-band) image. The initially subtracted background sky is added to this image in order to have a 'proper' image for GALFIT which needs background flux to successfully fit the galaxies. For the fit, we used the same GALFIT setup that was tested in §3, but GALFIT did not derive a sigma-image

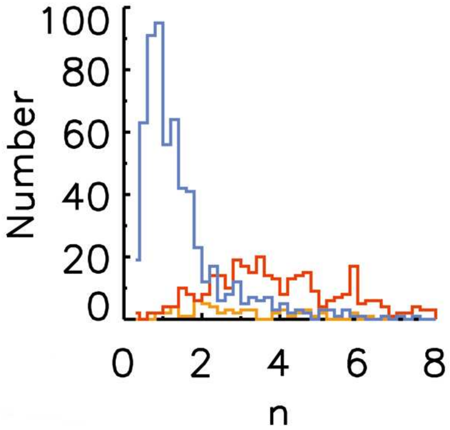


Figure 4.9 This plot (from Bell *et al.* (2004a) shows the distribution of Sérsic indices of visually classified E/S0 galaxies (red), Sb – Sdm galaxies (blue), and Sa galaxies (yellow). Although automated classification by dividing early- form late-type galaxies by using $n = 2.5$ as a cut, returns relatively robust (80% reliability) classifications of galaxies, significant contamination of both samples is still apparent. The inclusion of S0 galaxies in either of the late- or early-type galaxy bins does not significantly affect this conclusion.

internally, but we created one in this step and implemented it in the GALFIT start file. This new sigma image was necessary because the estimated restframe images had count levels which were sometimes substantially different from z_{obs}, and whose uncertainty could not be estimated from the count level + readnoise + gain alone. The sigma image was created using all S/N maps from the smoothing steps, the image itself, readnoise and uncertainties in the sky estimation. In the end, to avoid strong pixel-to-pixel variations, the sigma image was smoothed by convolution with a gaussian (FWHM=2 pixels).

The product of this pipeline was a g_{rest} image for each galaxy, the according GALFIT startfile and all information needed for the fit. In total, a number of 9163 have been created and prepared for the fit (see next section).

4.3 Identification of Blue Spheroidal Galaxies

As has already been described in detail in § 3, we fit Sérsic profiles to all galaxies using GALFIT. From § 3 it becomes clear, that GALFIT is capable of recovering the true parameter of the galaxies using the same masking and simultaneous fitting scheme as tested in § 3, simply using a different sigma image for differently weighting the individual pixels during the fit. The fitting results derived could, in principle, be used to automatically classify galaxies by using the Sérsic index n, classifying all galaxies with $n > 2.5$ as early-type and galaxies with $n < 2.5$ as late-type, as was done by several authors before (e. g. McIntosh *et al.* 2005, and others). This method indeed gives a rough classification for very large samples of galaxies without too much work, a scheme that is required and very much appreciated today as large galaxy surveys (especially *HST* surveys like GEMS, COSMOS and STAGES) exist and visual classification of **all** galaxies is clearly challenging. Nontheless, Figure 4.9 shows Sérsic indices for a sample of 1500 visually classified galaxies from Bell *et al.* (2004a). As is obvious from this plot, when using automated classification by cutting at $n = 2.5$, significant contamination from disk galaxies in the 'early-type' galaxy sample would be apparent. In order to remove this contamination and to get a clean sample of spheroidal galaxies which was strongly required for our analysis, visual classification of at least **some** galaxies was required. We decided to classify all galaxies with $n > 2$ (from the same fitting pipeline that was used for the code-testing in §3), to be slightly more conservative than the usual automated classification and to remove the contamination from our sample, without having to visually classify galaxies with $n < 2$. For time reasons, only half of the galaxies have been classified at this stage. For the rest of this analysis, we only use this half of the galaxy sample where the classification was completed (3501 galaxies).

For our classification we used several different images from our dataset:

- The most important image was the restframe image as derived by §4.2 as this was the image that was used for galaxy fitting during all this analysis.
- Not to be affected by the smoothing in the estimation of the restframe images, which hides weak internal structure, e. g. spirals that were an important criteria during the classification, we used the nearest observed image to restframe-B (the observed V-band image up to a galaxy redshift of z=0.5, the observed z-band image for galaxies beyond z=0.5.
- During the restframe estimation, a color-map of the galaxies was created. This colormap was used for the classification, because it sometimes showed internal structure that was not visible in any of the observed images. Some (clear) spirals only became obvious in the colormap.
- We used the fitting-residual from the z-band fits as they sometimes revealed spiral arms, dustlanes and/or other internal galaxy structures.

Although our internal classification was more complicated, the important step for this work is the separation of late-type and early-type galaxies (while still removing galaxy interactions and irregular galaxies). As early-type galaxies we use ellipticals and S0 galaxies, we will call the sum of these 'spheroidal' for the rest of this work. Some examples for galaxy classifications are show in Figure 4.11. As becomes obvious from this plot, some galaxies with $n > 2$ are clearly disk galaxies with a central cusp of light. These were removed from our sample of 'early-type' galaxies.

At this stage, only one classifier gas classified the galaxy sample (B. Häußler). Visual tests have suggested a good degree of reliability; however, results from multiple classifiers would be clearly desirable and will be secured and analyzed soon.

4.4 Results

Figure 4.10 shows color-magnitude-diagrams for galaxies from GEMS and COMBO-17 for different redshift bins. As all plots in this chapter, the values for magnitudes and colors are derived from COMBO-17 (Wolf *et al.* 2003a), GEMS values were used for Sérsic indices and galaxy sizes.

In the left hand column, galaxy classes from automatic classification using Sérsic indices to separate 'early-type' ($n > 2.5$, black symbols) from 'late-type' ($n < 2.5$, grey symbols) are shown. The red sequence, the blue cloud and the gap in between (the 'green valley') show up clearly. In the right column of plots, we overplot the visually classified spheroidal galaxies as classified according to §4.3. As one can see, not all high-Sérsic-index galaxies have been classified as early-type (black dots without surrounding diamonds or boxes), so the visual classification was required to remove this contamination and agrees roughly with the color-bimodality observed in the color-magnitude diagram. Red galaxies turn out to be mostly visually classified spheroids, though edge-on spirals with strongly reddened colors do also form a fraction of the red-sequence. Blue galaxies are basically spiral (disk-dominated, non-spheroidal) galaxies. This is true at all redshifts, although significant contamination is obvious.

By using the 'green valley', indicated by the solid line in the left plots and the dashed lines in the right plots, the visually classified spheroids can be divided into a red (black boxes) and a blue (black diamonds) population. Please note, that this cut 'evolves' with redshift, the slope of the cut (-0.08, parallel to the red-sequence) is derived by Bell *et al.* (2004b), but a Y-offset has been applied to better fit the 'green valley' for the different redshifts. Figure 4.11 shows examples of different types of galaxies that were classified by eye. Recalling that only galaxies with $n > 2$ were visually classified, it helps to understand why some galaxies with high Sérsic index n where clearly classified as disks. Most of them are disk galaxies with a central cusp of light, e. g. a faint AGN, a bright central cluster or a foreground star. Removing the clear disk galaxy contamination from the sample is important in order to get a sample of spheroidal galaxies as cleanly as possible.

We can use figure 4.10 to estimate how much BSGs at a given magnitude have to redden in order to form a typical red-sequence galaxy. Although passive evolution includes significant changes of luminostity (fading) this number should be a good estimate for the reddening that is required. This scheme is indicated in the right column of plots as the vertical lines from the BSGs to the fit of the red-sequence. We measure this number for all spheroidal galaxies, RSGs as well as BSGs.

Figure 4.12 shows magnitude-size-diagrams for the visually classified spheroidal galaxies. For each redshift we individually estimate the fit to the population of red spheroids (left plots and grey squares in right plots), by excluding very bright ($M_{B,abs} < -17\ mag$; 6 galaxies in the redshift bin $0 < z < 0.4$), very big ($r_e > 100\ kpc$; 1 galaxy in the highest redshift bin having several much brighter neighbors, a combination which makes the galaxy very hard to fit. The z-band fit returns a size 15 times smaller) and very small galaxies ($r_e < 0.3 kpc$; 7 galaxies in total, size is too small to represent a reliable

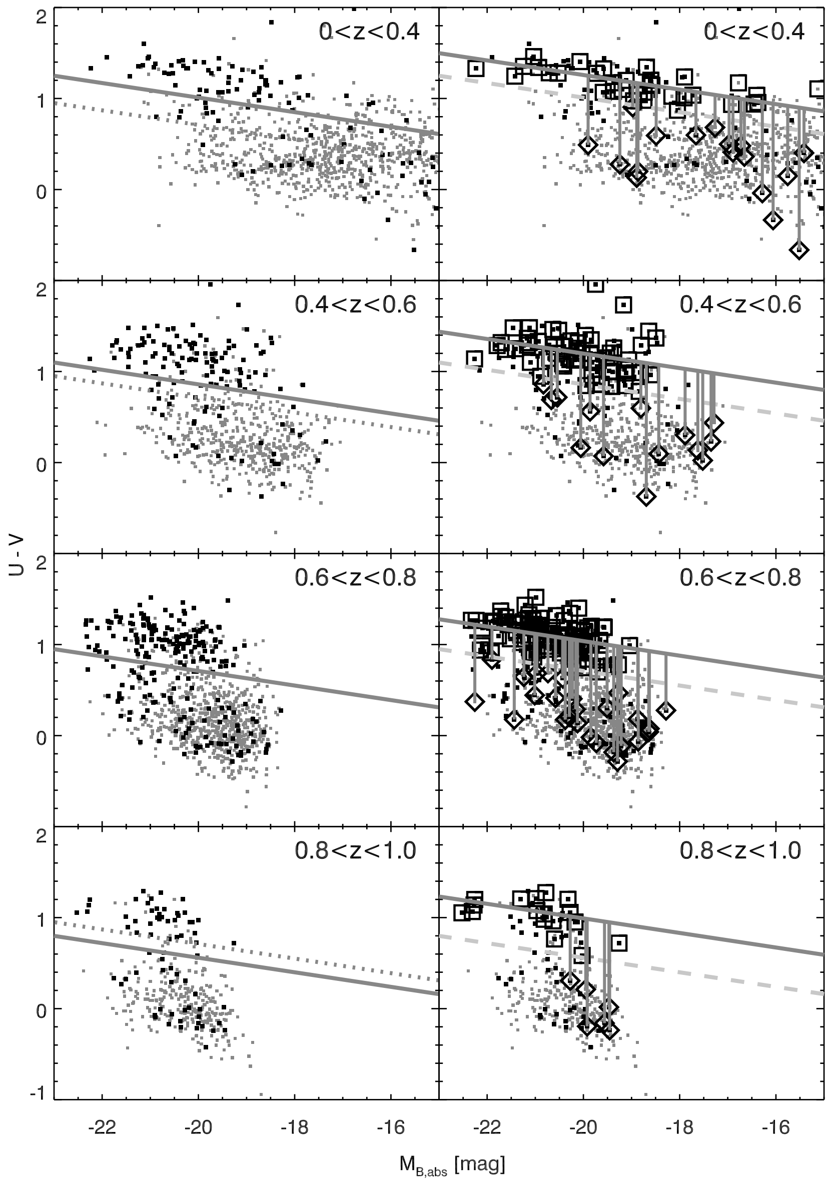


Figure 4.10 The left plots show color-magnitude diagrams of galaxies from the GEMS survey for different redshifts as labeled in the plots. X- and Y-axes, as for all plots in this chapter if not labeled differently, show parameters from COMBO-17 (Wolf *et al.* 2003a). The different grey scaled dots show the galaxy type as derived from automatical classification cutting galaxies at a Sérsic index $n = 2.5$, light grey indicating low n galaxies (late-type), black indicating high n galaxies (early-type). All galaxies with $n > 2$ were classified by eye. The diamonds indicate galaxies classified as early-type galaxies in this classification. (...caption continues on next page...)

[To figure 4.10:] As one can see, the classification roughly agrees with the color bimodality of galaxies. The diagonal lines in each plot show the cut that was used to define blue (bottom) and red (top) galaxies, directly defining the sample of blue spheroidal Galaxies. These cuts are also shown as dashed lines in the right plots. To guide the eye, the dividing line of the $z \sim 0.7$ bin is indicated in all plots as a dashed line. On the right side, visually classified spheroids have been overplotted, squares representing red spheroids, diamond representing blue spheroids as devided by the dashed line. The solid diagonal line in the right plots show the fit to the red-sequence. The slope (-0.08) is taken from Bell *et al.* (2004b), the y-offset is determined to fit the data. We measure the distance of all blue spheroids to the fit of the red-sequence (vertical solid lines) asking: "At a given magnitude, how much do galaxies have to redden in order to form a typical red-sequence spheroidal?"

fit to the data). This fit is shown in both plots as solid line. In the right plot, we also show the population of blue spheroidal galaxies (diamonds). For each galaxy we measure the distance to the fit of the RSGs at a given fixed size. Assuming that galaxies just fade passively during their evolution, this answers the question how much galaxies (at a given fixed size) have to fade (or brighten) in order to form a typical red-sequence spheroidal. As one can see from this plot, in fact most BSGs would have to *brighten* in order to form a typical red-sequence galaxy.

Given the fact that the magnitude of galaxies changes significantly during their lifetime, it is a 'bad' parameter to estimate evolution effects. A better parameter to do this are galaxy masses that should not evolve if passive evolution is assumed. We use stellar masses of galaxies derived by Borch *et al.* (2006) from the color-M/L-relation (Bell *et al.* 2003b; Borch *et al.* 2006) by using stellar population models. These masses have been tested dynamically on a small subset of galaxies from the COMBO-17 survey.

In figure 4.13 we show mass-size diagrams for GEMS galaxies. As becomes obvious from these plots, RSGs (grey open boxes) on one hand and disk galaxies (both visual and automated classification using $n < 2$, grey dots) on the other hand form two very distinct classes of galaxies. At a given r_e, RSGs are more than 10 times more massive than typical disk galaxies at all redshifts (i.e. spheroids have $\sim$ 10x surface density within r_e than disks). In contrast to other projects, our sample is large enough to be able to divide the identified BSGs into different subpopulations, which turns out to be important in order to examine these types of galaxies in more detail. Using a mass-density cut of $3 \cdot 10^8 M_\odot kpc^{-2}$ (dashed line) which was suggested by Kauffmann *et al.* (2006) as the mass-density where star formation histories of galaxies changes qualitatively, separating disk-dominated galaxies from bulge-dominated systems, we can divide the blue spheroids (black open diamonds) in 2 subclasses, low- and high-mass-density.

- The high-density BSGs at a given mass have similar sizes as the RSGs. Given that McIntosh *et al.* (2005) found no evidence for evolution in the mass-size relation of spheroids from $z = 1$ to the present day, a natural expectation is that high-density BSGs will become present-day red spheroid-dominated galaxies.

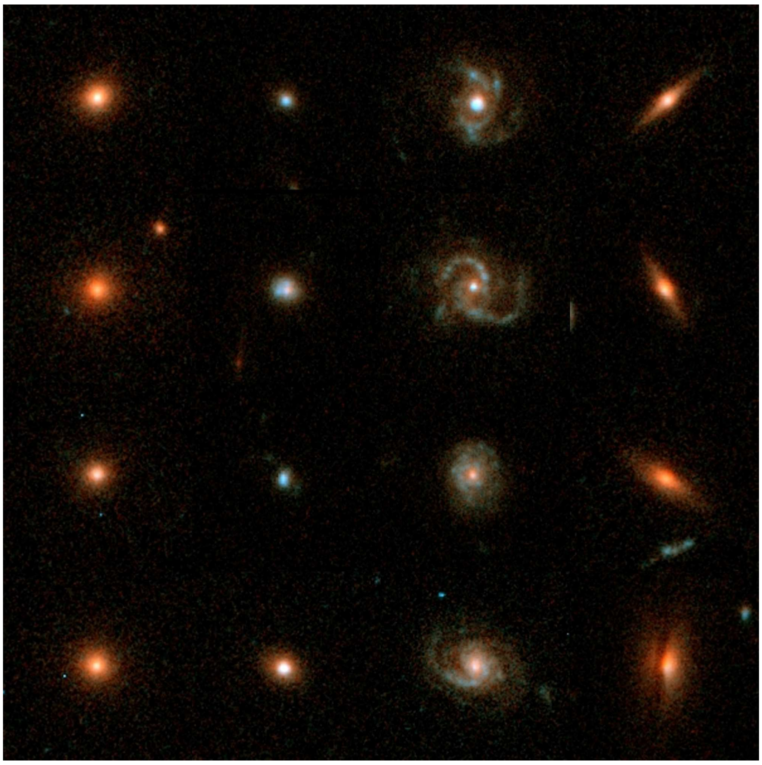


Figure 4.11 This figure shows examples for different types of visually classified $n > 2$ galaxies from the redshift bin $0.6 < z < 0.8$. Recalling that only galaxies with $n > 2$ were classified, these examples explain why many of these galaxies have high Sérsic index, although they are clearly classified as disks. From left to right we show: red spheroids, blue spheroids, blue disks, red disks. Please note that the images do only roughly show similar physical scales (all galaxies being from the same redshift bin $0.6 < z < 0.8$). Disk galaxies with high n turn out to be disk galaxies with a central cusp of light, either faint AGN, bright central clusters or other features. A single-component Sérsic fit to an exponential disk profile with some extra-light in the center returns higher Sérsic index. The difference between blue and red high-n disks is the viewing angle. Red disks turn out to be edge-on spirals that are reddened by dust. In some of them the dustlane is clearly visibly (e. g. the bottom right example). This figure illustrates clearly that automatic classification by using the Sérsic index as the dividing-criteria is insufficient to unambiguously identify spheroidal galaxies. Visual classification is strongly required to rule out significant contamination by disk-dominated galaxies in the spheroid-dominated galaxy sample.

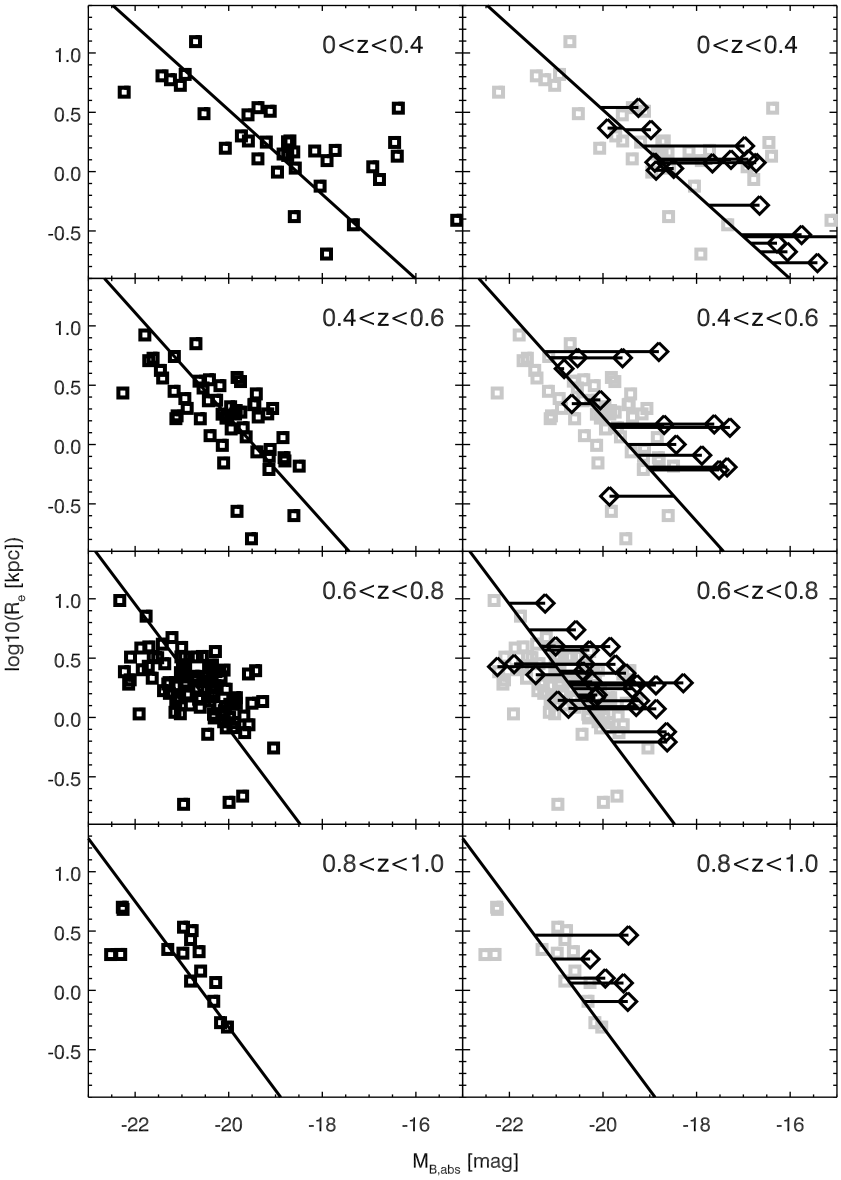


Figure 4.12 These plots show magnitude-size diagrams for the same redshift bins as in in Figure 4.10. The left columns of plots show red spheroidal galaxies and the line fit to these data points. The same data is shown in grey scale in the right column of plots. Overplotted, we show the distribution of blue spheroidal galaxies as diamonds. We measure the distance of these galaxies the the fit to the RSGs. How much do galaxies at a given size have to fade in order to form a typical RSG. When we talk about ‘fading’ of galaxies, one should note that in fact **most** of the BSGs would have to **brighten** in order to have true properties similar to typical red-sequence galaxies.

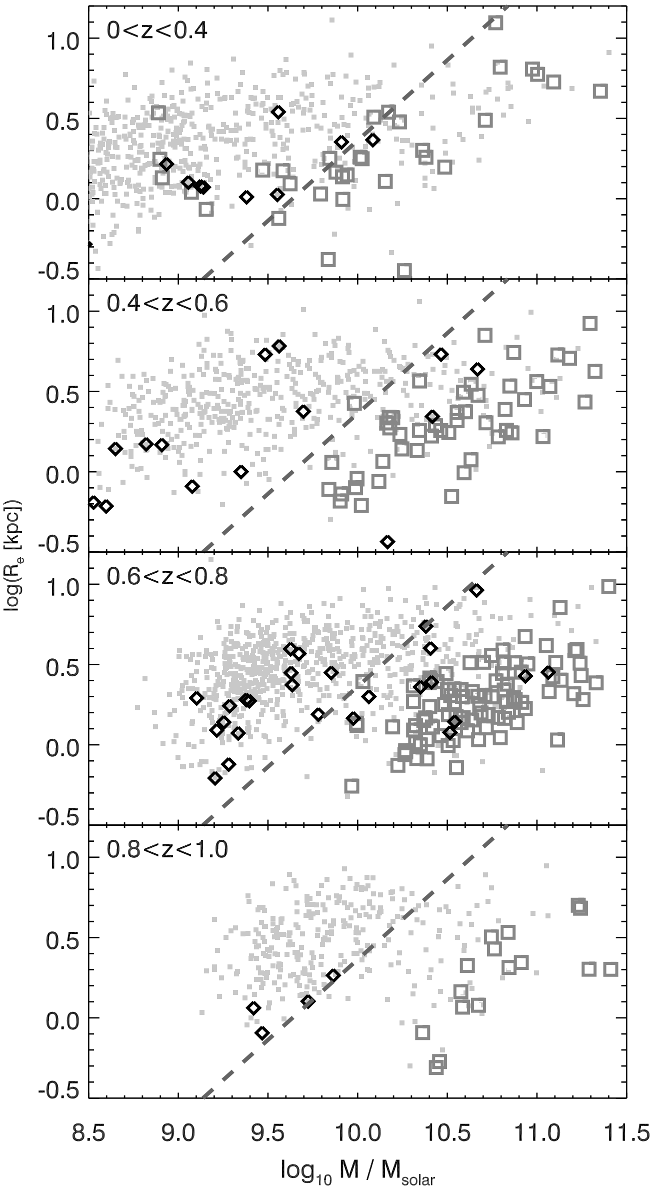


Figure 4.13 Mass-size diagrams for galaxies at different redshifts. The underlying grey dots show disk galaxies, both visually and automatically classified ($n < 2$), grey squares show red spheroidal from the visual classification and by color defined by figure 4.10. Diamonds show blue spheroidal galaxies. There are two classes of BSGs, one showing the same mass-size distribution as the RSGs, one showing sizes at a given mass much bigger, more in consistency with disk galaxies. The dashed line shows the mass-desnity cut of $3 \cdot 10^8 M_\odot kpc^{-2}$ (from Kauffmann *et al.* 2006) and is the same for all redshift bins. It is used to divide the BSGs into a low-density and a high-density sample.

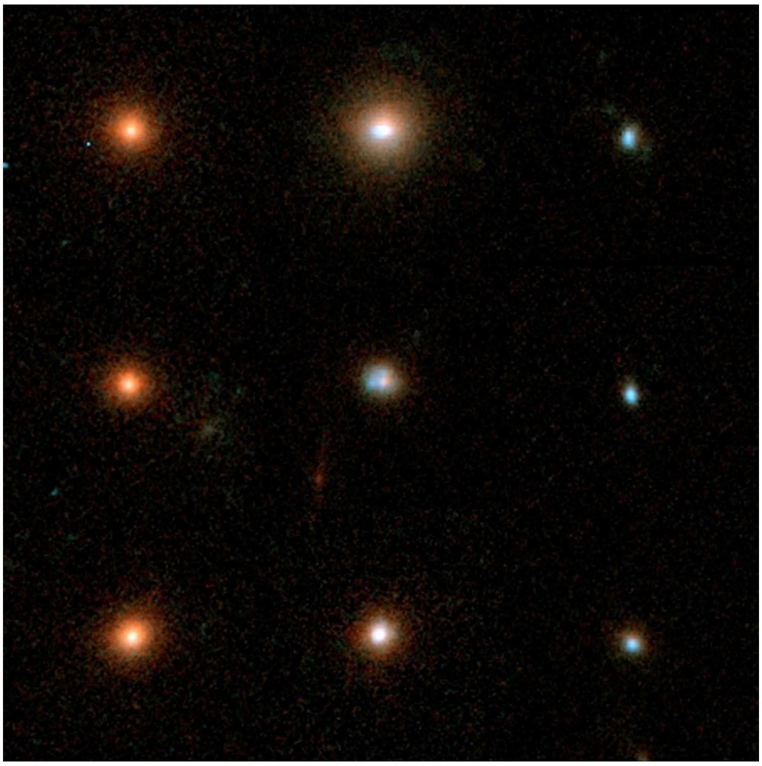


Figure 4.14 We show three examples for each sample of classified spheroidal galaxies, red spheroids, high-density blue spheroids and low-density blue spheroids (from left to right). Note again that the images do only roughly show cutouts of comparable physical size. The blue-red offset in the middle image is unphysical and comes from poor centering of one of the images used to create this color-image.

- Low-density BSGs are **much** bigger at a given mass than red spheroidal or high-density BSGG galaxies. Given their overlap with disks in the mass-size plane, we speculate that these galaxies my evolve into disk-dominated galaxies.

Figure 4.14 shows three examples of each of the three different populations, red spheroids, high- and low-density BSGs (from right to left). Note that these images show roughly the same physical size as they show galaxies from the $0.6 < z < 0.8$ redshift bin.

For these different classes of galaxies, we can plot the two quantities derived from figures 4.10 (the color distance to the red-sequence at a fixed magnitude in the color-magnitude diagram) and 4.12 (the magnitude distance to the red-sequence at a fixed size in the magnitude-size diagram). We show the result in figure 4.15, the X-axis showing the color difference, the Y-axis showing the magnitude difference. By definition, the red-sequence galaxies (open grey boxes) show no systematic offset in any direction, they clearly cluster around values of 0 along both axes. Low-density BSGs are shown as open diamonds, high-density BSGs are plotted as filled black diamonds. The overplotted model arrows connect 5 different stellar population models, to get a sense of the expected fading and reddening experienced by a galaxy after a starburst and truncation of the star formation rate. In all cases, the population started forming stars 5 Gyr before the epoch of observation. The star formation was truncated 0 (bluest), 0.5, 1.0, 2.5, and 4.5 (reddest) Gyrs before the epoch of observation. A starburst happens just before truncation, and the mass created in this burst brings up the models to have the same final stellar mass (i.e., these models have the same stellar mass, and the luminosity and color differences are the result of their different star formation histories). It is important to note that the expected fading and reddening of a truncation only, or a different mix of truncation and burst, are not significantly different from those shown in Fig. 4.15.

The arrows in each plot show how galaxies theoretically should fade and redden assuming passive evolution. They have been derived by stellar population models (see §4.2.1 for details) and show relative fading and reddening of stellar populations only, once the star formation has been shut off in the galaxy (arrows show the effect after 0.5, 1.0, 2.5 and 4.5 Gyrs).

One can see in figure 4.15 that only the high-mass-density galaxies are in place to evolve into typical red-sequence galaxies passively. Summarizing that they have both the same mass-size distribution as red-sequence galaxies and the right magnitudes and color, they are the ideal progenitors of the RSGs. Low-density galaxies are too faint to be able to evolve into red-sequence galaxies. Given the fact that they show a similar stellar mass-size distribution to typical disk galaxies, one could speculate that they are either old disk galaxies that form strong bulges through several minor mergers making them look rather spheroidal, or they are the remnants of mergers, but they are currently reforming a galactic disk by gas inflow, evolving them back into disk-dominated systems. They are clearly too extend at a given stellar mass (too low-mass at a given size) to be able to evolve into typical red-sequence galaxies.

4.5 Implications and Discussion

4.5.1 Comparison with Other Studies

Within our sample of spheroidal galaxies (high **and** low-density ellipticals, ignoring the lowest redshift bin $0 < z < 0.4$), we find that a total fraction of $\sim 22\%$ (47 out of 216 early-type galaxies in total, both spheroids and S0 galaxies) shows blue colors, so presumably star-formation happens or happened recently within these galaxies.

This is only roughly in agreement with other works which have found that approximately 30-50% of morphologically normal field E/S0 galaxies are quite blue because of recent star formation. For Example Abraham *et al.* (1999) find evidence for star formation which must have occurred within the past third of the galaxy ages at the epoch of observation for $\sim 40\%$ (4 out of their sample of 11 early-type field galaxies from the HDF with redshifts $0.4 < z < 1$), although they clearly suffer from number statistics due to their small field of view. Schade *et al.* (1999) also find that roughly 30% of spheroids (from a

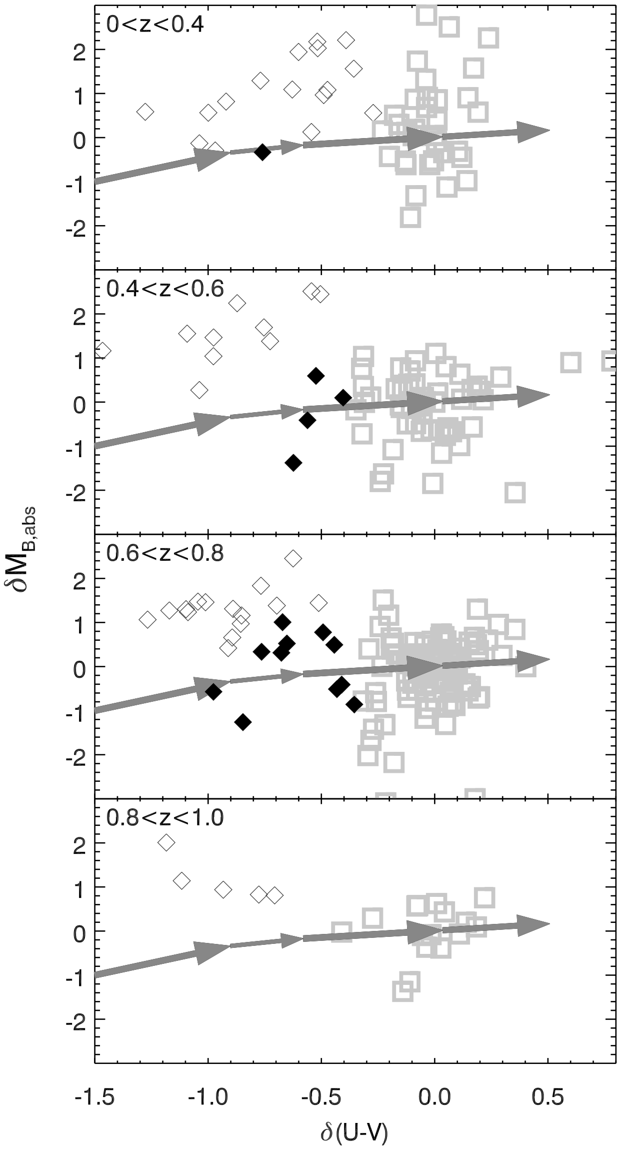


Figure 4.15 We plot the two quantities measured in figures 4.10 and 4.12, the distance to the fit of the red-sequence in the color-magnitude and the magnitude-size diagram, against each other for different types of galaxies and redshifts. Open diamonds represent low-density BSGs as defined by figure 4.13, filled diamond represent high-density BSGs. Open squares represent red-sequence galaxies. The arrows show the evolution according to stellar population synthesis models after 0.5, 1.0, 2.5, and 4.5 GYrs (for details about the models please see text). As one can see, high-density BSGs are well in place to evolve passively into red-sequence galaxies while the low-density BSGs are much too faint to be able to evolve into red-sequence galaxies that we typically see at all redshifts.

sample of 65 early-type galaxies) show star formation, identified by [OII] emission. Im *et al.* (2002) show a somewhat smaller number $\sim 16\%$, they used an automatically classified sample of 145 galaxies. Ferreras *et al.* (2005) find that $\sim 20\%$ of early-type galaxies in the GOODS survey show blue colors, well in agreement with our work.

It is important to notice that previous studies (except Ferreras *et al.* 2005) neglected low-density vs. high-density comparison. In fact, in our sample **most** of the BSGs, 48 out of 65, are low-density. Put differently, the red sequence has an observational mass limit of about $10^{10-10.3} M_{\odot}$ within our analysis, BSGs are seen at much lower masses, so the quantity $n_{bsg,9.5-12}/n_{rsg,10.3-12} + n_{bsg,9.5-12}$, where n_{bsg} is the number of blue spheroids, n_{rsg} is the number of red spheroids and numbers represent $\log{(mass[M_{\odot}])}$ and which is given in most of the older studies, compares different samples of galaxies. One would want to have a clean sample of galaxies by using $n_{bsg,10.3-12}/n_{rsg,10.3-12} + n_{bsg,10.3-12}$. This returns a fraction of only $\sim 9\%$ of high-mass spheroids being blue in our galaxy sample.

4.5.2 Are Mergers and Blue Spheroids Related?

Comparing the number of BSGs to the number of galaxy mergers is challenging as timescales have to be assumed for this kind of analysis. There are two key timescales: The timescale τ_{redden} over which remnants fade onto the red sequence, and the timescale τ_{merger} over which a merger remnant is recognizable as this. We assume a timescale of $\tau_{redden} \sim 1$ Gyr for the reddening and fading of starbursts triggered by galaxy mergers as given by the 'truncation' models presented in Fig. 4.15. That is, after a merger-induced truncation of star formation, we assume that the galaxy joins the red sequence in ~ 1 Gyr (a similar timescale was assumed by Im *et al.* 2002).

The second timescale, the timescale τ_{merger} over which merger remnants remain dynamically disturbed using 1 orbit depth *ACS* imaging, is considerably less certain, but has been argued to be ~ 0.4 Gyr (Conselice 2006). The 'BSG timescale' τ_{bsg} in which we would classify galaxies as BSG in our analysis is the difference between these two timescales — ~ 0.6 Gyr — that time where the galaxy is clearly blue, but shows weak/no signs of distortion from an early type galaxy light profile and appearance.

To be able to compare our work with the merger rate/close-pair fraction from Bell *et al.* (2006), we cut all spheroidal galaxies with masses $> 10^{10.5} M_{\odot}$ and $0.4 < z < 0.8$, roughly in agreement with their galaxy sample for their close-pair analysis. Given the fact that we assume 0.4 Gyrs as the dynamical timescale of a typical 1:1 merger, we should see ~ 1.5 times as many BSGs as we see galaxies in physical close pairs. Bell *et al.* (2006) find that roughly $5 \pm 1\%$ of all galaxies at these redshifts appear to be within close pairs. From the 101 chosen spheroids in our sample, 6 ($\sim 6\%$) show blue colors, which is roughly in agreement with the expectation. Given that the timescales are only rough approximations and given that the close-pair fraction estimated by Bell *et al.* (2006) includes some fraction of galaxies that are not physically correlated and some 'dry' (gasless) mergers (which would never be classified as blue spheroids), and thus overpredicts the number of galaxy mergers that will end up in blue spheroidal galaxies, there is no significant disagreement between these two numbers. While clearly this comparison is preliminary & could be improved and elaborated, the results are suggestive that the blue spheroid phase is an important of early-type galaxy evolution.

14 of the BSGs in our sample turn out to lie within the GOODS survey, which has 5 times deeper data. If one looks at these galaxies in this deeper data, 6-8 out of these 14 (43-57 %, 3-5 out of 11 in the low-density sample, 3 out of 5 in the high-density sample) galaxies indeed turn out to have very weak tidal features or other distortions, too weak to be seen in GEMS data. This shows that BSGs in 1-orbit *HST/ACS* data are indeed very good candidates as remnants of galaxy–galaxy mergers. The fraction of galaxies showing weak distortions is higher than expected from the timescales assumed, but given the fact that these timescales are only very rough approximations and that number statistics are very bad here, there is no significant disagreement.

4.5.3 The Cosmological Importance of Blue Spheroids

We have seen that the high-density blue spheroid population has the correct properties to be able to fade and redden on to present-day early-type galaxy scaling relations. Furthermore, the blue spheroid fraction for massive galaxies of $\sim 6\%$ compares favorably with the merger fraction, giving weight to the notion that blue spheroids are an intermediate step between galaxy merging and the creation of a red spheroid. In this section, we take the analysis one step further, and ask in very rough terms what the predicted growth in the spheroidal galaxy population is from integrating the blue spheroid creation rate as a function of redshift.

The blue spheroid fraction $n_{bsg}/n_{rsg} + n_{bsg}$ is assumed to be equal to 0.06 at $z = 0.6$. We assume an evolution of the blue spheroid fraction of similar form to those proposed for the merger rate $f_{bsg} \propto (1 + z)^{\alpha}$, where we take two limiting cases: $\alpha = 4$ following Bell *et al.* (2006) for the massive galaxy pair fraction evolution, and $\alpha = 2$ following Lotz *et al.* (2006a) and Bell *et al.* (2006) for the luminous galaxy merger fraction evolution. If $n_s = n_{bsg} + n_{rsg}$, where n_s is the number of spheroids, one can easily integrate the number of spheroids created per unit time:

$$n_s(z - \Delta z) = n_s(z)[1 + f_{bsg}(z)\, dt/dz\, \Delta z\, \tau_{bsg}^{-1}] \tag{4.5}$$

where Δz is the incremental change in redshift, dt/dz is the rate of change of cosmic epoch with redshift, and τ_{bsg} is the blue spheroid timescale, taken to be $\sim 0.6\,$Gyr. Integrating, we find that the number of spheroidal galaxies at $z = 0$ should be 1.4 (1.6) times the number at $z = 0.8$, assuming $\alpha = 4$ (2). Obviously, this is a highly uncertain exercise, bearing in mind the significant uncertainties in blue spheroid timescale and in the evolution of blue spheroid fraction. Nonetheless, it is clear that the measured massive galaxy blue spheroid fraction implies significant growth of the early-type galaxy population between $z = 0.8$ and $z = 0$.

4.6 Conclusions

Using stellar population models, we were able to estimate galaxy restframe images on a pixel-to-pixel basis from one-color (two-band) *HST/ACS* data from GEMS. These images, although clearly showing low-level artefacts because of the unavoidable effects of smoothing (required to measure accurate colors), are well-understood and were found to be reliable at the $\sim 10 - 20\%$ level. We intensively tested these images on multi-wavelength *HST* GOODS data.

From fitting these images, we were able to derive a sample of high-Sérsic-index galaxies, which were visually classified to remove contamination of disk galaxies with central light-cusps, that severely bias the sample of 'early-type' galaxies when automated classification is used. Visual classification returns a sample of 216 spheroidal galaxies out to $z = 1$, out of which 47 (22%) turn out to have blue colors. This is, to best of our knowledge, by far the biggest sample of blue spheroidal galaxies examined so far, enabling us to separate this sample of galaxies into high-mass-density and low-mass-density galaxies. We find fundamental differences between these two classes of galaxies.

- High-density blue spheroids have the same mass–size distribution as typical red sequence galaxies; that is, they have colors and magnitudes at a given size which allow passive fading and reddening onto the red sequence.
- The low-density blue spheroids are more in agreement with disk galaxies in their mass–size distribution. As we have shown, they are not able to passively evolve into early-type galaxies; in fact, most of them would have to brighten at a given size to fall on local early-type galaxy scaling relations. What these galaxies are — perhaps merger remnants reforming a galaxy disk through further gas accretion, or disk galaxies forming bulges through minor merging — we cannot tell from this analysis.

We measure a total BSG fraction of $\sim$ 22% at $z \sim 0.6$, comparing reasonably well with previously published determinations of the BSG fraction. Yet, importantly, we find that the vast majority of BSG galaxies are low-density. Restricting our attention to high density BSG galaxies alone, we find that some $\sim 6\%$ of spheroidal galaxies are blue. This fraction of $\sim 6\%$ at $z \sim 0.6$ is a much more physically-meaningful indicator of the growth rate of the early-type galaxy population through the addition of 'young' elliptical galaxies. We estimate a $\sim 0.6\,\mathrm{Gyr}$ BSG timescale; adopting such a timescale we find a reasonable degree of consistency between the massive galaxy merger fraction and the blue spheroid fraction. Such a consistency gives support to a picture of merger-driven growth of the early-type galaxy population at $z \lesssim 0.8$. Finally, we estimate the growth in the number of early-type galaxies since $z = 0.8$ from galaxies which pass through the blue spheroid stage: the number of spheroid-dominated galaxies at the present-day is predicted to be ~ 1.5 times higher than the number of such galaxies at $z = 0.8$. Thus, blue spheroids represent signposts of cosmologically-significant growth of the spheroidal galaxy population at intermediate redshifts.

Chapter 5
GEMS: The Size Evolution of Disk Galaxies[12]

Abstract

We combine HST imaging from the GEMS survey with photometric redshifts from COMBO-17 to explore the evolution of disk-dominated galaxies since $z \lesssim 1.1$. The sample is comprised of all GEMS galaxies with Sérsic indices $n < 2.5$, derived from fits to the galaxy images. We account fully for selection effects through careful analysis of image simulations; we are limited by the depth of the redshift and *HST* data to the study of galaxies with $M_V \lesssim -20$, or equivalently $\log(\mathcal{M}/\mathcal{M}_\odot) \gtrsim 10$. We find strong evolution in the magnitude–size scaling relation for galaxies with $M_V \lesssim -20$, corresponding to a brightening of $\sim$1 mag arcsec^{-2} in rest-frame V-band by $z \sim 1$. Yet, disks at a given absolute magnitude are bluer and have lower stellar mass-to-light ratios at $z \sim 1$ than at the present day. As a result, our findings indicate weak or no evolution in the relation between stellar mass and effective disk size for galaxies with $\log(\mathcal{M}/\mathcal{M}_\odot) \gtrsim 10$ over the same time interval. This is strongly inconsistent with the most naive theoretical expectation, in which disk size scales in proportion to the halo virial radius, which would predict that disks are a factor of two denser at fixed mass at $z \sim 1$. The lack of evolution in the stellar mass–size relation is consistent with an "inside-out" growth of galaxy disks on average (galaxies increasing in size as they grow more massive), although we cannot rule out more complex evolutionary scenarios.

5.1 Introduction

The last eight billion years have witnessed strong evolution of the disk galaxy population. Both 'archaeological' studies of local disk-dominated galaxies and 'look-back' studies of the evolution of disk galaxies suggest a steady build-up in their stellar masses since $z \sim 1$ (Rocha-Pinto *et al.* 2000; Flores *et al.* 1999; Bell *et al.* 2005; Hammer *et al.* 2005). Insights into how this growth occurs are accessible through the study of disk galaxy scaling relations, such as the luminosity–rotation velocity (Tully-Fisher) relation or the luminosity–size relation (e.g. Vogt *et al.* 1996; Lilly *et al.* 1998; Simard *et al.* 1999;

---

[12] This chapter has been adapted from Barden *et al.* (2005).

Böhm *et al.* 2004). Yet, owing to sample size limitations, selection effects, and differences in analysis techniques, these studies have come to widely divergent conclusions. In this chapter, we explore the evolution of the luminosity–size and stellar mass–size relations over the last 8 Gyr (since $z \sim 1$) using a sample of almost 5700 disk-dominated galaxies from the HST GEMS survey (Galaxy Evolution from Morphology and SEDs Rix *et al.* 2004).

In the Cold Dark Matter (CDM) picture of structure formation, collapsing dark matter perturbations acquire angular momentum through tidal torques and mergers (Peebles 1969; Maller *et al.* 2002; Vitvitska *et al.* 2002). Some fraction of this angular momentum is conserved, leading to the formation of cold, rotationally-supported disks. The typical magnitude of the specific angular momentum predicted in this framework leads to the formation of present day disks with approximately the correct distribution of radial sizes, if the specific angular momentum of the gas is similar to that of the dark matter and is mostly conserved during the formation process (Fall & Efstathiou 1980).

A difficulty is that this idealized picture does not correspond to the outcome when the process of galaxy formation is simulated in detail within the cosmological context of CDM. In hydrodynamical simulations, the gas tends to lose a large fraction of its initial angular momentum, resulting in disks that are too small compared to observed nearby galaxies (Navarro & White 1994; Sommer-Larsen *et al.* 1999; Navarro & Steinmetz 2000; D'Onghia & Burkert 2004). Furthermore, very few 'late-type' disks are formed in such simulations: galaxies tend to suffer mergers that thicken and destroy their disks (Steinmetz & Navarro 2002). It is not yet established whether this problem represents a fundamental difficulty with the 'standard' CDM paradigm (i.e., a result of excess small scale power), a reflection of our incomplete ability to understand and simulate the complexities of star formation and supernova feedback, or inadequacies in numerical resolution.

Many proposed solutions to this problem involve *delaying* gas collapse and disk formation to later times, either by adopting an alternate power spectrum with reduced small scale power (such as Warm Dark Matter), in which structure formation occurs later (e.g. Sommer-Larsen & Dolgov 2001), or by invoking some form of feedback that prevents the gas from cooling until relatively late times $z \sim 1$ (Weil *et al.* 1998; Thacker & Couchman 2001). While these solutions would be consistent with an important build-up in the disk galaxy population at late times, the late formation times implied by these models may be in conflict with the old ages of disk stars in the Milky Way and M31 (Rocha-Pinto *et al.* 2000; Ferguson & Johnson 2001). Additional constraints can be gleaned from so-called 'backwards evolution' models, in which the ages and metallicities of the stars in present-day disk galaxies are used to constrain the formation history of different components within our and other galaxies (Chiappini *et al.* 1997; Boissier & Prantzos 1999). Direct measurements of the mass–size scaling relations and radial size distributions of disk galaxies at earlier epochs will provide an important counterpoint to these arguments by providing direct constraints on the angular momentum content of stars at these earlier times.

A number of previous studies have used the *Hubble Space Telescope* (*HST*) to quantify the evolution of disk galaxies by measuring their absolute sizes and magnitudes as a function of redshift. Magnitude and size are strongly correlated; a line of constant surface brightness falls *almost* parallel to the distribution of observed galaxies, making the evolution of galaxy surface brightness a natural choice for parameterizing the evolution of galaxy sizes. However, the results of studies measuring average rest-frame surface brightnesses as a function of redshift have proven controversial, ranging from detecting no evolution to rather strong evolution in the range of 1-2 mag arcsec^{-2} brightening by redshift $z \sim 1$. For example, Lilly *et al.* (1998) found an average increase of the surface brightness of ~ 1 mag by redshift $z \sim 1$. This result is supported by observations of galaxies at high redshifts ($z \sim 2 - 3$), detected in very deep ground-based near-infrared images (Labbé *et al.* 2003). Trujillo *et al.* (2004) estimate that the average rest-frame surface brightness of these objects is more than 2-3 mag arcsec^{-2} brighter than in the local universe.

Simard *et al.* (1999) pointed out that selection effects play a crucial role in such analysis. After accounting for the different sources of incompleteness, Simard *et al.* (1999) and Ravindranath *et al.* (2004) argue that the luminosity–size relation of disk galaxies evolves by less than 0.4 mag arcsec^{-2} over the interval $0.25 < z < 1.25$. Yet, in order to reproduce the observations, both groups found it necessary to introduce a new population of high surface brightness galaxies in the highest redshift bin ($z \sim 1$). A different interpretation was suggested by Trujillo & Aguerri (2004), who find strong evolution of the average rest-frame V-band surface brightness of ~ 0.8 mag arcsec^{-2} at a redshift $z \sim 0.7$, also including a full treatment of completeness.

In this work, we present the results from a new sample of disk-dominated galaxies from the GEMS survey. Each of our galaxies has a spectrophotometrically-measured redshift, a spectral energy distribution (Wolf *et al.* 2004, SED), and a stellar mass estimate (Borch 2004) from COMBO-17 . We combine these SED constraints with light-profile shapes and sizes determined from deep high-resolution *HST Advanced Camera for Surveys* (*ACS*) images. We reassess the evolution of the magnitude–size and stellar mass–size relation as a function of redshift over the range $0.1 \lesssim z \lesssim 1.1$, taking particular care to model the impact of the selection function. We suggest a resolution to the conflicting previous results by presenting a coherent picture of strong surface brightness evolution with redshift without the need for a new population of high surface brightness galaxies.

The layout of this chapter is as follows. In § 5.2 we present the GEMS data set and describe the sample selection, the galaxy fitting techniques and the corrections we applied to the data. We explain in some more detail our modeling of the sample completeness in § 5.3. In § 5.4, we explore the evolution of the magnitude–size and stellar mass–size relations for disk-dominated galaxies. We show that there is a trend of increasing average surface brightness with redshift and that there is little evolution of the surface mass density. In § 5.5 we discuss our results in comparison with previous studies in the literature, and compare them with theoretical expectations. We summarize our results in § 5.6. Throughout this chapter we use the concordance cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$ (Spergel *et al.* 2003). Unless indicated otherwise we use Vega-normalized magnitudes.

5.2 Sample Definition

5.2.1 Imaging Data

GEMS, Galaxy Evolution from Morphologies and SEDs (Rix *et al.* 2004), has imaged an area of ~ 800 arcmin^2 centred on the Chandra Deep Field South (CDFS), using the *ACS* on-board *HST*. Of these 78 *ACS* tiles the central 15 were incorporated from the GOODS project (Giavalisco *et al.* 2004). With integration times of ~ 35 min in each of two filters (F606W and F850LP) the point source detection limits reached $m_{\rm AB}$ (F606W) $= 28.3$ (5σ) and $m_{\rm AB}$ (F850LP) $= 27.1$ (5σ), respectively. Details about the image mosaic and data reduction will be explained in a subsequent paper (Caldwell et al. 2007, in prep.).

5.2.2 COMBO-17 Data

The *HST* imaging data is complemented by low resolution spectrophotometric data from COMBO-17 (Wolf *et al.* 2004). COMBO-17 has provided precise redshift estimates ($\sigma_z/(1+z) \sim 0.02$) for approximately 9000 galaxies down to $m_R < 24$. Rest-frame absolute magnitudes and colors, accurate to ~0.1 mag, are also available for these galaxies. Furthermore, using a simple parameterized star formation history and the photometry in the 17 COMBO-17 bands, Borch (2004) computed stellar mass estimates for each galaxy in our sample, assuming a Kroupa *et al.* (1993) stellar initial mass function (IMF). These mass estimates are consistent with those derived using a one-color-based transformation from light to mass as described in Bell & de Jong (2001) and Bell *et al.* (2003b). While such estimates suffer from uncertainties in the IMF, ages, dust, and metallicity, it is encouraging to note that several studies (Bell *et al.* 2003b; Drory *et al.* 2004) find good agreement between masses based on broad-band colors and those from spectroscopic (e.g. Kauffmann *et al.* 2003a,b) and dynamical (Drory *et al.* 2004) techniques.

5.2.3 Source Detection

For source detection we use the SEXTRACTOR software (Bertin & Arnouts 1996) on the F850LP image. In contrast to the standard single-pass approach, we apply a two-step process, running SEXTRACTOR twice on each image to find an acceptable compromise between deblending and detection threshold (see Rix *et al.* 2004). Combining the source lists from each tile, taking care to remove duplicate objects that were detected in two neighbouring tiles, we end up with over 40,000 galaxies.

5.2.4 Galaxy Fitting and Disk Selection

For the purpose of this chapter we wish to isolate the subset of galaxies whose light is dominated by a disk component. We start by identifying all galaxies that can be reasonably well-fit by any single Sérsic profile (Sersic 1968) using the two-dimensional fitting code GALFIT (Peng *et al.* 2002). The Sérsic profile is a generalisation of a de Vaucouleurs profile with variable Sérsic index n:

$$\Sigma\left(R\right) = \Sigma_e \times \exp\left(-\kappa\left[\left(R/R_e\right)^{1/n} - 1\right]\right), \tag{5.1}$$

where R_e is the effective or half-light radius, Σ_e is the effective surface density, $\Sigma\left(R\right)$ is the surface density as a function of radius and $\kappa = \kappa\left(n\right)$ is a normalization constant. An exponential profile has $n = 1$ while a de Vaucouleurs profile has $n = 4$. GALFIT convolves Sérsic profile galaxy models with the point spread function of the *ACS* (Jahnke *et al.* 2004b, Jahnke et al., in preparation) and then determines the best fit by comparing the convolved models with the science data using a Levenberg-Marquardt algorithm to minimize the χ^2 of the fit. The best-fit model is given by 7 parameter values and their associated uncertainties, including the half-light radius, the Sérsic index and the total magnitude. Initial GALFIT starting guesses for the model parameters were obtained from the SEXTRACTOR source catalogues. Typically, neighbouring galaxies were excluded from each model fit using a mask, but in the case of closely neighbouring galaxies with overlapping isophotes the galaxies were fitted simultaneously. The sky level for each galaxy was carefully measured using flux growth curves, masking out detected neighbouring sources. Lacking an estimate for the Sérsic index from SEXTRACTOR, we started all fits with $n = 1.5$. In addition, all galaxies with $0.65 < z < 0.75$ were fitted with GIM2D (Simard *et al.* 2002). Estimates for magnitudes, sizes and Sérsic indices from the two codes agree very well (see Bell et al. 2004; see also § 3). Morphological quantities quoted in the present chapter were derived using GALFIT.

For this study, we estimate structural and morphological parameters from the z-band images (F850LP). In the optical (and the near-infrared), young stars make a progressively smaller contribution with increasing wavelength. Therefore, galaxy morphologies in F850LP are smoother than those in F606W, leading to a more robust detection and deblending of extended sources. The F850LP band corresponds to rest-frame R, V, and B-bands at $z \sim 0.4$, 0.7, and 1 respectively.

Selection of a galaxy sample for this kind of study is a multi-step process. First we merge the GEMS catalogue with the COMBO-17 redshift catalogue, then we select disk-dominated objects, and finally we remove sources with poor fits (see Fig. 5.1). We start by matching the GEMS sources to the COMBO-17 catalogue. To account for the relatively high source density in the GEMS images we pick the closest neighbour in the COMBO-17 catalogue within 0.5 arcsec as the corresponding match for a GEMS galaxy. Only at matching distances exceeding 1 arcsec does one start to include uncorrelated pairs. We are left with about 8000 matched sources in our sample.

We isolate disk-dominated galaxies for further study by cutting the sample based on the Sérsic profile fits. We adopt $n = 2.5$ as the dividing line between disk- ($n < 2.5$) and spheroid-dominated ($n > 2.5$) galaxies. This cut discriminated between visually-classified early- and late-type galaxies from GEMS with $0.65 < z < 0.75$ with 80% reliability and less than 25% contamination (Bell *et al.* 2004a). This cut is also consistent with the analysis conducted by the Sloan Digital Sky Survey (SDSS; see Shen *et al.*

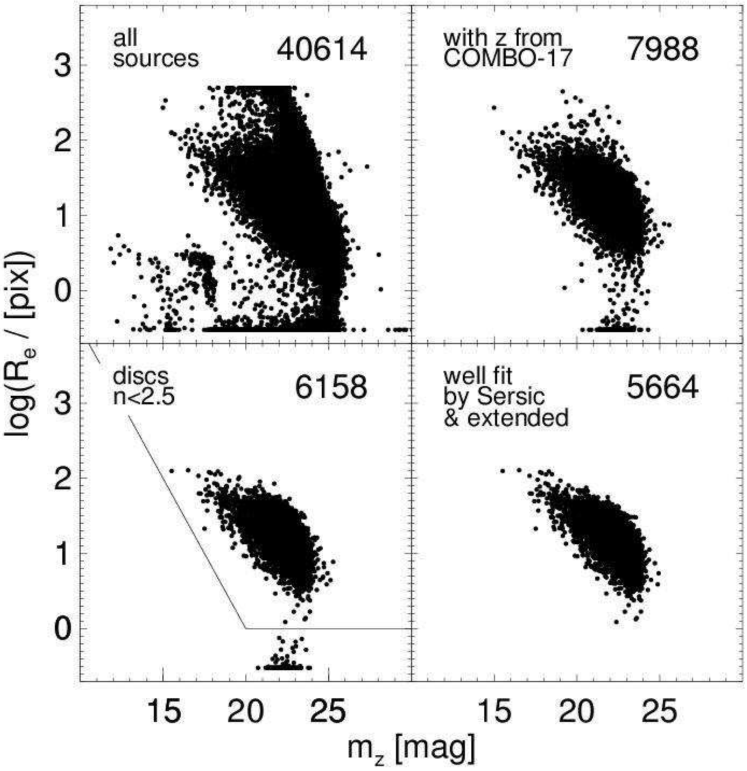


Figure 5.1 The disk sample definition in the apparent magnitude–size plane (apparent effective radius: $R_e^{\rm app}$, apparent z-band magnitude: m_z). Top left: all galaxies detected in the GEMS tiles. Top right: galaxies with COMBO-17 redshifts. Bottom left: disk galaxies with Sérsic index $n < 2.5$. Bottom right: disk galaxies with reliable GALFIT fits (see text for details on selection criteria). In each panel, we give the total number of galaxies in the upper right.

2003). Furthermore, Ravindranath *et al.* (2004) have redshifted a sample of local galaxies to show that the Sérsic index is still a useful indicator at redshifts $z \sim 0.5 - 1$. Selecting galaxies with $n < 2.5$ left us with ~ 6200 disk-dominated objects.

To ensure that the extracted galaxy profile parameters are reliable we remove objects from our source list that have relative formal errors in Sérsic index n and effective radius R_e of more than 25% ($\delta n/n > 0.25$, $\delta R_e/R_e > 0.25$)[13]. We also exclude objects that reach the boundary conditions for n ($0.2 < n < 8$) or R_e ($0.3 < R_e$ [pixel] < 500). Furthermore, we require that the GALFIT magnitudes coincide with the SEXTRACTOR magnitudes to within 0.6 mag ($|m_{\rm GALFIT} - m_{\rm SEx} + 0.166| < 0.6$). Finally, we remove compact sources with $\log R_e^{\rm app} <$ max $[8 - 0.4 \times m_z, 0]$ (indicated by the solid line in the bottom left panel in Fig. 5.1). While slightly more galaxies with low surface brightness were

---

[13] GALFIT formal errors underestimate the true uncertainties, as assessed using simulated galaxy images. The true uncertainties for the bulk of the sample are $\sim 35\%$ in R_e and ~ 0.2 mag in m_z.

removed by these additional cuts than high surface brightness galaxies, no pronounced bias was introduced. It is important to note that the simulated galaxy samples were also subjected to these same cuts for the construction of the completeness maps; thus, the completeness maps account fully for any biases introduced by these (necessary) extra sample cuts.

This sample selection should provide a fair representation of the disk-dominated galaxy population at all redshifts. The final catalogue contains 5664 disk galaxies with absolute rest-frame B- and V-band magnitudes, redshifts and stellar masses obtained from COMBO-17 and apparent half-light radii and Sérsic indices from GALFIT.

5.2.5 The Local Comparison Sample

In order to compare our measurements to a local reference point we have opted to use the NYU Value-Added Galaxy Catalog (VAGC; Blanton *et al.* 2004), which is based on the second data release (DR2) of the SDSS (Abazajian 2004). It contains Sérsic fits for 28089 galaxies in the redshift range $0.0033 < z < 0.05$. For this chapter, we use the VAGC elliptical aperture Sérsic fits for estimates of r_{AB}-band half-light radius, Sérsic index and -flux; coupled with extrapolated circular aperture $u_{AB}, g_{AB}, r_{AB}, i_{AB}, z_{AB}$ fluxes. The magnitudes were converted to absolute galactic foreground extinction-corrected magnitudes using the latest K-CORRECT routines, which were also used for the original data (Blanton *et al.* 2003a). We apply the following correction to convert the SDSS elliptical half-light r_{AB}-band sizes to rest-frame V-band (see § 5.2.6): $R_e\,(V) = R_e\,(r) \times 1.011$. The redshift of the individual SDSS sources does not impact significantly on this correction factor. To obtain a rest-frame B-band size for the SDSS galaxies we use: $R_e\,(B) = R_e\,(V) \times 1.017$.

We have chosen the VAGC rather than the fits to the magnitude–size and stellar mass–size planes by Shen *et al.* (2003) for various reasons. Using the VAGC we have full control over all estimated parameters including the photometric system, k-corrections, etc. Specificially, the fits by Shen *et al.* (2003) where performed on circularized size estimates while we use elliptical Sérsic measurements. The half-light sizes and absolute magnitudes by Shen *et al.* (2003) were provided only in SDSS filters, necessitating the use of color transformations and of additional luminosity function convolutions in order to obtain mean values for the same selection and photometric system as the GEMS data. Furthermore, the VAGC allows us to repeat the same analysis procedure that was also used for the GEMS data. Finally, the VAGC incorporates the latest version of the SDSS pipeline, leading to more robust Sérsic indices, fainter apparent limiting magnitudes and fewer problems with deblending large sources. Since the VAGC and the data used by Shen *et al.* (2003) have $\sim 20,000$ sources in common we could verify that the measured parameters broadly agree with each other.

The VAGC does not contain stellar masses. Therefore, we have used the prescription given in Bell *et al.* (2003b) to convert a $(g-r)_{AB}$ color into a SDSS r_{AB}-band stellar mass-to-light ratio:

$$\log\left(\mathcal{M}/L_r\right) = -0.306 + 1.097 \times (g-r)_{AB} - 0.15. \tag{5.2}$$

We have applied a correction of -0.15 to convert to a Kroupa IMF, in accord with our GEMS stellar masses. The stellar mass was then obtained from the following relation:

$$\log(\mathcal{M}) = \log(\mathcal{M}/L_r) - 0.4 \times (M_{r,S} - r_\odot), \qquad (5.3)$$

with the absolute rest-frame Sérsic magnitude $M_{r,S} = r_S - 5\log(D_L) - 25$, the apparent rest-frame Sérsic magnitude r_S, the luminosity distance D_L and the absolute magnitude of the sun $r_\odot = 4.67$ in SDSS r_{AB}. Calculating a stellar mass in the same fashion for the lowest redshift GEMS galaxies and comparing this estimate with the SED-based masses (Borch 2004) reveals no apparent systematic offsets.

5.2.6 Rest-Frame V-band Sizes

Galaxies are known to exhibit radial color gradients. As a result of this, galaxy sizes vary as a function of wavelength and the measured physical size evolution of the galaxy population could be skewed by the effects of band shifting with redshift. Therefore, we have not simply converted our apparent half-light sizes R_e^{app} measured in the F850LP filter to a physical value, but instead have applied a color gradient correction to each individual galaxy according to its redshift to correct the size to the rest-frame V-band. For a sample of local galaxies, de Jong (1996) presents the relative disk scale lengths, which for a pure disk corresponds to $R_e = 1.678 \times R_d$, in the B-, V-, R-, I-, H- and K-bands. Figure 5.2 illustrates this ratio of the disk scale lengths in one band to the size measured in the V-band, as a function of the corresponding wavelength. A linear fit with the intercept fixed to 1 at the V-band results in a slope of $a_R = -0.184$, corresponding to correction factors varying by only $\pm 3\%$ over the whole redshift range. All future references to effective radii R_e are to sizes corrected to the rest-frame V-band.

In order to obtain rest-frame sizes for the SDSS data we have calculated the ratio of the circularized half-light sizes in the five SDSS bands, divided by the size in the SDSS g_{AB}-band. We overplot the resulting values in Fig. 5.2, minimizing in a simultaneous fit the offset between the SDSS points and the other V-band normalized measurements. The agreement between the various measurements is striking. This supports the validity of the *average* correction to obtain rest-frame sizes, bearing in mind the 20% galaxy-by-galaxy scatter, and that this method, strictly speaking, applies only to nearby galaxies.

Given the possible rapid evolution of galaxy disks in the last 8 billion years, it is not inconceivable that the 'average' disk color gradient has evolved considerably since $z \sim 1$. In a subsequent chapter (compare § 4) we will reconstruct the rest-frame B-band for individual galaxies and estimate sizes directly from this image to account for this effect. As an interim solution, we have tested the applicability of the local average relation on distant galaxies in GEMS. We have fit all GEMS galaxies in the F606W band using exactly the same approach used to fit in F850LP. Owing to significant differences in the depth of the F606W and F850LP data, and F606W's extra sensitivity to ongoing star formation, we consider the F606W fits at this stage to be preliminary[14]. From these fits we selected those

---

[14] While many galaxy fits were reasonably successful, a non-negligible fraction of the fits are substantially

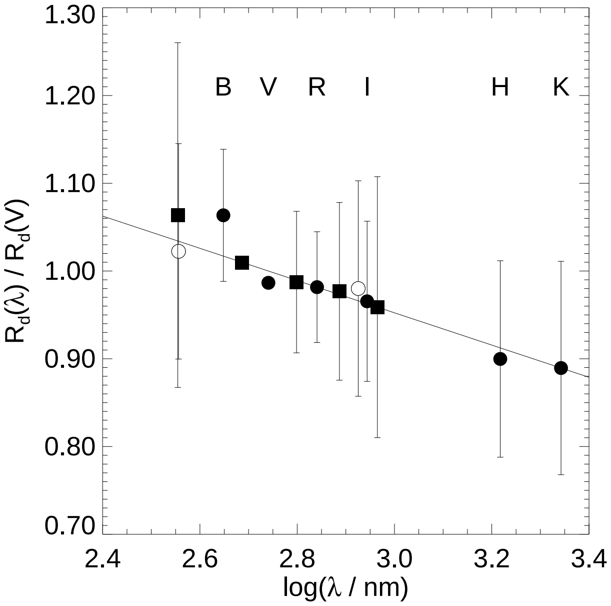


Figure 5.2 The average ratio of the disk scale length $R_d\,(\lambda)$ measured in various bands (B, V, R, I, H, K) over $R_d\,(V)$ measured in the V-band for the de Jong (1996) data as a function of corresponding wavelength λ (solid dots). The solid line marks a linear fit f, which is constrained to $f\,(2.74) = 1$ at the V-band (not strictly requiring coincidence with the data point at the V-band). Over the redshift range sampled by the GEMS data the size corrections as inferred from this plot are of order $\pm 3\%$. Errors indicate the dispersion of the distribution of $R_d\,(\lambda)\,/R_d\,(V)$; they do not represent errors of the mean values plotted here. Solid boxes mark data points from the SDSS DR2 data set (see § 5.2.5 for details). Since there is no direct measurement in the V-band available for these objects, the data points are simultaneously fit minimizing the total offset between SDSS and de Jong values using the g_{AB}-band as the reference filter. Open symbols represent measurements from the GEMS survey where in a certain redshift bin one of the filters matched the rest-frame V-band.

sources for which one of the bands corresponds to the rest-frame V-band and measure the size ratio at $z \sim 0.08$ (F606W $\sim V_{\rm rest}$) and at $z \sim 0.64$ (F850LP $\sim V_{\rm rest}$). The average values from these measurements are overplotted in Fig. 5.2. They confirm the trend seen in the de Jong (1996) and SDSS data, supporting the validity of the correction we have applied to the data[15].

---

in error. Thus, while on average, the F606W fits are reliable, it is impossible at this stage to use a weighted sum of the F606W and F850LP fits to directly estimate the rest-frame B- or V-band sizes on a galaxy-by-galaxy basis.

[15] It is worth recalling that the corrections implied by this relation are rather small, $\sim 3\%$ for the average GEMS galaxy. Furthermore, the evolution of average rest-frame V-band surface brightness is dominated by galaxies with $z \gtrsim 0.6$, where the F850LP samples rest-frame V-band almost directly, and by the SDSS

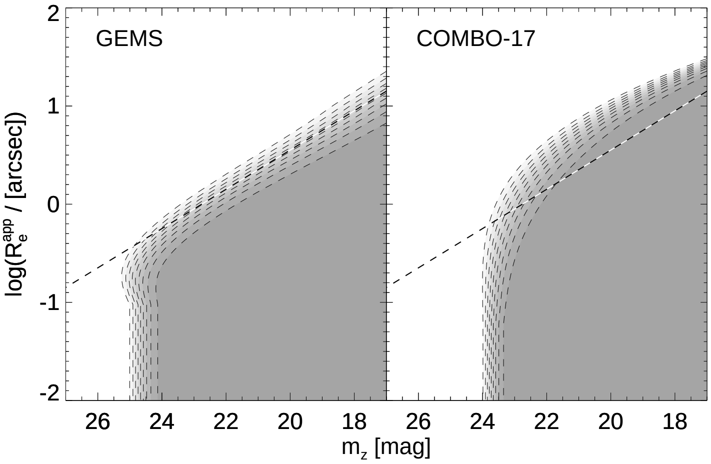


Figure 5.3 The GEMS (left panel) and the COMBO-17 (right panel) detection probabilites as a function of apparent z-band magnitude m_z and apparent half-light radius $R_e^{\rm app}$. The contours indicate different detection probability levels (from light to dark shades of grey: 10%, 20%, 30%,... 90%). For comparison, in both plots a line of constant apparent surface brightness $\mu_z^{\rm app} = 24$ mag arcsec^{-2} is shown.

5.2.7 Completeness

In order to estimate the limitations of the GEMS survey we have performed extensive simulations of artificial disk galaxy light profiles (see § 3). By inserting a number of such artificial disk images with purely exponential profiles (Sérsic index $n = 1$), and subsequently re-running the source detection and fitting process (including removal of bad fits according to § 5.2.4), we calculate our success rate: the completeness as a function of apparent effective radius $R_e^{\rm app}$ and apparent magnitude m_z. It turns out that the contours of constant detection probability in the $R_e^{\rm app}$-m_z-plane (see Fig. 5.3) lie along lines of constant apparent surface brightness:

$$\mu_z^{\rm app} = m_z + 2.5 \log(2\pi q) + 5 \log(R_e^{\rm app} / \,[\rm arcsec]), \tag{5.4}$$

in the limit of bright magnitudes (q is the axis ratio). At the faint magnitude limit, however, the lines of constant detection probability are at constant magnitude. The precise location of such a line depends also on the axis ratio of the objects. In the absence of dust, an object with high inclination has a higher detection probability than a source of the same apparent magnitude but viewed face-on.

We model the detection probability as a function of the apparent magnitude. A double exponential model provides a good fit to the data (for a detailed description see appendix 5.7.1). Both the shape and the characteristic magnitude limit at which a specific detection probability is reached depend on the apparent size and the axis ratio.

---

data at $z \leq 0.05$; thus, further reducing our sensitivity to any errors in the size correction.

Our final sample contains only the objects with redshift estimates from COMBO-17 and therefore we must also account for the COMBO-17 completeness limit. Wolf *et al.* (2003a) have calculated the completeness of COMBO-17 as a function of apparent R-band aperture magnitude $m_{\mathrm{R,aper}}$, redshift and $U - V$ color. In order to show the COMBO-17 completeness contours on Figs. 5.3, 5.4, 5.6 and 5.7, we statistically transform the COMBO-17 completeness map into the $m_z - R_e$ plane (Appendix B). We adopt this analytic approximation to the COMBO-17 completeness in the rest of this chapter, but note that the use of either the true COMBO-17 completeness map or the analytical mapping of the completeness maps onto the $m_z - R_e$ plane in the analysis that follows does not affect our conclusions.

We combine the GEMS detection probability and the COMBO-17 completeness by multiplying the two values for each individual object:

$$p = p_{\mathrm{GEMS}} \times p_{\mathrm{COMBO\text{-}17}}. \tag{5.5}$$

We can now estimate the combined detection probability p of individual galaxies. Since later on we weight galaxies by the inverse of the detection probability we have taken special care when using very low detection probability values. In order to avoid attributing large weights to any given galaxy (which would then dominate the whole sample), we remove any object with $p < 5\%$ from the sample (a total of 14 sources). For the main analysis presented here, we only include objects with a detection probability $p > 50\%$. In appendix 5.7.3 we discuss in more detail how the detection probability will impact on the evaluation of the data especially in the magnitude–size plane, which is also the reason for not removing galaxies with $0.05 < p < 0.5$ from the sample altogether. In Fig. 5.4 we illustrate the resulting detection probability function in the R_e^{app}-m_z-plane.

The completeness p_{SDSS} of the SDSS data is parameterized as a function of surface brightness $\mu_{50,r}$ and position on the sky RA (α) and dec (δ) (Blanton *et al.* 2004): $p_{\mathrm{SDSS}}\left(\mu_{50,r}, \alpha, \delta\right) = f_{ti}\left(\mu_{50,r}\right) \times f_{sp}\left(\mu_{50,r}\right) \times f_{ph}\left(\mu_{50,r}\right) \times f_{got}\left(\alpha, \delta\right)$, where f_{ti} is the "tiling" fraction, f_{sp} is the spectroscopic completeness, f_{ph} is the photometric completeness and f_{got} is the fraction of main targets for which a classification was obtained in this object's sector, as described in more detail in Blanton *et al.* (2004). The resulting completeness as a function of surface brightness we present in Fig. 5.5 for the case $f_{got}\left(\alpha, \delta\right) = 1$. Note that the rapid drop of the completeness at high surface brightnesses directly results from the improper deblending of the largest nearby galaxies. We have approximated the data points given in Blanton *et al.* (2004) with the following analytical formula:

$$\begin{aligned} p_{\mathrm{SDSS}} &= 0.99 \times \exp\left(-\exp\left(\frac{\mu_{50,r} - 23.6}{0.6}\right)\right) \\ &\quad \times \left(1 - \exp\left(-\exp\left(\frac{\mu_{50,r} - 18.1}{0.7}\right)\right)\right) \end{aligned} \tag{5.6}$$

In the subsequent analysis we only consider objects with a completeness $p_{\mathrm{SDSS}} \geq 0.5$, in order to match the selection of the GEMS galaxies.

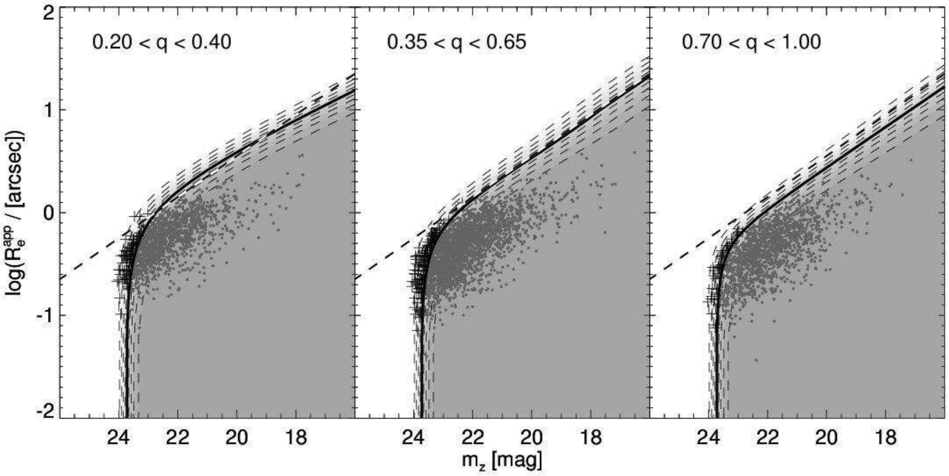


Figure 5.4 The combined GEMS / COMBO-17 completeness (dashed contours indicating 10%, 20%, 30%,... 90%) in the R_e^{app}-m_z-plane. The solid contour shows the 50% completeness limit. Dots and pluses indicate disk galaxies with a combined GEMS detection / COMBO-17 redshift estimation probability $p > 50\%$ and $p < 50\%$, respectively. The panels show the completeness contours and data points for three different axis ratio ranges, indicated in the top of each plot. The diagonal line indicates a constant apparent surface brightness $\mu_z^{\text{app}} = 24$ mag arcsec^{-2}.

5.3 Analysis of Completeness and Selection Effects

In the following sections we evaluate the magnitude–size and stellar mass–size relations as a function of redshift. We have subdivided our sample of disk galaxies into five redshift bins, each of which spans a range of 0.2 in redshift, centred on $z = 0.2, 0.4, 0.6, 0.8, 1.0$, plus an additional redshift bin at $z \sim 0.0$ for the SDSS data.

In Fig. 5.4 we show the combined completeness map with observed disks with $0.35 < q < 0.65$ overplotted. The galaxies in the sample form a relatively tight relation in the apparent magnitude–size plane. Inspecting the slope α of this relation one realizes that it is close to, but not exactly equal to a line of constant surface brightness. A linear fit provides a slope $\alpha \sim -0.15$. In physical quantities this slope closely matches that of a line of constant volume density, i.e. a law such that the ratio of flux and the cube of the radius is constant ($\alpha = -0.1\overline{3}$), rather than a line of constant surface density ($\alpha = -0.2$). The fact that the slope does not match that of a constant surface brightness implies that measuring average surface brightnesses depends to some extent on the range in magnitudes over which the average is calculated. Thus, in order to quantify the evolution of the surface brightness one has to make sure that the same range of absolute magnitudes is observed at all redshifts.

The sample becomes approximately magnitude-limited at $m_z \sim 23.5$. This limit

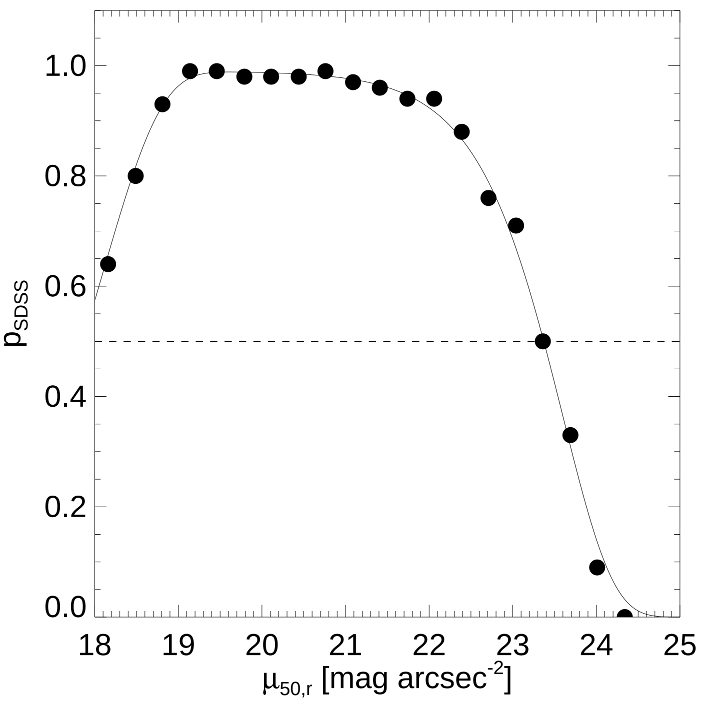


Figure 5.5 The SDSS completness as a function of apparent surface brightness $\mu_{50,r}$. The dots indicate the actual surface brightness values as given in Blanton *et al.* (2004). The solid line represents our analytical fit to that relation.

is imposed by the COMBO-17 redshifts; fainter galaxies cannot be assigned reliable redshifts. Furthermore, we find no galaxies at brighter magnitudes $m_z < 23$ with detection probabilities less than 50%. Since we show in Fig. 5.4 galaxies of all redshifts, this implies that *the* GEMS *data are not limited in surface brightness at any redshift*, even at the highest bin. Therefore, our subsequent analysis is not affected by a possible completeness-induced truncation of the surface brightness distribution of the galaxy population at any redshift. We conclude that the combined GEMS + COMBO-17 sample is essentially magnitude-limited only, with surface brightness playing a minor role. This conclusion is robust to the detailed choice of axis ratios.

We have translated these completeness contours to the absolute magnitude–size plane in Fig. 5.6. To estimate the absolute magnitude, we fit a third-order polynomial to the "average apparent z minus rest-frame apparent V color" $\langle m_z - m_V^{\rm rest}\rangle$ of our sample as a function of redshift. Obviously a redshift dependence cannot fully model this color, leading to a small additional scatter of the data relative to the transformed completeness map[16].

In Fig. 5.6 we also overplot the SDSS completeness. At the low surface brightness

---

[16] These transformed completeness maps are not used in the science analysis; rather, they are included in the figures for presentational purposes alone.

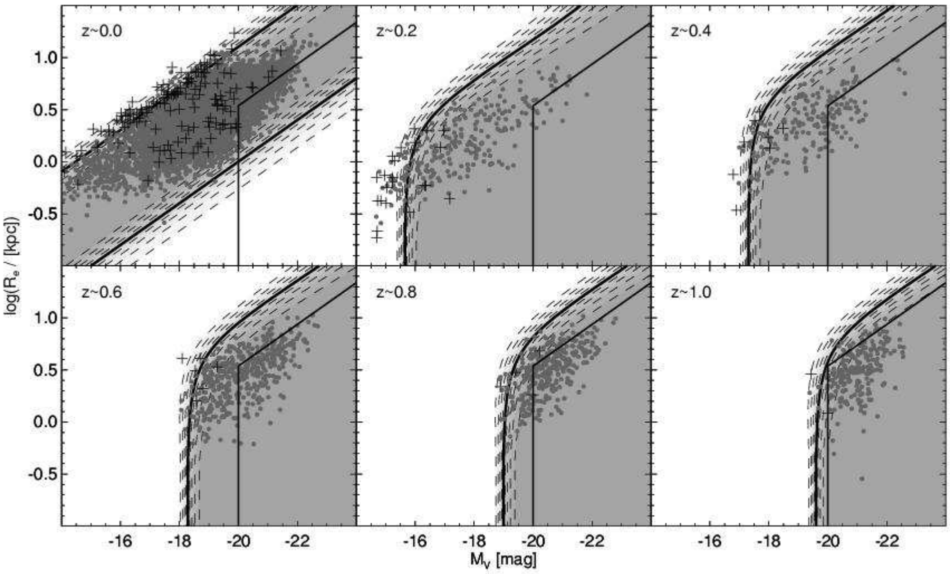


Figure 5.6 The completeness as a function of redshift in the absolute magnitude–size plane. The same completeness contours and objects from the middle panel of Fig. 5.4 are plotted. The $z \sim 0$ bin shows the SDSS data. The contours were computed for the central redshift of each bin as indicated in the top left of each panel. Galaxies with low detection probability are found along the absolute magnitude limit, but not at bright magnitudes and faint surface brightnesses. The box in each panel encloses a selection with $\mu_V^{\rm lim} = 22$ mag arcsec^{-2} and $M_V^{\rm lim} = -20$. This plot shows that the GEMS data ($0.2 < z < 1.0$) are not limited in surface brightness. Galaxies are observed in regions where the completeness contours indicate detection probabilities $p < 5\%$. This is a result of the finite width of the redshift bins (see § 5.4.1 and Fig. 5.7 for further explanations).

edge a fairly large number of SDSS objects are found with very low completeness values; the VAGC does not sample the full distribution of surface brightnesses. We adopt an absolute magnitude cut of $M_V < -20$ in this chapter: brighter than this limit the size distribution is sufficiently narrow that the full range of surface brightnesses is well-sampled. In order to estimate where the apparent magnitude limit $m_r^{\rm lim} = 17.77$ starts to affect the galaxy distribution we convert $m_r^{\rm lim}$ into an absolute magnitude $M_V^{\rm lim} = -18.8$ for the highest redshift in the VAGC using a color transformation for a typical Sbc (Fukugita *et al.* 1995). Again, this limit is fainter than our adopted absolute magnitude cut.

Inspection of Fig. 5.6 shows that the sample reaches $M_V < -20$ in the highest redshift bin; therefore in what follows we restrict our analysis to this absolute magnitude range at all redshifts. This selection leaves 3584, 76, 176, 704, 671 and 559 disk galaxies in the respective redshift bins $z = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$; a total of $3584 + 2186$ galaxies. This magnitude cut implies that our results are applicable only over this brightness range.

We have explored in detail the influence on the average surface brightnesses and surface densities of varying the $p > 50\%$ criterion, the surface brightness range over which one averages, and the absolute magnitude range considered. The influence of the p cut is negligible; the surface brightness and magnitude ranges do affect the average surface brightnesses/densities, and great care must be taken to choose appropriate integration ranges. These issues are discussed where relevant in §5.5, and in great detail in Appendix 5.7.3.

5.4 Analysis of the magnitude–size and Stellar mass–size Relation

For our subsequent analysis we define the absolute rest-frame effective surface brightness in the V-band as:

$$\mu_V = M_V + 5 \log R_e + 2.5 \log q + 38.568, \tag{5.7}$$

with the absolute rest-frame magnitude in the V-band from COMBO-17 and the half-light radius R_e in kpc. The constant 38.568 results from using sizes in kpc and luminosity distances in Mpc. Note that this formula is correct even for a general Sérsic profile. In the analysis of the evolution of μ_V we will only address the bright galaxy population with $M_V < M_V^{\mathrm{lim}} = -20$. Moreover, we define the "equivalent" absolute rest-frame surface mass density

$$\log \Sigma_{\mathcal{M}} = \log \mathcal{M} - 2 \log R_e - \log (2\pi q), \tag{5.8}$$

where the SED-estimated stellar galaxy mass $\mathcal{M}$ is given in $\mathcal{M}_\odot$. In the case of $\log \Sigma_{\mathcal{M}}$ we restrict the sample to galaxies with $\log (\mathcal{M}/\mathcal{M}_\odot) > \log \mathcal{M}^{\mathrm{lim}} = 10$. We calculate average values of the surface brightness $\langle \mu_V (z) \rangle$ and the surface mass density $\langle \log \Sigma_{\mathcal{M}} (z) \rangle$, correcting for incompleteness by weighting indiviual galaxies by the inverse of their detection probability as a function of redshift. We obtain errors on the estimated mean values by performing an extensive Monte-Carlo analysis (see appendix 5.7.3).

5.4.1 The Magnitude–Size Relation

In Fig. 5.7 we present the magnitude–size relation for disk galaxies in six redshift bins extending to $z \sim 1.1$. We stress that the completeness contours shown in the figure are only indicative as they were calculated for a fixed axis ratio $q = 0.5$ and the central redshift of the corresponding bin (see also Fig. 5.3). Therefore, especially in the $z = 0.2$ redshift bin, we see many galaxies "spilling over" into the incompleteness regions, which is a result of the non-negligible range of M_V cutoffs over redshifts $0.1 < z < 0.3$. To illustrate this effect we overplot vertical lines corresponding to an apparent magnitude $m_z = 24$ at the centre, low and high end of each redshift bin (for $z \geq 0.2$). With increasing redshift (co-moving volume) the spread of the completeness becomes smaller. The detection probabilities for individual galaxies, however, were calculated according to their exact magnitude, size and axis ratio and not relative to the plotted completeness contours. In the case of the $z \sim 0.0$ redshift bin we only indicate the brightness level, below which the highest redshift galaxies are not fully sampled.

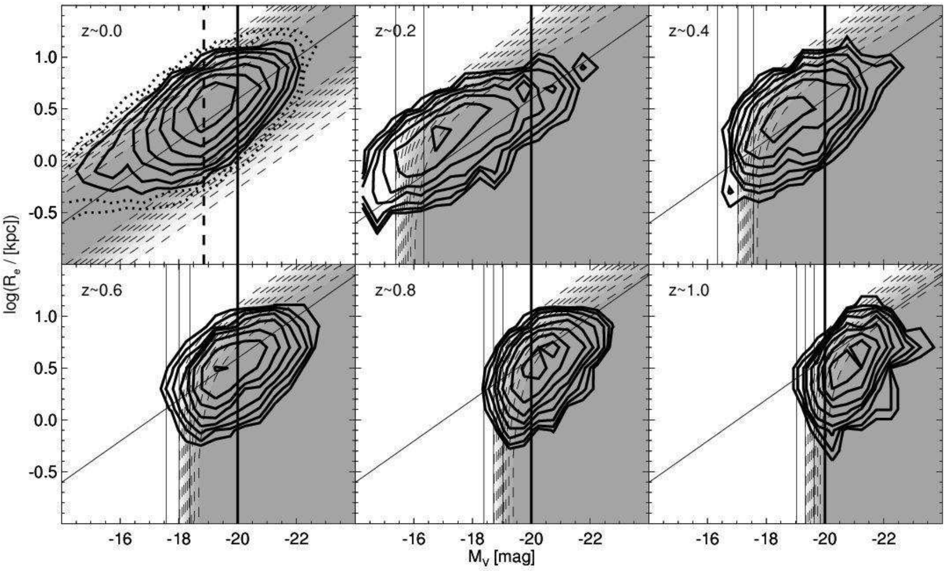


Figure 5.7 The magnitude–size relation for six different redshift bins. The solid contours indicate 11 levels of fractional number of objects per unit area. To obtain the contours the density field was normalised by the total number of objects being plotted in each redshift bin and the following scaling was used: $(\exp(x/2) - 1) \times 0.0005$, with $x = [1, 2, ..., 11]$. The dotted contours are only shown for the lowest redshift bin ($z \sim 0$; SDSS data) for being too noisy in the other bins. The grey-scale contours mark the same completeness levels as shown in Fig. 5.6. The thin vertical lines in each panel symbolise an apparent magnitude $m_z = 24$ converted to the rest-frame at three different redshifts corresponding to the center, low and high end of each redshift bin. This illustrates why apparently so many galaxies especially in the low redshift bin were detected at completeness values close to zero. The diagonal line in each panel corresponds to the average surface brightness $\mu_V = 20.84$ mag arcsec^{-2} in the lowest redshift bin. The thick vertical line represents the lower magnitude limit ($M_V = -20$), below which we exclude galaxies from the analysis. The thick dashed line shows the magnitude limit of the highest redshift SDSS data; at fainter magnitudes the VAGC data become incomplete.

As the completeness function limits us to detecting only the bright galaxies at high redshift, we limit our analysis to galaxies with $M_V < -20$. Recall also that we have demonstrated in §5.3 and Appendix 5.7.3 that we are not limited in absolute surface brightness even at the highest redshifts. Therefore, to evaluate the evolution of disk galaxies in the magnitude–size plane we have calculated the average rest-frame absolute surface brightness $\langle \mu_V(z) \rangle$ as a function of redshift including weighting of individual galaxies according to their detection probabilities. In Fig. 5.8 we show the weighted histograms of $\mu_V(z)$ for each redshift bin. Indicated in each panel (at each redshift bin z_0) are the estimated mean surface brightnesses $\langle \mu_V(z_0) \rangle$ together with the mean values of the preceeding redshift bins $\langle \mu_V(z < z_0) \rangle$ for comparison. This plot demonstrates clearly that there is a

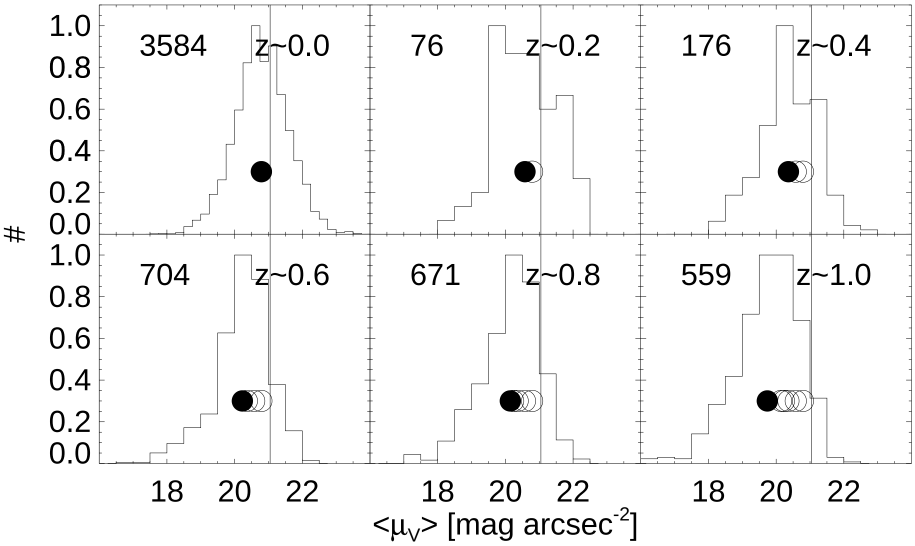


Figure 5.8 Histograms of the absolute rest-frame surface brightness as a function of redshift ($z \sim 0$ from SDSS; $0.2 < z < 1.0$ from GEMS). The histograms include weighting of individual objects according to their completeness (y-axis peak normalized). Only objects with a completeness exceeding 50% and with $M_V < -20$ were included in order to minimize the impact of selection effects. The black circle marks the estimated mean value $\langle \mu_V(z) \rangle$ for each redshift bin. In each redshift bin with $z > 0.0$ we overplotted as open circles values of $\langle \mu_V(z) \rangle$ of all lower redshift bins to visualize the evolution in surface brightness. The vertical line in each panel indicates the Freeman (1970) surface brightness converted to the V-band (conversion given in Fukugita *et al.* 1995). The numbers in the top left of each panel shows the number of sources used in each panel (not including weighting).

significant trend of increasing surface brightness with increasing redshift.

We demonstrate how the mean surface brightness of the disk galaxy population evolves by plotting $\langle \mu_V(z) \rangle$ as a function of redshift in Fig. 5.9. Fitting a linear function to the data we find an intercept and slope of $\langle \mu_V(z = 0) \rangle = 20.84 \pm 0.03$ mag arcsec^{-2} and d$\langle \mu_V(z) \rangle$/d$z = -0.99 \pm 0.06$, respectively, thus an evolution of ~1 mag to $z = 1$.

5.4.2 The Stellar Mass–Size Relation

These stellar mass estimates allow us to investigate the evolution of the analogous quantity to the magnitude–size relation: the stellar mass–size relation. Working in terms of stellar mass is useful not only because it is one step closer to the quantities actually predicted by theory, but also because it removes the evolution that is simply due to the aging of the stellar populations. We present the stellar mass–size relation in Fig. 5.10. Again, iso-density contours show the distribution of galaxies in the R_e-$\mathcal{M}$ rest-frame plane. We

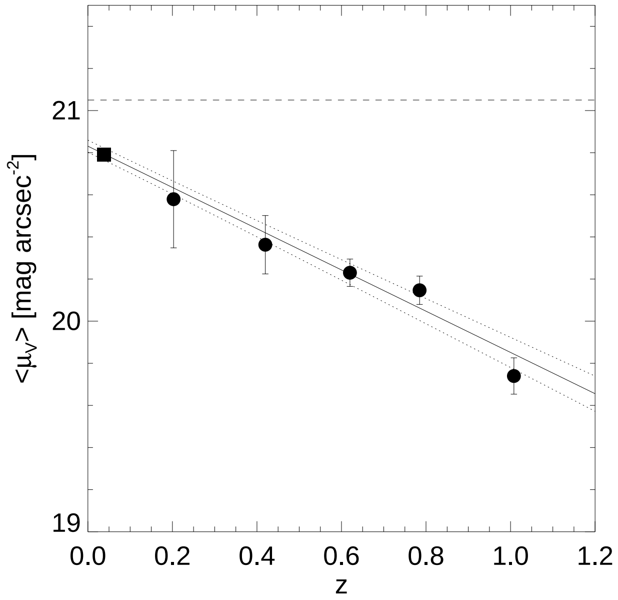


Figure 5.9 Evolution of the average absolute surface brightness $\langle \mu_V\,(z) \rangle$. Solid dots show the GEMS data for the individual redshift bins; the box symbol at $z \sim 0.05$ indicates the SDSS data point. The error bars mark 2σ statistical errors. The horizontal line at $\mu_V = 21.05$ represents the Freeman (1970) surface brightness converted to the V-band. The solid and dotted lines mark a linear fit to all data points plus the 1σ confidence limits, respectively.

use the same method as before to correct the size estimates to the rest frame V-band (it is important to note that ideally we would prefer to study stellar mass vs *stellar mass weighted size*, but we do not attempt this further correction here). As in the case of the average surface brightness $\langle \mu\,(z) \rangle$ we estimated the average stellar surface mass density $\langle \log \Sigma_{\mathcal{M}}\,(z) \rangle$, as defined in eq. 5.8, for each redshift bin.

We found that, as in the case with the surface brightness, the distribution of galaxies in the stellar mass–size plane does not fall exactly along a line of constant stellar surface mass density, but is of somewhat shallower slope. However, here the effect is much less pronounced (also due to the width of the distribution) and therefore, the precise cut-off in stellar mass, which is the equivalent of absolute magnitude, is not as important.

Plotting mass as a function of magnitude for different redshift bins we find that $\log\left(\mathcal{M}/\mathcal{M}_\odot\right) = 10$ is a good approximation of the limiting mass in the highest redshift bin. In the calculation of $\langle \log \Sigma_{\mathcal{M}}\,(z) \rangle$ we include the effects of completeness in exactly the same way as before, i.e. we compute $\langle \log \Sigma_{\mathcal{M}}\,(z) \rangle$ using a cut in stellar mass $\log\left(\mathcal{M}/\mathcal{M}_\odot\right) \geq 10$ and we weight galaxies with the detection probabilities derived from Fig. 5.3.

In Fig. 5.11 we plot $\langle \log \Sigma_{\mathcal{M}}\,(z) \rangle$ as a function of redshift and find that the average

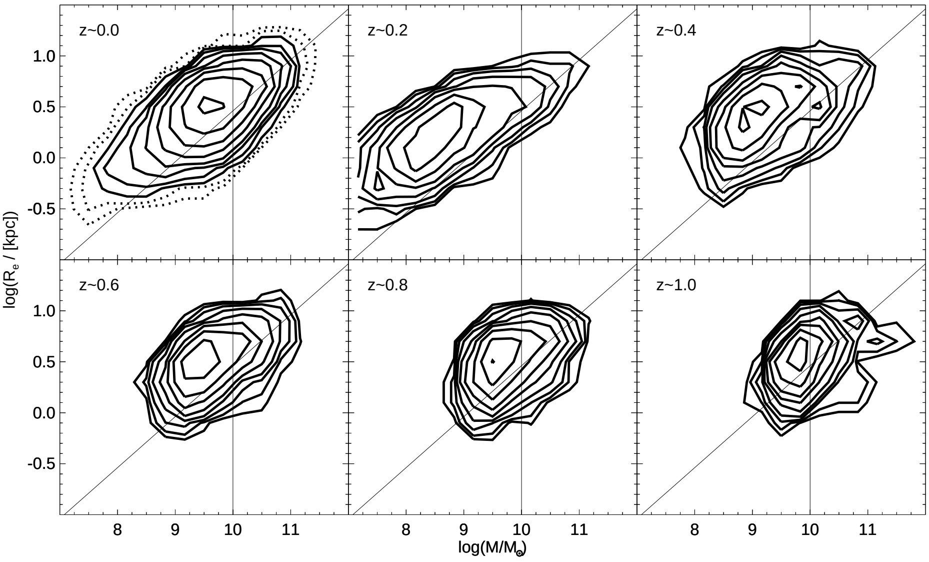


Figure 5.10 The stellar mass–size relation for six different redshift bins ($z \sim 0$ from SDSS; $0.2 < z < 1.0$ from GEMS). The contours indicate levels of the same fractional number of objects per unit area as in Fig. 5.7. The vertical line in each panel marks $\log(\mathcal{M}/\mathcal{M}_\odot) = 10$, which corresponds to the limiting stellar mass $\log \mathcal{M}^{\rm lim}$ applied to each redshift bin. The diagonal line in each panel corresponds to the average surface mass density measured from all redshift bins $\langle \log \Sigma_{\mathcal{M}} (z = 1) \rangle = 8.50$.

surface mass density, to first order, does not evolve significantly with redshift. The overall data values are found within $8.44 < \langle \log \Sigma_{\mathcal{M}} (z) \rangle < 8.57$. This is also illustrated in Fig. 5.12 where we plot the histograms of the stellar surface mass density for the individual redshift bins. The deviation of the lowest and the highest data point corresponds to only 34% in surface mass density. Fitting a line with constant slope zero to the data yields $\langle \log \Sigma_{\mathcal{M}} (z) \rangle = 8.50 \pm 0.03$. We stress that the validity of this estimate does depend strongly on systematic errors in the measurement of the stellar masses. The error bars do not account for such effects and therefore might present a somewhat oversimplied view.

The constancy of the stellar mass–size relation above $\log(\mathcal{M}/\mathcal{M}_\odot) \sim 10$ since $z \sim 1$ comprises a strong constraint on models of disk galaxy evolution. The simplest possible interpretation of the data is that galaxies grow inside-out: assuming that galaxies can only increase their stellar mass with time, in order to stay on the stellar-mass size relation as they grow in mass, galaxies must increase their scale-lengths accordingly. Yet, clearly, more complex and physically-motivated models will also be capable of fitting the data.

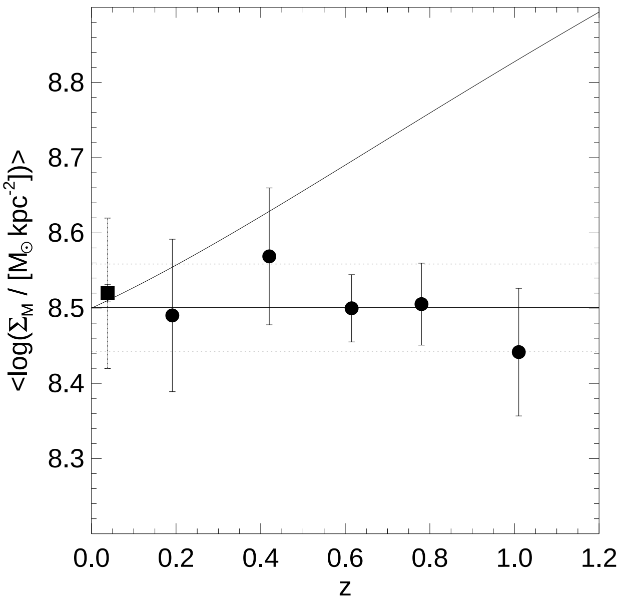


Figure 5.11 Evolution of the average stellar surface mass density $\langle\log\Sigma_{\mathcal{M}}(z)\rangle$. Solid dots show the GEMS data for the individual redshift bins; the box symbol at $z\sim0.05$ indicates the SDSS data point. The solid error bars indicate the 2σ statistical errors; the dotted error bar for the SDSS data point marks the 2σ systematic error resulting from the conversion of a mass-to-light ratio. The horizontal solid and dotted lines at $\langle\log\Sigma_{\mathcal{M}}(z)\rangle = 8.50\pm0.03$ represent a linear fit to the data with a constant slope of zero and the 1σ confidence limits, respectively. The diagonal line indicates the evolution as obtained from Mo *et al.* (1998).

5.5 Discussion

5.5.1 Surface Brightness Evolution

In order to facilitate comparison with previous studies, we repeat the analysis in the rest-frame B-band (using absolute B-band magnitudes from COMBO-17 and correcting the sizes to B-band). We convert the effective surface brightnesses to central surface brightnesses assuming effective size and disk scale length scale as $R_e = 1.678\times R_d$:

$$\mu_{0,B} = M_B + 5\log R_e - 5\log(1.678) + 38.568, \tag{5.9}$$

This is strictly true only for pure disk galaxies, but should be a reasonable approximation since the peak of our Sérsic index distribution roughly coincides with the exponential case $n\sim1$. As before, we find strong evolution in the rest B-band surface brightness with redshift. For the intercept and slope in the rest-frame B-band we find $\langle\mu_{0,B}(z=0)\rangle = 21.11\pm0.03$ mag arcsec^{-2} and -1.43 ± 0.07, respectively (see Fig. 5.13).

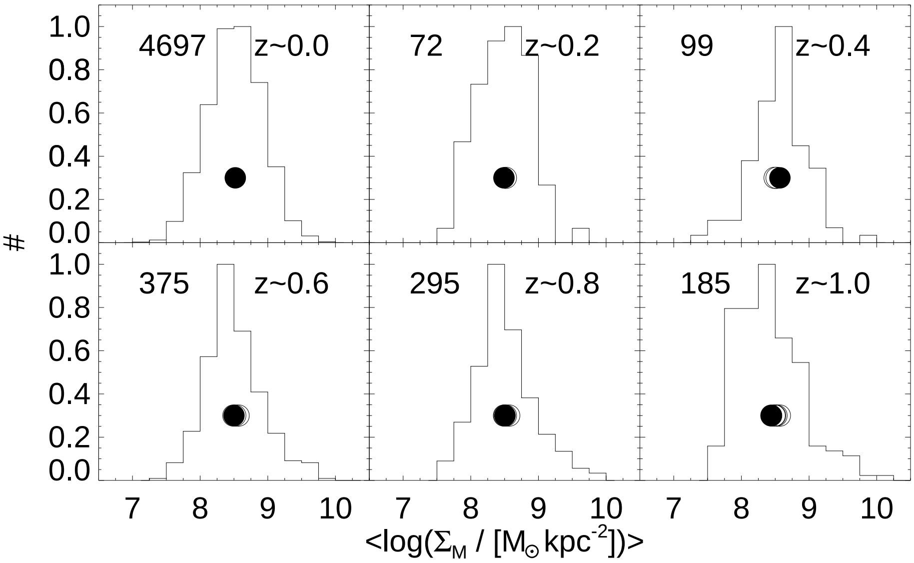


Figure 5.12 Histograms of the stellar surface mass density as a function of redshift ($z \sim 0$ from SDSS; $0.2 < z < 1.0$ from GEMS). The histograms include weighting of individual objects according to their completeness (y-axis peak normalized). Only objects with a completeness exceeding 50% and with $\log\left(\mathcal{M}/\mathcal{M}_{\odot}\right) > \log\mathcal{M}^{\mathrm{lim}} = 10$ were included in order to minimize the impact of selection effects. The black circle marks the estimated mean value $\langle\log\Sigma_{\mathcal{M}}\left(z\right)\rangle$ for each redshift bin. In each redshift bin with $z > 0.0$ we overplotted as open circles values of $\langle\log\Sigma_{\mathcal{M}}\left(z\right)\rangle$ of all lower redshift bins to visualize the evolution in surface mass density. The numbers in the top left of each panel shows the number of sources used in each panel (not including weighting).

In contrast to this picture of strong evolution, several previous authors have found results consistent with weak or no evolution in the average surface brightness out to $z \sim 1$ (e.g., Simard *et al.* 1999; Ravindranath *et al.* 2004). In this section, we discuss how these apparently contradictory findings, based on similar data, can be reconciled.

5.5.1.1 Are the Datasets Significantly Different?

We can rule out differences in the datasets as the source of our divergent conclusions. Owing to the similarity of the datasets, we can reproduce the analysis of Ravindranath *et al.* (2004) in some detail. Ravindranath *et al.* (2004) assessed the average B-band central surface brightness of their sample as a function of redshift, limited in surface brightness to $\mu_{0,B}^{\mathrm{lim}} < 20.6$. For both the GOODS and the GEMS data sets the Ravindranath *et al.* (2004) surface brightness limit implies removing half to two thirds of all galaxies at $0.25 < z < 0.50$ that are detected above the absolute magnitude limit and have a measured redshift. Note that only $\sim 5\%$ of all galaxies were excluded at the highest redshift $(1.00 < z < 1.25)$. Obviously, by using only one third of galaxies with the highest surface

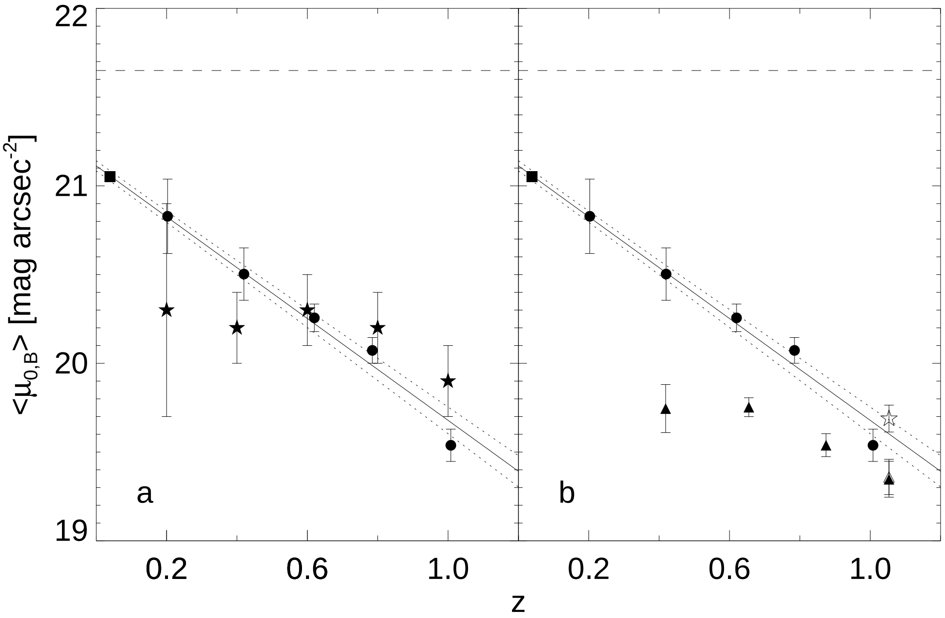


Figure 5.13 Comparison of the evolution of the average absolute central B-band surface brightness $\langle \mu_{0,B}\,(z) \rangle$ with the literature. Solid dots show the GEMS data for the individual redshift bins; the box symbol at $z \sim 0.05$ indicates the SDSS data point. The error bars mark 2σ statistical errors. The horizontal line at $\mu_{0,B} = 21.65$ represents the Freeman (1970) surface brightness. The solid and dotted lines mark a linear fit to all data points (including the SDSS value) plus the 1σ confidence limits, respectively. a) Comparison with Simard *et al.* (1999, star symbols). b) Reproducing the Ravindranath *et al.* (2004) analysis. Triangles show values obtained from the GEMS data using the selection in limiting surface brightness and absolute magnitude as chosen by Ravindranath *et al.* (2004). The open symbols indicate the impact of a population of high surface brightness galaxies at the highest redshift bin using the definition of Ravindranath *et al.* (open triangle, just barely visible above the filled triangle) and the definition by Simard *et al.* (open star). Especially when using the latter definition, the GEMS data with the Ravindranath *et al.* selection are consistent with no surface brightness evolution.

brightness, one introduces a strong bias in the measurement of $\langle \mu_{0,B}\,(z) \rangle$ and the derived value will therefore not represent the average properties of disk galaxies at that redshift. Adoption of the surface brightness limit used by Ravindranath *et al.* (2004) yields very consistent results to theirs for $\langle \mu_{0,B}\,(z) \rangle$ (right-hand panel of Fig. 5.13). For the GEMS data we find evolution at less than the 0.4 mag arcsec^{-2} level using their selection criteria. As expected the high redshift data points are the least affected by their surface brightness limit. However, at lower redshift the results achieved using their selection criteria start to deviate systematically from the analysis we presented earlier. Specifically, the lowest redshift point with the surface brightness cut is more than 10σ off the expected value (as estimated from our linear relation) without such a cut. Simard *et al.* (1999) adopted a very similar strategy, and also found very weak evolution, although in their case low number statistics are also an important source of uncertainty (there are only 5 and 6 galaxies in their lowest two redshift bins, respectively).

5.5.1.2 Are the Analysis Techniques Different?

We argue that the divergence between our conclusions and those of Simard *et al.* (1999) and Ravindranath *et al.* (2004) is driven primarily by important differences in the analysis techniques.

The analysis of Simard *et al.* (1999) and Ravindranath *et al.* (2004), justifiably, imposed the selection function of high-redshift galaxies on the low-redshift galaxy population, and asked whether the average surface brightness of galaxies which one could have in principle seen at $z \sim 1$ has evolved. Clearly, because of cosmological surface brightness dimming the bulk of nearby galaxies would be invisible if placed at $z \sim 1$, and are omitted from consideration. One then finds little difference in the population of local galaxies that would be observable at $z \sim 1$.

In this chapter, we adopt a different approach. In essence, we step gradually outwards from low redshift to higher redshift, asking at each stage if there is any evidence that the results are significantly biased due to cosmological surface brightness dimming. The $z \sim 0$ SDSS data are clearly not surface brightness limited for galaxies with $M_V < -20$. Stepping outwards to $z \sim 0.2$ in the GEMS data, the surface brightness limits are well clear of the observed drop-off in galaxy number density for galaxies with $M_V < -20$. Similarly for $z \sim 0.4$, $z \sim 0.6$, and $z \sim 0.8$; at each redshift we have clearly detected both sides of the size distribution in a region where completeness is $> 90\%$, and the observed drop-off is real. Thus, the observed evolution, at least out to $z \sim 0.8$, is a genuine property of the entire disk galaxy population and is unaffected by surface brightness dimming.

At $z \sim 1$, it is less obvious that the data are well clear of the selection boundaries — we correct for incompleteness using the estimates obtained by applying our pipelines to artificial galaxies. Yet, even at the $z \sim 1$ bin we reach well beyond the peak of the surface brightness distribution (see Fig. 5.8), within the limits that we can confidently correct for incompleteness. Therefore, either the evolution we measure in that bin is roughly correct, or the galaxy surface brightness distribution would have to be bimodal. In that case we could not observe a hypothetical second peak of low surface brightness galaxies. Furthermore, these galaxies would have to fade significantly (and faster than the "normal" galaxy population) with time, because otherwise we would detect these objects at lower redshifts. So far there are neither observational nor theoretical grounds on which to expect such a population of low surface brightness galaxies.

It is worth noting that if the galaxy population did not evolve towards higher surface brightness at higher redshift, we would have seen that in the data, as the sample out to $z \sim 0.8$ is clearly deep enough to probe the entire $M_V < -20$ galaxy distribution in the high completeness region.

5.5.2 A New Population of High Surface Brightness Galaxies at High Redshift?

Simard *et al.* (1999) and Ravindranath *et al.* (2004) suggest that at $z \sim 1$ a distinct population of very high surface brightness galaxies emerges that is not detected at lower redshifts. Simard *et al.* (1999) describe these objects as sources with very high surface brightness $\mu_{0,B} \lesssim 18$ (they found 9 candidates; 18% of all galaxies detected at that redshift). Ravindranath *et al.* (2004) delineate this group of objects as compact ($R_e < 0.8$ kpc) and bright ($M_V < -21.5$). Using their classification, $< 5\%$ of the galaxies at $z \sim 1$ fall into this category.

One might conjecture that the introduction of a new population of high surface brightness galaxies at $z \sim 1$ simply arises in order to interpret the increasing average surface brightness within a global picture of a non-evolving $\langle\mu_{0,B}\,(z)\rangle$. Our results suggest that the *whole distribution* of surface brightnesses shifts with redshift, naturally leading to a larger number of high surface brightness galaxies at higher redshift. Furthermore, we find no evidence that the surface brightness distribution changes its shape (at the 10% level, from inspection of Fig. 5.8).

Interestingly, both Simard *et al.* (1999) and Ravindranath *et al.* (2004) introduce the appearance of this new group of objects just at the high redshift limits of their surveys. At those redshifts $z \sim 1 - 1.2$ their values for the average surface brightness are generally in agreement with the GEMS data points. We have shown in Fig. 5.13b that we can reproduce the effect of a flattening in the evolution of $\langle\mu_{0,B}\,(z)\rangle$ by introducing a hard upper surface brightness cut. However, the bulk of the remaining evolution appeared in the highest redshift bin (and hence one could propose the introduction of a new class of high surface brightness galaxies to account for this). After removing the highest surface brightness galaxies as classified by Ravindranath *et al.* (2004) or Simard *et al.* (1999), even the GEMS data do not show a significant redshift-dependent trend in $\langle\mu_{0,B}\,(z)\rangle$ (see Fig. 5.13b). Although this line of reasoning appears to be consistent it nevertheless has a major drawback. At the lowest redshift the results should agree with the average surface brightness obtained from the SDSS. This fact alone should raise strong concerns regarding the global sampling of the local galaxy population. Only strong evolution of $\langle\mu_{0,B}\,(z)\rangle \sim 21.1 - 1.43z$ can account for both convergence with the local data point and the high redshift results from Simard *et al.* (1999), Ravindranath *et al.* (2004) and the results presented in this chapter.

To summarize, we believe that the weak surface brightness evolution found by Ravindranath *et al.* (2004) and Simard *et al.* (1999), and the emergence of a 'new population' of high surface brightness galaxies at $z \sim 1$ results from differences between their analysis technique — which imposes the high-redshift selection function on galaxies at all redshifts — and our analysis technique, which implicitly steps out gradually from the local towards the high redshift universe, asking if there is any evidence for the galaxy distribution running into the surface brightness detection limits. Applying the same selection criteria as Ravindranath *et al.* (2004), we also found weak surface brightness evolution. The disadvantage of this approach is that it cannot yield quantitative statements about the evolution of the global ensemble of disk galaxies, especially at low redshift.

5.5.3 Comparison with Theoretical Expectations

The basic picture of disk formation within a hierarchical universe posits that the dark matter and gas are 'spun up' by tidal torques in the early universe. The internal angular momentum is generally characterized by the dimensionless spin parameter, λ. Assuming that the gas does not suffer significant loss of specific angular momentum during collapse, the size of the resulting disk R_d is expected to scale as $R_d \propto \lambda r_i$, where r_i is the radius enclosing the gas before collapse (see e.g. Mo *et al.* 1998). In cosmological N-body simulations, it is found that the distribution of values of λ for dark matter halos follows a characteristic log-normal form, and that the value of λ does not correlate with halo mass, nor does the distribution of λ evolve with time (Bullock *et al.* 2001a). Thus, to first order, we expect the size of a disk of fixed mass to scale with time in proportion to the virial radius of the dark matter halo:

$$R_d\left(z\right) = R_d\left(0\right) \times \left[\frac{H\left(z\right)}{H\left(0\right)}\right]^{-2/3}, \tag{5.10}$$

where $R_d\left(0\right)$ is the scale length at $z = 0$ and $H(z)$ is the Hubble parameter as a function of redshift (Mo *et al.* 1998). Using the definition of the surface mass density eq. 5.8 we find:

$$\log \Sigma_{\mathcal{M}}\left(z\right) = \log \Sigma_{\mathcal{M}}\left(0\right) + \frac{4}{3} \log \frac{H\left(z\right)}{H\left(0\right)}, \tag{5.11}$$

with the surface mass density at redshift zero $\log \Sigma_{\mathcal{M}}\left(0\right)$. Since we are interested in the relative evolution only, we normalize the curve to the observed value $\log \Sigma_{\mathcal{M}}\left(0\right) = 8.5$ and show the redshift dependence in Fig. 5.11. The expectation of this very naive model is that disks at $z \sim 1$ should be a factor of two denser at fixed mass than they are at the present day, in clear contradiction with the observational results.

In reality, however, we expect there to be several other competing factors. For example, the internal density profile of the dark matter halo, as commonly characterized by the concentration c, will also impact the final size of the disk, in the sense that halos with higher concentration will produce smaller, denser disks. The average concentration at fixed halo mass is a function of epoch, scaling as $c \propto (1 + z)^{-1}$ (Bullock *et al.* 2001b). Thus, the fact that halos were *less* concentrated at $z \sim 1$ by about a factor of two will tend to counteract the strong evolution in surface density indicated above. As well, there are numerous other complications: there is certainly not a straightforward relationship between halo mass and the mass of baryons that collapse to form a disk; the specific angular momentum of the baryons that comprise the disk may not be equal to that of the dark matter halo; the disk size can be affected by the presence of a pre-existing bulge; and halos with low spin parameters and/or large disk masses may not be able to support a stable disk (Mo *et al.* 1998). In addition, a proper comparison of the predicted evolution of the disk mass-size relation with the data requires a careful treatment of the observational selection effects. We defer this analysis to a future work (Somerville et al. in prep).

5.6 Summary

Based on two-dimensional fits to the light profiles of all GEMS sources we have compiled a complete and unbiased sample of disk galaxies. Our disk sample was defined by its radial profile, specificially by Sérsic profiles with concentrations lower than $n = 2.5$. COMBO-17 provided us with redshifts, rest-frame absolute magnitudes and stellar masses for ~ 5700 sources. In order to compare the GEMS data to a local reference we have obtained the VAGC, containing the same information as provided by GEMS for $\sim 28,000$ nearby ($z < 0.05$) SDSS galaxies. Inspecting the magnitude–size and the stellar mass–size relation for disk galaxies as a function of redshift we have come to the following conclusions:

- At high redshifts $z \sim 1$ the GEMS survey is complete only for galaxies with absolute magnitudes $M_V \lesssim -20$ or stellar masses $\log\left(\mathcal{M}/\mathcal{M}_{\odot}\right) \gtrsim 10$. In order to properly address the potentially severe biases that arise when one attempts to explore the evolution of the galaxy population over this redshift range, we have computed a detailed 2-dimensional selection function and introduced a lower limiting absolute magnitude cut.

- Treating completeness and selection effects carefully, we find that the average surface brightness of disk galaxies increases with redshift, by about 1 magnitude from $z \sim 1$ to the present in the rest-frame V-band.

- The values calculated in our study are consistent at the high redshift end with the results of Ravindranath *et al.* (2004) and Simard *et al.* (1999) and at the low redshift end with the value estimated from the SDSS VAGC. We have shown that the reasons the studies of Simard *et al.* (1999) and Ravindranath *et al.* (2004) reached rather different conclusions from our own (weak or no surface brightness evolution over the same redshift range) are primarily related to the way the data were analyzed, as well as to problems with small number statistics in the lower redshift bins. In particular, applying a hard lower surface brightness cut leads to removing substantial numbers of galaxies in the low redshift bins, and to a strong bias in the estimated value of the *average* surface brightness. This approach yields average surface brightness estimates at low redshift $z \sim 0.2$–0.4 that do not converge with the "zero redshift" results from SDSS. We confirmed that when we apply the same selection criteria to the GEMS data, we obtain results that are consistent with those of Ravindranath *et al.* (2004).

- In contrast to the conclusions of Simard *et al.* (1999) and Ravindranath *et al.* (2004), we find that there is no need to appeal to a new population of high surface brightness galaxies, which makes its appearance at high redshift. The increased number of high surface brightness galaxies at high redshift is a natural result of the surface brightness evolution that we have detected.

- While the magnitude–size relation shows strong evolution with redshift, we show that the stellar mass–size relation stays constant with time.
- The most naive theoretical expectation is that disks of fixed mass should be about a factor of two denser at $z \sim 1$, in clear contradiction with our results. Several competing factors probably conspire to produce the weaker evolution that we observe.
- As the stellar mass of galaxies increases with time, the fact that the surface mass density does not evolve as a function of redshift implies that *on average* disk galaxies form inside-out, i.e. through increasing their disk scale lengths with time as they grow in mass.

5.7 Appendix

5.7.1 Parameterization of the Detection Probability

The detection probability p for the GEMS data as a function of apparent magnitude m is well fitted by a double exponential function:

$$p_{\text{GEMS}} = \exp\left(-\exp\left(\frac{m - m_0}{\sigma}\right)\right) \tag{5.12}$$

with the slope σ and the characteristic magnitude m_0. Both the slope $\sigma = \sigma\left(R_e^{\text{app}}, q\right)$ and the characteristic magnitude $m_0 = m_0\left(R_e^{\text{app}}, q\right)$ are a function of the apperant half-light radius R_e^{app} [pix] and the axis ratio q. Since the smallest objects included in the simulations have half-light radii $R_e^{\text{app}} \geq 0.3$ pix we hold the GEMS completeness fixed at sizes $R_e^{\text{app}} < 0.5$ pix: $\log R = \max\left[\log\left(R_e^{\text{app}}\right), 0.5\right]$. The slope σ is defined as:

$$\sigma = \sigma_0\left(q\right) + \sigma_1\left(q\right) \times R \tag{5.13}$$

with

$$\begin{aligned} \sigma_0\left(q\right) &= 0.0860 + 0.118 \times q \\ \sigma_1\left(q\right) &= 0.308 - 0.0634 \times q \end{aligned} \tag{5.14}$$

The characteristic magnitude m_0 is defined as:

$$\begin{aligned} m_0 = \min[&22.5 + \tilde{m}\,(q) \times \log R + (7.37 - 1.83 \times \tilde{m}\,(q)) \times \log R^2 + \\ &(-3.44 + 0.60 \times \tilde{m}\,(q)) \times \log R^3 + \cos\left(7 \log R - 1.75\right) \\ &\times \exp\left(-4.2 \log R\right) + 5 \log R + 2.5 \log\left(q\right), \mu_{\mathrm{max}}\,(q)] \\ &-5 \log R - 2.5 \log\left(q\right) \end{aligned} \tag{5.15}$$

with

$$\begin{aligned} \tilde{m}\,(q) &= 5.325 + 5.373 \times q - 2.128 \times q^2 \\ \mu_{\mathrm{max}}\,(q) &= 29.80 + 0.0933 \times q \end{aligned} \tag{5.16}$$

Similarly, we fit the detection probability for the COMBO-17 data by a double exponential function:

$$p_{\text{COMBO-17}} = n \times \exp\left(-\exp\left(\frac{m - m_{0,c}}{\sigma_c}\right)\right) \tag{5.17}$$

Since the COMBO-17 detection probability does not depend on the axis ratio, both $\sigma_c = \sigma_c\,(R_e^{\mathrm{app}})$ and $m_{0,c} = m_{0,c}\,(R_e^{\mathrm{app}})$ take much simpler forms as functions of R_e^{app} only. The slope σ_c is defined as:

$$\begin{aligned} \sigma_c = \; & 0.168 + 0.388 \times \\ & \exp\left(-\frac{1}{2}\left(\frac{{}^{\mathrm{m}}\left[-\log \frac{R_e^{\mathrm{app}}}{2}, 0\right] - 2.131}{0.895}\right)^2\right) \end{aligned} \tag{5.18}$$

and the characteristic magnitude $m_{0,c}$ is defined as:

$$m_{0,c} = 23.85 - 0.274 \times \tilde{R} + 0.507 \times \tilde{R}^2 - 0.403 \times \tilde{R}^3 \tag{5.19}$$

with $\tilde{R} = \max\left[\log R_e^{\mathrm{app}}, 0.3\right]$. The normalisation n differs slightly from unity due to the effect of redshift focussing:

$$n = 1.014 + 0.00112 \times \tilde{R} \tag{5.20}$$

5.7.2 Incorporating the Wolf *et al.* (2003a) Completeness Map

The completeness map given in Wolf *et al.* (2003a) contains values for the COMBO-17 detection probability $p_{\text{COMBO-17}} = p_{\text{COMBO-17}}\,(m_{\mathrm{R,aper}}, z, (U-V)_{\mathrm{rest}})$ as a function of the apparent R-band aperture magnitude $m_{\mathrm{R,aper}}$, the redshift z and the $(U\text{-}V)_{\mathrm{rest}}$ rest-frame color. In order to convert this completeness map into our magnitude–size frame $p_{\text{COMBO-17}} = p_{\text{COMBO-17}}\,(m_z, R_e^{\mathrm{app}})$, we take the following approach.

We start with a simulated catalogue containing a uniform distribution of apparent z-band magnitudes, apparent half-light sizes (uniformly distributed in $\log R_e^{\rm app}$) and redshifts. Then, we convert the GEMS z-band magnitude m_z into a COMBO-17 total R-band magnitude $m_{\rm R,tot} = m_{\rm R,tot}\,(m_z, z)$. The following polynomial fit to the data is an adequate description (where RND denotes a normally distributed random number with a mean of zero and a standard deviation of 1):

$$\begin{aligned} m_{\rm R,tot} \;=\; & m_z - [\mathrm{RND} \times (-0.308 + 0.0253 \times m_z) \\ & -0.202 - 0.340 \times z + 3.984 \times z^2 - 13.881 \times z^3 \\ & +13.918 \times z^4 - 4.264 \times z^5 \\ & +2.951 - 0.376 \times m_z + 0.0110 \times m_z^2] \end{aligned} \tag{5.21}$$

Next, we relate $m_{\rm R,tot}$ to $m_{\rm R,aper}$ (the COMBO-17 completeness map is expressed in terms of $m_{\rm R,aper}$). By assuming that the aperture loss for the disk galaxies in COMBO-17 is a function of the half-light radius. For the $n < 2.5$ disk sample we find a linear correlation between the difference of total and aperture COMBO-17 magnitude $m_{\rm R,tot} - m_{\rm R,aper}$ and $\sqrt{R_e^{\rm app}}$:

$$m_{\rm R,aper} = m_{\rm R,tot} - 0.508 + 0.254\sqrt{R_e^{\rm app}} - 0.226 \times \mathrm{RND} \tag{5.22}$$

The scatter about this relation is only 0.23 mag.

Finally, we estimate the COMBO-17 $(U - V)$ color given the GEMS m_z, redshift and size. We find that the following description, which is a function of m_z and redshift, is an adequate representation of the data:

$$\begin{aligned} (U - V)_{\rm rest} \;=\; & \mathrm{RND} \times 0.270 \\ & +0.480 - 0.534 \times z + 0.125 \times z^2 \\ & +2.417 - 0.107 \times m_z \end{aligned} \tag{5.23}$$

Using these transformations, we assign a COMBO-17 detection probability from the completeness map given in Wolf *et al.* (2003a) to each mock GEMS galaxy, where the detection probability is a function of m_z, $R_e^{\rm app}$ and z. We have compared the results of the modeling to the direct values from the Wolf *et al.* (2003a) completeness map. *Statistically the agreement is good and our subsequent conclusions are unaffected by which particular method is chosen.* We have carried out the analysis using both methods, arriving at the same conclusions. In the chapter we refer to our statistical approach in order to clearly demonstrate the fact that COMBO-17 is somewhat deeper in terms of surface brightness.

5.7.3 Analysis of Completeness and Selection Effects

In order to address the effects of our completeness correction and sample selection we take the following approach: For each redshift bin we calculate histograms of $\mu_V\,(z)$ and $\log \Sigma_{\mathcal{M}}\,(z)$ using the inverse detection probability p of each object as a weight. From these "weighted" histograms we measure average values for the rest-frame absolute surface brightness in the V-band, $\langle \mu_V(z) \rangle$ and stellar surface mass density, $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$. These average values and the corresponding errors we obtain by constructing 1,000 Monte-Carlo realisations of the GEMS data for each redshift bin. Each realisation consists of a random subsample of the whole data set containing as many sources as the full set, but allowing for duplicate data points. The adopted average values originate from the average mean value of the 1,000 simulations, while the error bars in $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ were calculated from the scatter of the 1,000 mean value estimations. Using such a procedure, we are able to correct for galaxies missing in the GEMS survey down to the level where we can reliably estimate the detection probability p when calculating average mean values.

The calculation of $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ is affected by three limitations. At some limiting magnitude m_z^{lim} the detection probability p drops to zero. The same occurs at some limiting surface brightness $\mu_z^{\mathrm{app,lim}}$. Both effects limit the range of absolute magnitude and surface brightness that is covered by the GEMS data. The higher the redshift, the brighter is the corresponding limiting absolute magnitude M_V^{lim} and the limiting rest-frame surface brightness μ_V^{lim}. As a result of this, we have to restrict the study of the average galaxy population at each redshift bin to the galaxies brighter than M_V^{lim} and μ_V^{lim} corresponding to the highest redshift bin. Finally, the value of p^{lim}, at which one does not include objects in the calculation of $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$, also potentially impacts on the analysis. In the case of $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ the absolute magnitude limit M_V^{lim} translates into a limiting mass $\log \mathcal{M}^{\mathrm{lim}}$.

As will be shown in our subsequent analysis we find $M_V^{\mathrm{lim}} = -20$ and $\log \mathcal{M}^{\mathrm{lim}} = 10$. We will also provide further proof for the fact that the GEMS data are not limited in surface brightness even at the highest redshift bin. Furthermore, we will show that our results are fairly independent of the choice of p^{lim}. We adopt a rather conservative value of $p^{\mathrm{lim}} = 0.5$. In order to demonstrate these results, we calculate $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ for various combinations of p^{lim}, μ_V^{lim}, M_V^{lim} and $\log \mathcal{M}^{\mathrm{lim}}$.

In Fig. 5.14 we plot $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ as a function of the adopted p^{lim} while holding $M_V^{\mathrm{lim}} = -20$, $\log \mathcal{M}^{\mathrm{lim}} = 10$ and $\mu_V^{\mathrm{lim}} = \infty$ constant. Both $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ do not vary significantly, i.e. one would obtain the same results for $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ using $p^{\mathrm{lim}} = 0.2$ or $p^{\mathrm{lim}} = 0.8$. The reason for this is two-fold. On one hand the absolute magnitude limit $M_V^{\mathrm{lim}} = -20$ and the lower stellar mass limit $\log \mathcal{M}^{\mathrm{lim}} = 10$ are chosen rather conservatively, leading to a removal of almost all sources with $p < 0.8$. On the other hand, once the absolute magnitude limit reaches the region where the detection probability drops, galaxies fall along a line of constant apparent magnitude (see Fig. 5.4, at $m_z \sim 23.75$ the 50% completeness contour is almost vertical). Thus, calculation of a mean surface brightness (or surface mass density) is evenly (un-)affected by the completeness correction (independent of surface brightness). Both these arguments arise from the fact that the GEMS data is not limited in surface brightness.

We repeat this exercise for μ_V^{lim} (see Fig. 5.15). This time we hold $p^{\mathrm{lim}} = 0.5$, $M_V^{\mathrm{lim}} = -20$ and $\log \mathcal{M}^{\mathrm{lim}} = 10$ fixed. We find that there is a characteristic surface brightness at each redshift at which the estimated values of $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ systematically start to deviate towards higher surface brightnesses. This has to be interpreted as the surface brightness at which one starts removing galaxies from the sample with the lowest

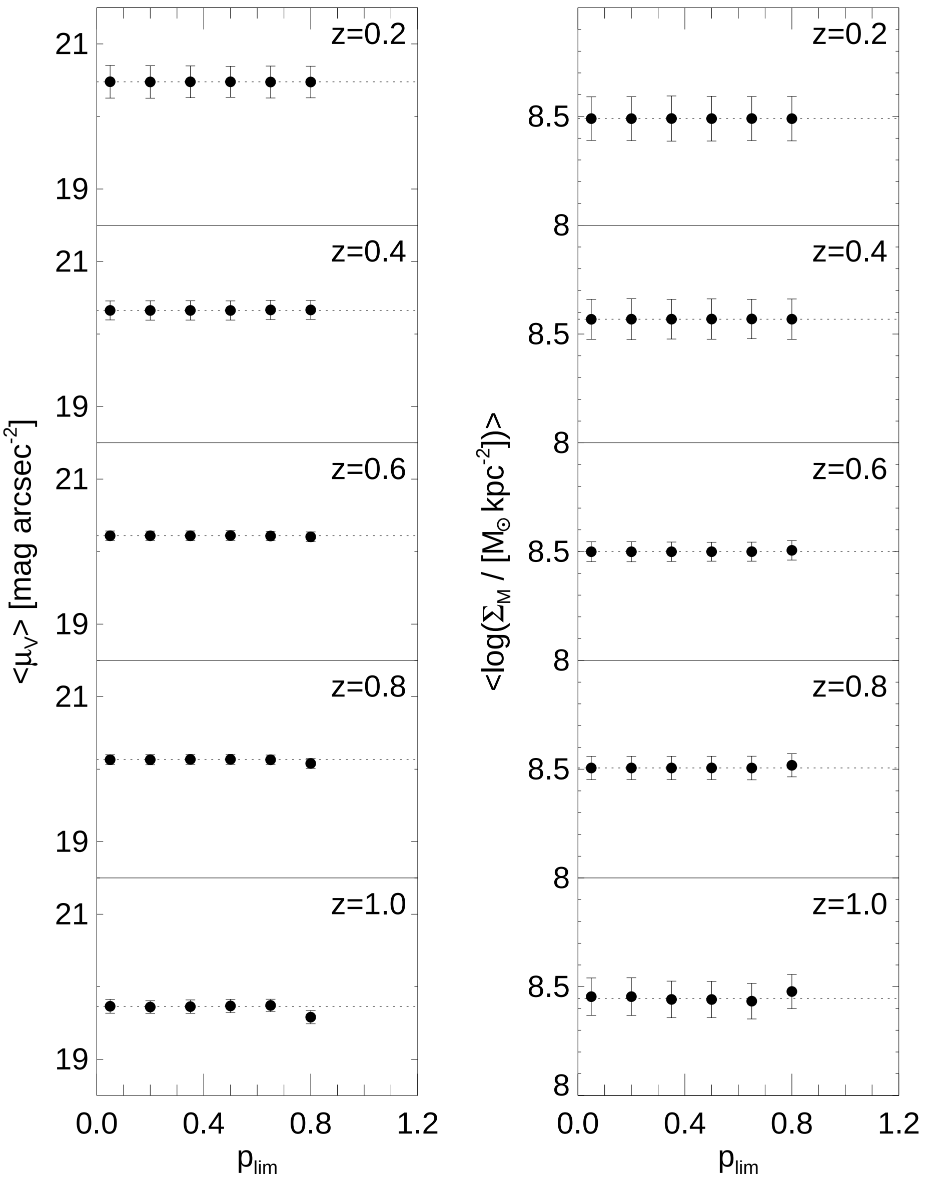


Figure 5.14 The average surface brightness $\langle \mu_V\,(z) \rangle$ (left panel) and the average surface mass density $\langle \log \Sigma_{\mathcal{M}}\,(z) \rangle$ (right panel) for the different redshift bins as a function of the adopted cut-off detection probability $p_{\rm lim}$. Horizontal dotted lines mark the means of each redshift bin. In both panels an absolute limiting magnitude $M_V^{\rm lim} = -20$ was applied.

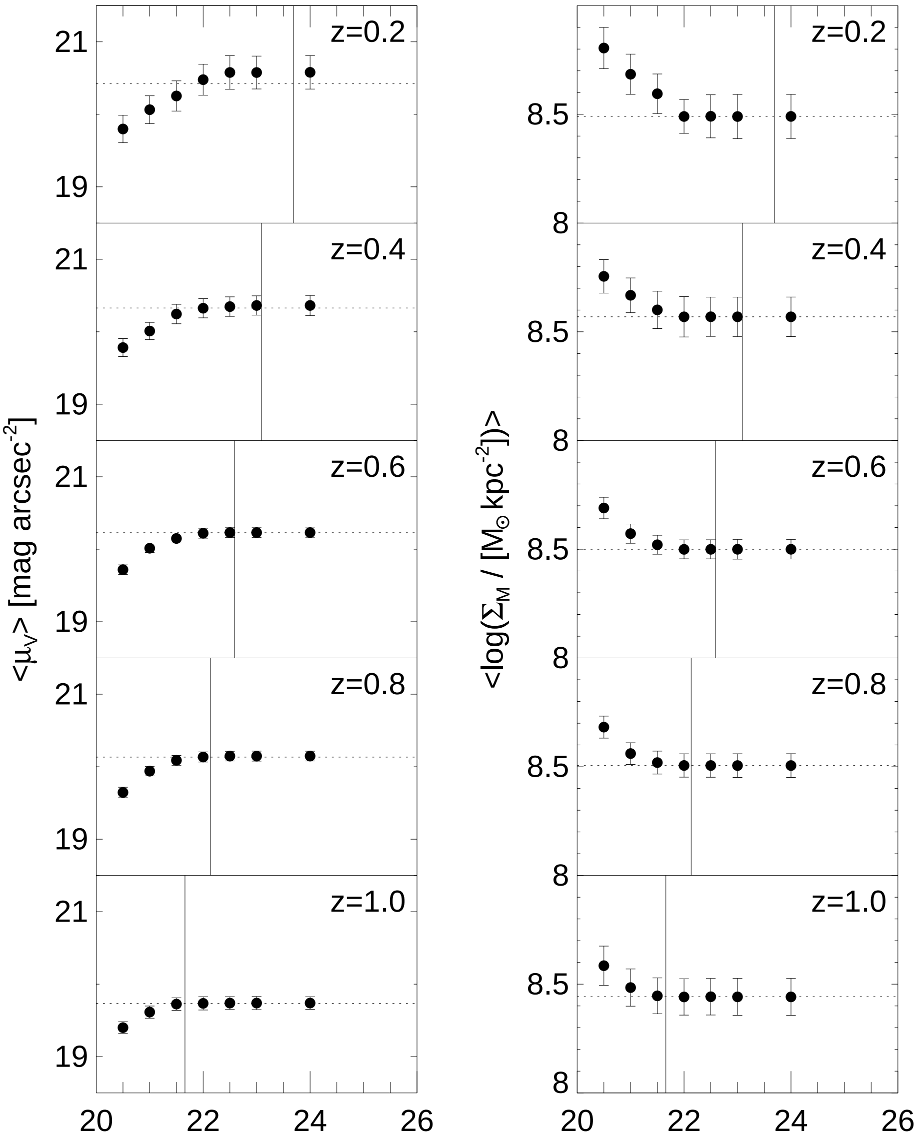


Figure 5.15 The average surface brightness $\langle \mu_V(z) \rangle$ (left panel) and the average surface mass density $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ (right panel) for the different redshift bins as a function of the adopted surface brightness limit $\mu_V^{\rm lim}$. Horizontal dotted lines mark the means of each redshift bin. The values at $\mu_V^{\rm lim} = 24$ are plotted to demonstrate the results in the case were no fixed surface brightness limit ($\mu_V^{\rm lim} = \infty$) is chosen. In both panels an absolute limiting magnitude $M_V^{\rm lim} = -20$ and a detection probability cut $p_{\rm lim} = 0.5$ was applied. The vertical lines correspond to the 50% completeness limiting surface brightness for an axis ratio $q = 0.5$ at the indicated redshift.

surface brightness, thus shifting the average to higher surface brightnesses. Measuring constant mean values at the lowest surface brightnesses, however, does not necessarily imply that the average does not shift. It rather means that we run into our completeness limit eventually, i.e. we do not detect the galaxy population at all that might yet exist at such a faint level. The question is, whether we reach a plateau in $\langle \mu_V(z) \rangle$ or $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ before we run into the GEMS surface brightness limit. To test this, we convert the apparent z-band surface brightness limit, i.e. where a completeness level of 50% is reached, as obtained from the GEMS completeness map $\mu_z^{\text{app},50\%} \sim 23.9$ into a rest-frame surface brightness limit $\mu_V^{50\%}$ in the V-band for each redshift bin using the following relation:

$$\mu_V^{50\%} = \mu_z^{\text{app},50\%} + \left[m_V^{\text{rest}} - m_z\right] - 10\log(1+z) \tag{5.24}$$

with a redshift z dependent color term $m_V^{\text{rest}} - m_z$. The term $-10\log(1+z)$ arises from the surface brightness dimming $\propto (1+z)^4$, which has to be accounted for when converting an apparent surface brightness to an absolute one. From a fit to the GEMS data we obtain:

$$\begin{aligned} m_V^{\text{rest}} - m_z &= M_V - m_z + 5\log D_L + 25 \\ &= 0.562 - 0.111 \times z + 1.160 \times z^2 - 0.841 \times z^3 \end{aligned} \tag{5.25}$$

with the luminosity distance D_L. The values obtained in this manner are indicated in Fig. 5.15 as vertical lines. We find that we can calculate the average galaxy population representing values for $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ at $z = 0.2$ down to a limiting surface brightness $\mu_V^{50\%} \sim 23.7$. At higher redshift the corresponding value has dropped significantly, $\mu_V^{50\%} \sim 21.7$ at $z = 1.0$. Fortunately, even at that redshift we see that the average value $\langle \mu_V(z) \rangle$ has already flattened out, thus implying that even at the high redshift end of the GEMS survey we do sample the full distribution of surface brightnesses. The same line of reasoning also applies to $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$.

Finally, we examine the effect of the choice of M_V^{lim} and $\log \mathcal{M}^{\text{lim}}$ while holding $p^{\text{lim}} = 0.5$ and $\mu_V^{\text{lim}} = \infty$ fixed (see Fig. 5.16). Similarly to Fig. 5.15 we overplot the 50% detection limit $m_z^{50\%} \sim 23.7$ at the faint magnitude end of the completeness map converted to a rest-frame absolute magnitude limit $M_V^{50\%}$. Now the conversion reads:

$$\begin{aligned} M_V^{50\%} &= m_z^{\text{app},50\%} + \left[m_V^{\text{rest}} - m_z\right] - 5\log D_L - 25 \\ &= m_z^{\text{app},50\%} + \left[M_V - m_z\right] \end{aligned} \tag{5.26}$$

with the same definitions as above. We have not attempted to construct a similar relation for the case of $\log \mathcal{M}^{\text{lim}}$. We find that both $\langle \mu_V(z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}}(z) \rangle$ vary systematically as a function of M_V^{lim} and $\log \mathcal{M}^{\text{lim}}$. The reason for this is that the distribution of galaxies in the magnitude–size plane does not exactly fall along a line of constant surface brightness, but has a slightly steeper slope. This is most obvious at the lower redshift bins where we have the largest dynamic range in absolute manitudes. Therefore, our results are strictly true only for the adopted limiting magnitude $M_V^{\text{lim}} = -20$. If one were to repeat our evaluation with deeper data, thus reaching fainter absolute limiting magnitudes at the

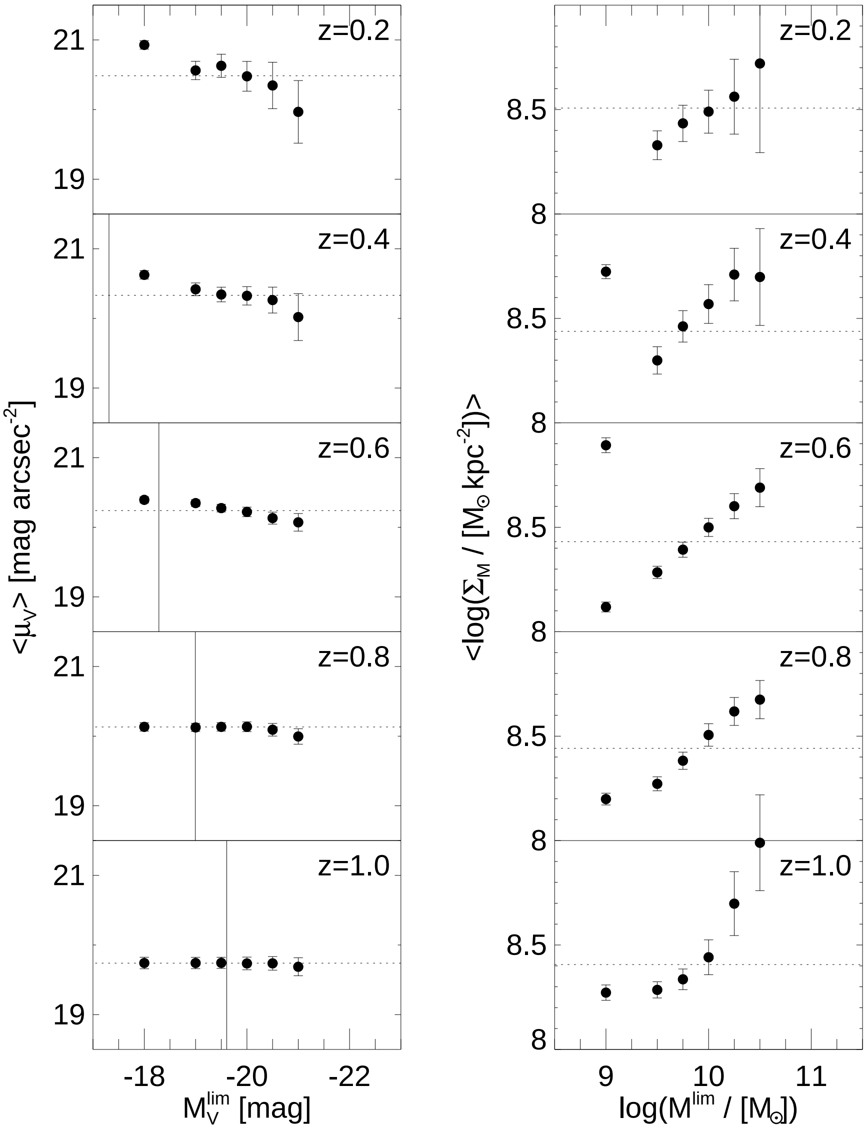


Figure 5.16 The average surface brightness $\langle \mu_V\,(z) \rangle$ (left panel) and the average surface mass density $\langle \log \Sigma_{\mathcal{M}}\,(z) \rangle$ (right panel) for the different redshift bins as a function of the adopted absolute magnitude limit $M_V^{\rm lim}$ or the limiting stellar mass $\log \mathcal{M}^{\rm lim}$ (left and right panel, respectively). Horizontal dotted lines mark the means of each redshift bin. The values at $M_V^{\rm lim} = -18$ (left) and $\log \mathcal{M}^{\rm lim} = 9$ (right) are plotted to demonstrate the results if no fixed absolute magnitude limit ($M_V^{\rm lim} = -\infty$) or stellar mass limit ($\log \mathcal{M}^{\rm lim} = -\infty$) is chosen. In both panels a detection probability cut $p_{\rm lim} = 0.5$ was applied. The vertical lines in the left panel correspond to the 50% completeness limiting absolute magnitude at the indicated redshift.

highest redshift, one should expect to find slightly different absolute values for $\langle \mu_V (z) \rangle$ and $\langle \log \Sigma_{\mathcal{M}} (z) \rangle$. However, if the distribution of galaxies does not change with redshift in the magnitude–size plane, which would imply differential evolution, one would measure the same relative differences. The same line of reasoning of course also holds for $\log \mathcal{M}^{\rm lim} = 10$.

The only way to circumvent this problem would be to move from measuring evolution in the surface brightness to a new variable ρ, which matches the observed slope of the low redshift population[17]. Fitting the slope in our lowest redshift bin, one reads off approximately $\log \rho \propto M + 2.5 \times 3 \log R$, with magnitude M and radius R. This quantity has the physical dimensions of a volume density instead of a surface density, being proportional to the radius cubed.

---

[17] In fact, it should match the slope at all redshifts. However, we have found that usually at higher redshift the dynamic range in absolute magnitude is too small to reasonably constrain the slope.

Chapter 6
Skywalking GEMS and UDF[18]

Recently large high resolution space based imaging surveys for galaxies have been conducted with the Hubble Space Telescope and its Advanced Camera for Surveys (ACS). Very prominent are the Hubble Ultra Deep Field (UDF, Beckwith 2004), the observations in the GOODS field (Giavalisco *et al.* 2004) and our own GEMS survey ('Galaxy Evolution from Morphologies and SEDs', Rix *et al.* 2004), ordered by increasing sky coverage. While the UDF covers one ACS pointing or about $200'' \times 200''$, GEMS covers 70 times as much, nearly $28' \times 28'$. With the observed size and depth of the observations — the UDF is 1–1.5 mag deeper than the Hubble Deep Field — these images are very rich in galaxies of all sizes, shapes, and — all mentioned surveys obtained multi-band images — also colours. The UDF contains almost 10 000 galaxies, GEMS about 40 000.

Both the UDF and GEMS teams created "true colour" images from their data, largely for outreach activity. HST plus its ACS camera have a number of advantages over other instruments, with one combination making it fully unique: the large field-of-view and a $0\farcs05$ spatial sampling. While this is wonderful for science it poses one problem for outreach applications: How to view these images? The UDF in a computer screen sized resolution shows a number of coloured dots and hides most of the beauty of the very deep space, the GEMS colour mosaic viewed at a resolution of 1000 pixel in width is largely black. We think that the GEMS colour mosaic would be best viewed in a printout about 10 meters on a side (still $\sim$6 pixel per mm!), at 50 cm distance to see all details, which might be difficult to realise for a widespread audience. While the underlying amount of image data in JPEG compression is comparably small compared to the science grade images, it is still 10 Mb for UDF and 175 Mb in case of GEMS thus not very friendly for download without access to institute quality internet.

To nevertheless allow people to enjoy the colour mosaics, we put together a JavaScript web application, that allows to pan around in the 1:1 UDF and 2:1 binned GEMS images via the WWW, without downloading all of the images at once, but only the part viewed. In this way access even with analog modems is possible. We dubbed these applications "GEMS Skywalker" and "UDF Skywalker" (names not sponsored by Lucasfilm). They are available online for free use at

`http://www.aip.de/groups/galaxies/sw/gems/` (GEMS Skywalker)
`http://www.aip.de/groups/galaxies/sw/udf/` (UDF Skywalker).

---

[18] This chapter has been published as Jahnke *et al.* (2004a).

They should work with most Netscape, Internet Explorer and Opera Versions.

For the realisation of the Skywalkers, we would like to thank both the UDF and GEMS teams for their work that generated these images. Also a big Thank You goes to the developers of the DynAPI JavaScript library (Steinman 2004) that made programming easy.

Chapter 7
Conclusions & Outlook

7.1 This Work in the Scientific Context of GEMS-Results

GEMS is the largest *HST/ACS* color survey in existence. The combination of its high spatial resolution (to be able to study galaxy morphologies at high redshifts), its big field of view and redshifts from COMBO-17 provides large samples of galaxies needed to examine the evolution of even rare objects. This gives an excellent opportunity to study several important aspects of galaxy evolution out to a redshift of $z \sim 1$. In this thesis, I have concentrated on two of these aspects: The evolution of blue spheroidal galaxies and their role in the picture of hierarchical galaxy assembly through mergers (see § 4), and the size evolution of disk galaxies (§ 5). Before I conclude with the results found for these galaxies, I would like to put my work in the big picture of other results on galaxy evolution.

In GEMS we can draw a consistent picture of galaxy evolution at redshifts $z \lesssim 1$. We examine disk galaxies in detail and from different viewing angles. Wolf *et al.* (2005) and Bell *et al.* (2005) show that disk galaxies host the majority of the star formation at all redshifts out to $z \sim 1$. During their mass assembly, scaling relations do not change significantly, galaxy disks grow in size as they grow in mass (see § 5, also see Barden *et al.* 2005), consistent with a inside-out growth of disk galaxies. We quantitatively examine their transformation into dynamically spheroidal galaxies (see § 4) through galaxy mergers (Bell *et al.* 2006). These newly-formed BSGs (at least the high-density BSGs) are then able to passively evolve into typical red spheroidal galaxies and end up on the red galaxy scaling relations (McIntosh *et al.* 2005), that do not significantly change since $z \sim 1$. AGN host galaxies in this context can be blue spheroid-dominated galaxies (Sánchez *et al.* 2004), well in agreement to the picture that AGN can play a role in the final stages of merging and could be associated with blue spheroids. They might indeed be triggered by galaxy mergers that also induce some star formation and, through feedback, might remove gas from these galaxies, turning them into dead, non-star-forming systems of spheroidal galaxies. This transformation of galaxies from the 'blue cloud' to the 'red sequence' is well observed today (Borch *et al.* 2006, see also § 4); a 50–100% increase in the number of spheroids from $z \sim 1$ to present day is observed through multiple approaches.

7.2 Conclusions

In order to quantitatively measure galaxy evolution since $z \sim 1$, it was necessary to measure structural galaxy parameters reliably. Especially galaxy sizes and Sérsic indices were needed for this kind of analysis. In order to make sure that the properties measured reflect the true properties of the galaxies examined, we intensively tested two different software packages, namely GALFIT and GIM2D.

We found that the initial setup of both fitting codes seriously influences the outcome and the reliability of the parameters returned. Proper usage and intensive testing are essential for every work based on these codes. At least when using GEMS-like data, the GIM2D setup recommended by the author, especially using `'dobkg'`=‘yes’ and `'initparams'`=‘yes’, was unable to recover the input parameters from our simulations of $n = 4$ galaxies that were fainter than the sky surface brightness. We *strongly* discourage users from using these setups, but instead fix the background at a local value as e. g. , returned by SEXTRACTOR, and to input very wide model parameter limits. This turned out to return much more reliable results. GALFIT returned the best results when the sky was fixed at an ‘isophotal’ value given by GALAPAGOS instead of using the SEXTRACTOR value.

Both GALFIT and GIM2D were able to recover the input values of $n = 1$ galaxies with only little bias, but suffered from much more serious bias and scatter when fitting $n = 4$ galaxies. The treatment of neighboring galaxies seemed to be essential during this process and turned out to be the main reason why GIM2D, being unable to simultaneously fit neighboring galaxies, returned significantly worse results than GALFIT. We recommend GALFIT in combination with GALAPAGOS for single Sérsic profile fitting in medium-depth *HST/ACS* data, as this combination of codes returned the best results, both being more reliable in the mean and showing lower scatter and less sensitivity to contamination by neighbors.

The nominal errors returned by each code did underestimate the true uncertainty of the fit by a large factor. Instead we showed a statistical approach to derive fitting uncertainties for individual galaxies and we recommend to carry out a similar analysis when using any of these codes.

From testing both codes on consistency when fitting the same galaxies on images of different depth, we have shown that real galaxies can be reasonably well described by general Sérsic light profiles.

Using stellar population models, we were able to estimate galaxy restframe images on a pixel-to-pixel basis from one-color (two-band) *HST/ACS* data from GEMS. These images, although clearly showing low-level artefacts because of the unavoidable effects of smoothing (required to measure accurate colors), are well-understood and were found to be reliable at the $\sim 10 - 20\%$ level. We intensively tested these images on multi-wavelength *HST* GOODS data.

From fitting these images, we were able to derive a sample of high-Sérsic-index galaxies, which were visually classified to remove contamination of disk galaxies with central light-cusps, that severely bias the sample of ‘early-type’ galaxies when automated classification is used. Visual classification returns a sample of 216 spheroidal galaxies out to $z = 1$, out of which 47 (22%) turn out to have blue colors. This is, to best of our knowledge, by far the biggest sample of blue spheroidal galaxies examined so far, enabling us to separate this sample of galaxies into high-mass-density and low-mass-density galaxies. We find fundamental differences between these two classes of galaxies.

- High-density blue spheroids have the same mass–size distribution as typical red sequence galaxies; that is, they have colors and magnitudes at a given size which allow passive fading and reddening onto the red sequence.
- The low-density blue spheroids are more in agreement with disk galaxies in their mass–size distribution. As we have shown, they are not able to passively evolve into early-type galaxies; in fact, most of them would have to brighten at a given size to fall on local early-type galaxy scaling relations. What these galaxies are — perhaps merger remnants reforming a galaxy disk through further gas accretion, or disk galaxies forming bulges through minor merging — we cannot tell from this analysis.

We measure a total BSG fraction of $\sim 22\%$ at $z \sim 0.6$, comparing reasonably well with previously published determinations of the BSG fraction. Yet, importantly, we find that the vast majority of BSG galaxies are low-density. Restricting our attention to high density BSG galaxies alone, we find that some $\sim 6\%$ of spheroidal galaxies are blue. This fraction of $\sim 6\%$ at $z \sim 0.6$ is a much more physically-meaningful indicator of the growth rate of the early-type galaxy population through the addition of 'young' elliptical galaxies. We estimate a $\sim 0.6\,$Gyr BSG timescale; adopting such a timescale we find a reasonable degree of consistency between the massive galaxy merger fraction and the blue spheroid fraction. Such a consistency gives support to a picture of merger-driven growth of the early-type galaxy population at $z \lesssim 0.8$. Finally, we estimate the growth in the number of early-type galaxies since $z = 0.8$ from galaxies which pass through the blue spheroid stage: the number of spheroid-dominated galaxies at the present-day is predicted to be ~ 1.5 times higher than the number of such galaxies at $z = 0.8$. Thus, blue spheroids represent signposts of cosmologically-significant growth of the spheroidal galaxy population at intermediate redshifts.

Using z-band fit-results of real galaxies from the fitting pipeline tested in this work, we were able to compile a complete and unbiased sample of ~ 5700 disk galaxies by automatic classification of galaxies with $n \leq 2.5$ with redshifts, rest-frame absolute magnitudes and stellar masses from COMBO-17. Using a large comparison sample from SDSS we were able to examine the evolution of the magnitude–size and the mass–size relation as a function of redshift.

We find that the average surface brightness of disk galaxies increases with redshift, by about 1 magnitude from $z \sim 1$ to the present in the rest-frame V-band Barden *et al.* (2005). This is consistent with the results from Ravindranath *et al.* (2004) and Simard *et al.* (1999), the reasons for the different conclusions of these two studies are primarily related to the way they analyzed their data, as well as small number statistics in the lower redshift bins.

While the magnitude–size relation shows strong evolution with redshift, we show that the stellar mass–size relation stays constant with time. The most naive theoretical expectation is that disks of fixed mass should be about a factor of two denser at $z \sim 1$, in clear contradiction with our results. As the stellar mass of galaxies increases with time, the fact that the surface mass density does not evolve as a function of redshift implies that *on average* disk galaxies form inside-out, i.e. through increasing their disk scale lengths with time as they grow in mass.

7.3 Outlook

Large surveys with multi-wavelength coverage in distinct fields, e. g. GEMS, STAGES, COSMOS, GOODS and others, datasets that are currently being build up or already exist, will help to rule out cosmic variations within this analysis, to understand the role of blue spheroidal galaxies and to solve the question whether they are really the progenitors of red-sequence galaxies. High-resolution data (e. g. *HST*) is essential in this kind of work to be able to get reliable galaxy classifications and properties as on ground-based data it is very hard to measure galaxy parameters like e. g. Sérsic indices and halflight-radii. Together, these surveys provide a huge galaxy sample with a wide range in both redshift and environment, making it indeed possible to examine the evolution of this rare class of galaxies. In this work, only half of the galaxy sample in GEMS has been taken into account, we will increase our number statistics significantly. Increasing the sample size is important as blue spheroidal galaxies are not very common objects, number statistics is the limiting factor in most current surveys.

STAGES, centered on the supercluster Abell901/902, will help to understand environmental effects of galaxy evolution in general and blue spheroids in particular. Galaxies in clusters are unlikely to merge with each other due to their high velocity difference, so when galaxies merge, they should in principle do it in galaxy groups, where the velocity difference is small, yet the galaxy density is high enough. If this is true, there should be more blue spheroids in galaxy groups than anywhere else and these galaxy groups should then fall into clusters with already existing elliptical galaxies. The GEMS data alone cannot answer this question; the field is too small to contain sufficient numbers of clusters and, using photometric redshifts, it is impossible to identify galaxy groups. Other large *HST* mosaics like COSMOS with its substantial spectroscopic data and/or STAGES could be able to help solve this problem.

We will use multi-wavelength data, e. g. Spitzer $24\mu m$ to help to identify galaxies with intrinsically high star-formation rates and X-ray data to investigate AGN-accretion in galactic centers, to further improve the sample selection and characterization of blue spheroidal galaxies. This, or similar, data exist for most of these surveys.

More clues about the nature of BSGs could be obtained by spectroscopic observations. If it is true that massive BSGs fade and redden passively into RSGs and low-mass BSGs form stars and evolve ‘back’ into disk dominated systems, then the massive BSGs should contain aging stars with absorption lines while low-mass BSGs with strong ongoing SF should have clear emission features in their spectra. Furthermore, spectroscopy is the key to dynamical masses, which are independent measurements from the stellar mass estimates that we used in our work (Borch *et al.* 2006). More importantly, they are not susceptible to strong changes in the stellar M/L ratio from merger-induced SF bursts. These changes, described by Bell & de Jong (2001), can exceed a factor of 3 in M/L, which translates to an error of the same order in stellar mass.

An interesting analysis of the relation between galaxy mergers and blue spheroidal galaxies would be to look at merging galaxies in the GEMS data that are morphologically distorted (the approach of Bell *et al.* 2006, was based on close pairs which overestimate the fraction of merging galaxies). An automated identification of mergers is not yet possible and only having classified galaxies with $n > 2$, our sample does not necessarily identify all galaxy mergers in the data set unless one classifies **all** galaxies in the GEMS field.

Appendix A
List of Acronyms

ACS	**A**dvanced **C**amera for **S**urveys
AGN	**A**ctive **G**alactic **N**uclei
BSG	blue spheroidal galaxy, in contrast to RSG
COMBO-17	**C**lassifying **O**bjects by **M**edium-**B**and **O**bservations — a spectrophotometric **17**-filter survey
COSMOS	The Cosmic Evolution Survey
CPU	**C**entral **P**rocessing **U**nit
GALAPAGOS	**G**alaxy **A**nalysis over **L**arge **A**reas: **P**arameter **A**ssessment by **G**ALFITting **O**bjects from **S**EXTRACTOR
GALFIT	Well, it's simply GALFIT
GASPHOT	**G**alaxy **A**utomated **S**urface **PHOT**ometry
GEMS	**G**alaxy **E**volution from **M**orphologies and **S**EDs
GIM2D	**G**alaxy **IM**age **2D**
GOODS	The **G**reat **O**bservatories **O**rigins **D**eep **S**urvey
Gyrs	Gigayear, 1 billion years
HST	**H**ubble **S**pace **T**elescope
PSF	**P**oint **S**pread **F**unction
RSG	red spheroidal galaxies, in contrast do BSGs
SED	**S**pectral **E**nergy **D**istribution
SF	star formation
SDSS	**S**loan **D**igital **S**ky **S**urvey
SEXTRACTOR	**S**ource **EX**tractor
STAGES	**S**pace **T**elescope **A**901/902 **G**alaxy **E**volution **S**urvey

Bibliography

Abazajian, K. *et al.* . 2004. The Second Data Release of the Sloan Digital Sky Survey. *AJ*, **128**(July), 502–512.

Abraham, R. G. 1998. Quantitative Morphology at High Redshifts. *ArXiv Astrophysics e-prints*, Feb.

Abraham, R. G., Tanvir, N. R., Santiago, B. X., Ellis, R. S., Glazebrook, K., & van den Bergh, S. 1996. Galaxy morphology to I=25 mag in the Hubble Deep Field. *MNRAS*, **279**(Apr.), L47–L52.

Abraham, R. G., Ellis, R. S., Fabian, A. C., Tanvir, N. R., & Glazebrook, K. 1999. The star formation history of the Hubble sequence: spatially resolved colour distributions of intermediate-redshift galaxies in the Hubble Deep Field. *MNRAS*, **303**(Mar.), 641–658.

Barden, M., Rix, H.-W., Somerville, R. S., Bell, E. F., Häußler, B., Peng, C. Y., Borch, A., Beckwith, S. V. W., Caldwell, J. A. R., Heymans, C., Jahnke, K., Jogee, S., McIntosh, D. H., Meisenheimer, K., Sánchez, S. F., Wisotzki, L., & Wolf, C. 2005. GEMS: The Surface Brightness and Surface Mass Density Evolution of Disk Galaxies. *ApJ*, **635**(Dec.), 959–981.

Barnes, J. E. 1992. Transformations of galaxies. I – Mergers of equal-mass stellar disks. *ApJ*, **393**(July), 484–507.

Barnes, J. E., & Hernquist, L. 1992. Dynamics of interacting galaxies. *ARA&A*, **30**, 705–742.

Beckwith, S. V. W. et al. 2004 (Mar.). *The Hubble Ultra Deep Field project*. Homepage http://www.stsci.edu/hst/udf/.

Bell, E. F. 2003. Estimating Star Formation Rates from Infrared and Radio Luminosities: The Origin of the Radio-Infrared Correlation. *ApJ*, **586**(Apr.), 794–813.

Bell, E. F. 2004. Galaxy Assembly. *ArXiv Astrophysics e-prints*, Aug.

Bell, E. F., & de Jong, R. S. 2001. Stellar Mass-to-Light Ratios and the Tully-Fisher Relation. *ApJ*, **550**(Mar.), 212–229.

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003a. A First Estimate of the Baryonic Mass Function of Galaxies. *ApJ*, **585**(Mar.), L117–L120.

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003b. The Optical and Near-Infrared Properties of Galaxies. I. Luminosity and Stellar Mass Functions. *ApJS*, **149**(Dec.), 289–312.

Bell, E. F., McIntosh, D. H., Barden, M., Wolf, C., Caldwell, J. A. R., Rix, H.-W., Beckwith, S. V. W., Borch, A., Häussler, B., Jahnke, K., Jogee, S., Meisenheimer, K., Peng, C., Sanchez, S. F., Somerville, R. S., & Wisotzki, L. 2004a. GEMS Imaging of Red-Sequence Galaxies at z˜0.7: Dusty or Old? *ApJ*, **600**(Jan.), L11–L14.

Bell, E. F., Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., Kleinheinrich, M., Wisotzki, L., & McIntosh, D. H. 2004b. Nearly 5000 Distant Early-Type Galaxies in COMBO-17: A Red Sequence and Its Evolution since z˜1. *ApJ*, **608**(June), 752–767.

Bell, E. F., Papovich, C., Wolf, C., Le Floc'h, E., Caldwell, J. A. R., Barden, M., Egami, E., McIntosh, D. H., Meisenheimer, K., Pérez-González, P. G., Rieke, G. H., Rieke, M. J., Rigby, J. R., & Rix, H.-W. 2005. Toward an Understanding of the Rapid Decline of the Cosmic Star Formation Rate. *ApJ*, **625**(May), 23–36.

Bell, E. F., Phleps, S., Somerville, R. S., Wolf, C., Borch, A., & Meisenheimer, K. 2006. The Merger Rate of Massive Galaxies. *ApJ*, **652**(Nov.), 270–276.

Bershady, M. A., Jangren, A., & Conselice, C. J. 2000. Structural and Photometric Classification of Galaxies. I. Calibration Based on a Nearby Galaxy Sample. *AJ*, **119**(June), 2645–2663.

Bertin, E., & Arnouts, S. 1996. SExtractor: Software for source extraction. *A&AS*, **117**(June), 393–404.

Bertola, F., & Capaccioli, M. 1975. Dynamics of Early Type Galaxies. *Pages 373–+ of:* Oppenheim, A. V., & Schafer, R. W. (eds), *IAU Symp. 69: Dynamics of the Solar Systems*.

Blanton, M. R., Brinkmann, J., Csabai, I., Doi, M., Eisenstein, D., Fukugita, M., Gunn, J. E., Hogg, D. W., & Schlegel, D. J. 2003a. Estimating Fixed-Frame Galaxy Magnitudes in the Sloan Digital Sky Survey. *AJ*, **125**(May), 2348–2360.

Blanton, M. R., Hogg, D. W., Bahcall, N. A., Baldry, I. K., Brinkmann, J., Csabai, I., Eisenstein, D., Fukugita, M., Gunn, J. E., Ivezić, Ž., Lamb, D. Q., Lupton, R. H., Loveday, J., Munn, J. A., Nichol, R. C., Okamura, S., Schlegel, D. J., Shimasaku, K., Strauss, M. A., Vogeley, M. S., & Weinberg, D. H. 2003b. The Broadband Optical Properties of Galaxies with Redshifts 0.02¡z¡0.22. *ApJ*, **594**(Sept.), 186–207.

Blanton, M. R., Hogg, D. W., Bahcall, N. A., Brinkmann, J., Britton, M., Connolly, A. J., Csabai, I., Fukugita, M., Loveday, J., Meiksin, A., Munn, J. A., Nichol, R. C., Okamura, S., Quinn, T., Schneider, D. P., Shimasaku, K., Strauss, M. A., Tegmark, M., Vogeley, M. S., & Weinberg, D. H. 2003c. The Galaxy Luminosity Function and Luminosity Density at Redshift z = 0.1. *ApJ*, **592**(Aug.), 819–838.

Blanton, M. R., Schlegel, D. J., Strauss, M. A., Brinkmann, J., Finkbeiner, D., Fukugita, M., Gunn, J. E., Hogg, D. W., Ivezic, Z., Knapp, G. R., Lupton, R. H., Munn, J. A., Schneider, D. P., Tegmark, M., & Zehavi, I. 2004. NYU-VAGC: a galaxy catalog based on new public surveys. *ArXiv Astrophysics e-prints*, Oct.

Böhm, A., Ziegler, B. L., Saglia, R. P., Bender, R., Fricke, K. J., Gabasch, A., Heidt, J., Mehlert, D., Noll, S., & Seitz, S. 2004. The Tully-Fisher relation at intermediate redshift. *A&A*, **420**(June), 97–114.

Boissier, S., & Prantzos, N. 1999. Chemo-spectrophotometric evolution of spiral galaxies – I. The model and the Milky Way. *MNRAS*, **307**(Aug.), 857–876.

Borch, A. 2004. Evolution of the stellar mass density of galaxies since redshift 1.0. *Ph.D. Thesis*, July.

Borch, A., Meisenheimer, K., Bell, E. F., Rix, H.-W., Wolf, C., Dye, S., Kleinheinrich, M., Kovacs, Z., & Wisotzki, L. 2006. The stellar masses of 25 000 galaxies at $0.2 \leq z \leq 1.0$ estimated by the COMBO-17 survey. *A&A*, **453**(July), 869–881.

Brown, M. J. I., Dey, A., Jannuzi, B. T., Brand, K., Benson, A. J., Brodwin, M., Croton, D. J., & Eisenhardt, P. R. 2006. The Evolving Luminosity Function of Red Galaxies. *ArXiv Astrophysics e-prints*, Sept.

Bullock, J. S., Dekel, A., Kolatt, T. S., Kravtsov, A. V., Klypin, A. A., Porciani, C., & Primack, J. R. 2001a. A Universal Angular Momentum Profile for Galactic Halos. *ApJ*, **555**(July), 240–257.

Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001b. Profiles of dark haloes: evolution, scatter and environment. *MNRAS*, **321**(Mar.), 559–575.

Caldwell, J. A. R., McIntosh, D. H., Rix, H., Barden, M., Beckwith, S. V. W., Bell, E. F., Borch, A., Heymans, C., Haeussler, B., Jahnke, K., Jogee, S., Meisenheimer, K., Peng, C. Y., Sanchez, S. F., Somerville, R. S., Wisotzki, L., & Wolf, C. 2007, in prep. GEMS Survey Data and Catalog.

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994. Dust extinction of the stellar continua in starburst galaxies: The ultraviolet and optical extinction law. *ApJ*, **429**(July), 582–601.

Caon, N., Capaccioli, M., & D'Onofrio, M. 1993. On the Shape of the Light Profiles of Early Type Galaxies. *MNRAS*, **265**(Dec.), 1013–+.

Charlot, S., & Fall, S. M. 2000. A Simple Model for the Absorption of Starlight by Dust in Galaxies. *ApJ*, **539**(Aug.), 718–731.

Chen, H.-W., Marzke, R. O., McCarthy, P. J., Martini, P., Carlberg, R. G., Persson, S. E., Bunker, A., Bridge, C. R., & Abraham, R. G. 2003. The Las Campanas Infrared Survey. IV. The Photometric Redshift Survey and the Rest-Frame R-Band Galaxy Luminosity Function at $0.5 \leqq z \leqq 1.5$. *ApJ*, **586**(Apr.), 745–764.

Chiappini, C., Matteucci, F., & Gratton, R. 1997. The Chemical Evolution of the Galaxy: The Two-Infall Model. *ApJ*, **477**(Mar.), 765–+.

Cimatti, A., Daddi, E., Mignoli, M., Pozzetti, L., Renzini, A., Zamorani, G., Broadhurst, T., Fontana, A., Saracco, P., Poli, F., Cristiani, S., D'Odorico, S., Giallongo, E., Gilmozzi, R., & Menci, N. 2002. The K20 survey. I. Disentangling old and dusty star-forming galaxies in the ERO population. *A&A*, **381**(Jan.), L68–L72.

Coil, A. L., Davis, M., Madgwick, D. S., Newman, J. A., Conselice, C. J., Cooper, M., Ellis, R. S., Faber, S. M., Finkbeiner, D. P., Guhathakurta, P., Kaiser, N., Koo, D. C., Phillips, A. C., Steidel, C. C., Weiner, B. J., Willmer, C. N. A., & Yan, R. 2004. The DEEP2 Galaxy Redshift Survey: Clustering of Galaxies in Early Data. *ApJ*, **609**(July), 525–538.

Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000. Hierarchical galaxy formation. *MNRAS*, **319**(Nov.), 168–204.

Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., Couch, W., Cross, N., Deeley, K., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Price, I., Seaborne, M., & Taylor, K. 2001. The 2dF Galaxy Redshift Survey: spectra and redshifts. *MNRAS*, **328**(Dec.), 1039–1063.

Conselice, C. J. 2006. Galaxy Mergers and Interactions at High Redshift. *ArXiv Astrophysics e-prints*, Oct.

Couch, W. J., Barger, A. J., Smail, I., Ellis, R. S., & Sharples, R. M. 1998. Morphological Studies of the Galaxy Populations in Distant "Butcher-Oemler" Clusters with the Hubble Space Telescope. II. AC 103, AC 118, and AC 114 at Z = 0.31. *ApJ*, **497**(Apr.), 188–+.

Daddi, E., Broadhurst, T., Zamorani, G., Cimatti, A., Röttgering, H., & Renzini, A. 2001. The spatial clustering of distant, z ˜ 1, early-type galaxies. *A&A*, **376**(Sept.), 825–836.

Davis, M., Faber, S. M., Newman, J., Phillips, A. C., Ellis, R. S., Steidel, C. C., Conselice, C., Coil, A. L., Finkbeiner, D. P., Koo, D. C., Guhathakurta, P., Weiner, B., Schiavon, R., Willmer, C., Kaiser, N., Luppino, G. A., Wirth, G., Connolly, A., Eisenhardt, P., Cooper, M., & Gerke, B. 2003 (Feb.). Science Objectives and Early Results of the DEEP2 Redshift Survey. *Pages 161–172 of:* Guhathakurta, P. (ed), *Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. Edited by Guhathakurta, Puragra. Proceedings of the SPIE, Volume 4834, pp. 161-172 (2003).*

de Jong, R. S. 1996. Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. II. A two-dimensional method to determine bulge and disk parameters. *A&AS*, **118**(Sept.), 557–573.

de Vaucouleurs, G. 1948. Recherches sur les Nebuleuses Extragalactiques. *Annales d'Astrophysique*, **11**(Jan.), 247+.

Di Matteo, T., Springel, V., & Hernquist, L. 2005. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, **433**(Feb.), 604–607.

Dickinson, M., Giavalisco, M., & The Goods Team. 2003. The Great Observatories Origins Deep Survey. *Pages 324–+ of:* Bender, R., & Renzini, A. (eds), *The Mass of Galaxies at Low and High Redshift.*

D'Onghia, E., & Burkert, A. 2004. Bulgeless Galaxies and Their Angular Momentum Problem. *ApJ*, **612**(Sept.), L13–L16.

Dressler, A. 1980. Galaxy morphology in rich clusters – Implications for the formation and evolution of galaxies. *Astrophys. J.*, **236**(Mar.), 351–365.

Drory, N., Bender, R., & Hopp, U. 2004. Comparing Spectroscopic and Photometric Stellar Mass Estimates. *ApJ*, **616**(Dec.), L103–L106.

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962. Evidence from the motions of old stars that the Galaxy collapsed. *ApJ*, **136**(Nov.), 748–+.

Ellis, R. S., Colless, M., Broadhurst, T., Heyl, J., & Glazebrook, K. 1996. Autofib Redshift Survey – I. Evolution of the galaxy luminosity function. *MNRAS*, **280**(May), 235–251.

Fall, S. M., & Efstathiou, G. 1980. Formation and rotation of disc galaxies with haloes. *MNRAS*, **193**(Oct.), 189–206.

Ferguson, A. M. N., & Johnson, R. A. 2001. Constraints on Galaxy Formation from Stars in the Far Outer Disk of M31. *ApJ*, **559**(Sept.), L13–L16.

Ferguson, H. C., Dickinson, M., Giavalisco, M., Kretchmer, C., Ravindranath, S., Idzi, R., Taylor, E., Conselice, C. J., Fall, S. M., Gardner, J. P., Livio, M., Madau, P., Moustakas, L. A., Papovich, C. M., Somerville, R. S., Spinrad, H., & Stern, D. 2004. The Size Evolution of High-Redshift Galaxies. *ApJ*, **600**(Jan.), L107–L110.

Ferreras, I., Lisker, T., Carollo, C. M., Lilly, S. J., & Mobasher, B. 2005. Evolution of Field Early-Type Galaxies: The View from GOODS CDFS. *ApJ*, **635**(Dec.), 243–259.

Fioc, M., & Rocca-Volmerange, B. 1997. PEGASE: a UV to NIR spectral evolution model of galaxies. Application to the calibration of bright galaxy counts. *A&A*, **326**(Oct.), 950–962.

Flores, H., Hammer, F., Thuan, T. X., Césarsky, C., Desert, F. X., Omont, A., Lilly, S. J., Eales, S., Crampton, D., & Le Fèvre, O. 1999. 15 Micron Infrared Space Observatory Observations of the 1415+52 Canada-France Redshift Survey Field: The Cosmic Star Formation Rate as Derived from Deep Ultraviolet, Optical, Mid-Infrared, and Radio Photometry. *ApJ*, **517**(May), 148–167.

Ford, H. C., Clampin, M., Hartig, G. F., Illingworth, G. D., Sirianni, M., Martel, A. R., Meurer, G. R., McCann, W. J., Sullivan, P. C., Bartko, F., Benitez, N., Blakeslee, J., Bouwens, R., Broadhurst, T., Brown, R. A., Burrows, C. J., Campbell, D., Cheng, E. S., Feldman, P. D., Franx, M., Golimowski, D. A., Gronwall, C., Kimble, R. A., Krist, J. E., Lesser, M. P., Magee, D., Miley, G., Postman, M., Rafal, M. D., Rosati, P., Sparks, W. B., Tran, H. D., Tsvetanov, Z. I., Volmer, P., White, R. L., & Woodruff, R. A. 2003 (Feb.). Overview of the Advanced Camera for Surveys on-orbit performance. *Pages 81–94 of:* Blades, J. C., & Siegmund, O. H. W. (eds), *Future EUV/UV and Visible Space Astrophysics Missions and Instrumentation. Edited by J. Chris Blades, Oswald H. W. Siegmund. Proceedings of the SPIE, Volume 4854, pp. 81-94 (2003).*

Freeman, K. C. 1970. On the Disks of Spiral and so Galaxies. *ApJ*, **160**(June), 811–+.

Fried, J. W., von Kuhlmann, B., Meisenheimer, K., Rix, H.-W., Wolf, C., Hippelein, H. H., Kümmel, M., Phleps, S., Röser, H. J., Thierring, I., & Maier, C. 2001. The luminosity function of field galaxies and its evolution since z=1. *A&A*, **367**(Mar.), 788–800.

Fukugita, M., Shimasaku, K., & Ichikawa, T. 1995. Galaxy Colors in Various Photometric Band Systems. *PASP*, **107**(Oct.), 945–+.

Giacconi, R., Rosati, P., Tozzi, P., Nonino, M., Hasinger, G., Norman, C., Bergeron, J., Borgani, S., Gilli, R., Gilmozzi, R., & Zheng, W. 2001. First Results from the X-Ray and Optical Survey of the Chandra Deep Field South. *ApJ*, **551**(Apr.), 624–634.

Giavalisco, M., & GOODS Team. 2003 (May). The Great Observatories Origins Deep Survey: Preliminary Results from HST/ACS imaging. *Pages 722–+ of: Bulletin of the American Astronomical Society.*

Giavalisco, M., Steidel, C. C., Adelberger, K. L., Dickinson, M. E., Pettini, M., & Kellogg, M. 1998. The Angular Clustering of Lyman-Break Galaxies at Redshift Z $\approx$ 3. *ApJ*, **503**(Aug.), 543–+.

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., Dickinson, M., Alexander, D. M., Bauer, F. E., Bergeron, J., Biagetti, C., Brandt, W. N., Casertano, S., Cesarsky, C., Chatzichristou, E., Conselice, C., Cristiani, S., Da Costa, L., Dahlen, T., de Mello, D., Eisenhardt, P., Erben, T., Fall, S. M., Fassnacht, C., Fosbury, R., Fruchter, A., Gardner, J. P., Grogin, N., Hook, R. N., Hornschemeier, A. E., Idzi, R., Jogee, S., Kretchmer, C., Laidler, V., Lee, K. S., Livio, M., Lucas, R., Madau, P., Mobasher, B., Moustakas, L. A., Nonino, M., Padovani, P., Papovich, C., Park, Y., Ravindranath, S., Renzini, A., Richardson, M., Riess, A., Rosati, P., Schirmer, M., Schreier, E., Somerville, R. S., Spinrad, H., Stern, D., Stiavelli, M., Strolger, L., Urry, C. M., Vandame, B., Williams, R., & Wolf, C. 2004. The Great Observatories Origins Deep Survey: Initial Results from Optical and Near-Infrared Imaging. *ApJ*, **600**(Jan.), L93–L98.

Graham, A. W., Driver, S. P., Petrosian, V., Conselice, C. J., Bershady, M. A., Crawford, S. M., & Goto, T. 2005. Total Galaxy Magnitudes and Effective Radii from Petrosian Magnitudes and Radii. *AJ*, **130**(Oct.), 1535–1544.

Gray, M. E., & STAGES collaboration. 2006 (June). STAGES: Linking Galaxy Morphology And Environment In the A901/902 Supercluster With HST. *Pages 66.05–+ of: American Astronomical Society Meeting Abstracts*.

Gunn, J. E., & Gott, J. R. I. 1972. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *ApJ*, **176**(Aug.), 1–+.

Hammer, F., Flores, H., Elbaz, D., Zheng, X. Z., Liang, Y. C., & Cesarsky, C. 2005. Did most present-day spirals form during the last 8 Gyr?. A formation history with violent episodes revealed by panchromatic observations. *A&A*, **430**(Jan.), 115–128.

Hogg, D. W., Blanton, M. R., Brinchmann, J., Eisenstein, D. J., Schlegel, D. J., Gunn, J. E., McKay, T. A., Rix, H.-W., Bahcall, N. A., Brinkmann, J., & Meiksin, A. 2004. The Dependence on Environment of the Color-Magnitude Relation of Galaxies. *ApJ*, **601**(Jan.), L29–L32.

Hubble, E. P. 1936. Realm of the Nebulae. *Yale University Press*.

Im, M., Simard, L., Faber, S. M., Koo, D. C., Gebhardt, K., Willmer, C. N. A., Phillips, A., Illingworth, G., Vogt, N. P., & Sarajedini, V. L. 2002. The DEEP Groth Strip Survey. X. Number Density and Luminosity Function of Field E/S0 Galaxies at z¡1. *ApJ*, **571**(May), 136–171.

Jablonka, P., Martin, P., & Arimoto, N. 1996. The Luminosity-Metallicity Relation for Bulges of Spiral Galaxies. *AJ*, **112**(Oct.), 1415–+.

Jahnke, K., Sanchez, S. F., & Haeussler, B. 2004a. Skywalking GEMS and UDF. *ArXiv Astrophysics e-prints*, Mar.

Jahnke, K., Sánchez, S. F., Wisotzki, L., Barden, M., Beckwith, S. V. W., Bell, E. F., Borch, A., Caldwell, J. A. R., Häussler, B., Heymans, C., Jogee, S., McIntosh, D. H., Meisenheimer, K., Peng, C. Y., Rix, H.-W., Somerville, R. S., & Wolf, C. 2004b. Ultraviolet Light from Young Stars in GEMS Quasar Host Galaxies at 1.8$< z <$2.75. *ApJ*, **614**(Oct.), 568–585.

Jogee, S., Barazza, F. D., Rix, H.-W., Shlosman, I., Barden, M., Wolf, C., Davies, J., Heyer, I., Beckwith, S. V. W., Bell, E. F., Borch, A., Caldwell, J. A. R., Conselice, C. J., Dahlen, T., Häussler, B., Heymans, C., Jahnke, K., Knapen, J. H., Laine, S., Lubell, G. M., Mobasher, B., McIntosh, D. H., Meisenheimer, K., Peng, C. Y., Ravindranath, S., Sanchez, S. F., Somerville, R. S., & Wisotzki, L. 2004. Bar Evolution over the Last 8 Billion Years: A Constant Fraction of Strong Bars in the GEMS Survey. *ApJ*, **615**(Nov.), L105–L108.

Katz, N., & Gunn, J. E. 1991. Dissipational galaxy formation. I – Effects of gasdynamics. *ApJ*, **377**(Aug.), 365–381.

Kauffmann, G., & Charlot, S. 1998. Chemical enrichment and the origin of the colour-magnitude relation of elliptical galaxies in a hierarchical merger model. *MNRAS*, **294**(Mar.), 705–+.

Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. *MNRAS*, **264**(Sept.), 201–+.

Kauffmann, G., Heckman, T. M., White, S. D. M., Charlot, S., Tremonti, C., Brinchmann, J., Bruzual, G., Peng, E. W., Seibert, M., Bernardi, M., Blanton, M., Brinkmann, J., Castander, F., Csábai, I., Fukugita, M., Ivezic, Z., Munn, J. A., Nichol, R. C., Padmanabhan, N., Thakar, A. R., Weinberg, D. H., & York, D. 2003a. Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey. *MNRAS*, **341**(May), 33–53.

Kauffmann, G., Heckman, T. M., White, S. D. M., Charlot, S., Tremonti, C., Peng, E. W., Seibert, M., Brinkmann, J., Nichol, R. C., SubbaRao, M., & York, D. 2003b. The dependence of star formation history and internal structure on stellar mass for 10^5 low-redshift galaxies. *MNRAS*, **341**(May), 54–69.

Kauffmann, G., Heckman, T. M., De Lucia, G., Brinchmann, J., Charlot, S., Tremonti, C., White, S. D. M., & Brinkmann, J. 2006. Gas infall and stochastic star formation in galaxies in the local universe. *MNRAS*, **367**(Apr.), 1394–1408.

Khochfar, S., & Burkert, A. 2005. On the origin of isophotal shapes in elliptical galaxies. *MNRAS*, **359**(June), 1379–1385.

Kormendy, J., & Bender, R. 1996. A Proposed Revision of the Hubble Sequence for Elliptical Galaxies. *ApJ*, **464**(June), L119+.

Kranz, T., Slyz, A., & Rix, H.-W. 2003. Dark Matter within High Surface Brightness Spiral Galaxies. *ApJ*, **586**(Mar.), 143–151.

Krist, J. 1993 (Jan.). Tiny Tim : an HST PSF Simulator. *Pages 536–+ of:* Hanisch, R. J., Brissenden, R. J. V., & Barnes, J. (eds), *ASP Conf. Ser. 52: Astronomical Data Analysis Software and Systems II.*

Kroupa, P., Tout, C. A., & Gilmore, G. 1993. The distribution of low-mass stars in the Galactic disc. *MNRAS*, **262**(June), 545–587.

Labbé, I., Rudnick, G., Franx, M., Daddi, E., van Dokkum, P. G., Förster Schreiber, N. M., Kuijken, K., Moorwood, A., Rix, H.-W., Röttgering, H., Trujillo, I., van der Wel, A., van der Werf, P., & van Starkenburg, L. 2003. Large Disklike Galaxies at High Redshift. *ApJ*, **591**(July), L95–L98.

Larson, R. B. 1974. Dynamical models for the formation and evolution of spherical galaxies. *MNRAS*, **166**(Mar.), 585–616.

Le Fèvre, O., Vettolani, G., Maccagni, D., Picat, J.-P., Garilli, B., Tresse, L., Adami, C., Arnaboldi, M., Arnouts, S., Bardelli, S., Bolzonella, M., Bottini, D., Buzzarello, G., Charlot, S., Chincarini, G., Contini, T., Foucaud, S., Franzetti, P., Guzzo, L., Gwyn, S., Ilbert, O., Iovino, A., Le Brun, V., Longhetti, M., Marinoni, C., Methez, G., Mazure, A., McCracken, H., Mellier, Y., Meneux, B., Merluzzi, P., Paltani, S., Pellò, R., Pollo, A., Radovich, M., Rippepi, P., Rizzo, D., Scaramella, R., Scodeggio, M., Zamorani, G., Zanichelli, A., & Zucca, E. 2003. The VIRMOS-VLT Deep Survey: a progress report. *The Messenger*, **111**(Mar.), 18–21.

Lenzen, R., Hartung, M., Brandner, W., Finger, G., Hubin, N. N., Lacombe, F., Lagrange, A.-M., Lehnert, M. D., Moorwood, A. F. M., & Mouillet, D. 2003 (Mar.). NAOS-CONICA first on sky results in a variety of observing modes. *Pages 944–952 of:* Iye, M., & Moorwood, A. F. M. (eds), *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. Edited by Iye, Masanori; Moorwood, Alan F. M. Proceedings of the SPIE, Volume 4841, pp. 944-952 (2003).*

Lilly, S., Schade, D., Ellis, R., Le Fevre, O., Brinchmann, J., Tresse, L., Abraham, R., Hammer, F., Crampton, D., Colless, M., Glazebrook, K., Mallen-Ornelas, G., & Broadhurst, T. 1998. Hubble Space Telescope Imaging of the CFRS and LDSS Redshift Surveys. II. Structural Parameters and the Evolution of Disk Galaxies to Z $\approx$ 1. *ApJ*, **500**(June), 75–+.

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996. The Canada-France Redshift Survey: The Luminosity Density and Star Formation History of the Universe to Z $\approx$ 1. *ApJ*, **460**(Mar.), L1+.

Lotz, J. M., Davis, M., Faber, S. M., Guhathakurta, P., Gwyn, S., Huang, J., Koo, D. C., Le Floc'h, E., Lin, L., Newman, J., Noeske, K., Papovich, C., Willmer, C. N. A., Coil, A., Conselice, C. J., Cooper, M., Hopkins, A. M., Metevier, A., Primack, J., Rieke, G., & Weiner, B. J. 2006a. The Evolution of Galaxy Mergers and Morphology at $z < 1.2$ in the Extended Groth Strip. *ArXiv Astrophysics e-prints*, Feb.

Lotz, J. M., Madau, P., Giavalisco, M., Primack, J., & Ferguson, H. C. 2006b. The Rest-Frame Far-Ultraviolet Morphologies of Star-forming Galaxies at z $\approx$ 1.5 and 4. *ApJ*, **636**(Jan.), 592–609.

Maller, A. H., Dekel, A., & Somerville, R. 2002. Modelling angular-momentum history in dark-matter haloes. *MNRAS*, **329**(Jan.), 423–430.

McIntosh, D. H., Bell, E. F., Rix, H.-W., Wolf, C., Heymans, C., Peng, C. Y., Somerville, R. S., Barden, M., Beckwith, S. V. W., Borch, A., Caldwell, J. A. R., Häußler, B., Jahnke, K., Jogee, S., Meisenheimer, K., Sánchez, S. F., & Wisotzki, L. 2005. The Evolution of Early-Type Red Galaxies with the GEMS Survey: Luminosity-Size and Stellar Mass-Size Relations Since z=1. *ApJ*, **632**(Oct.), 191–209.

Menanteau, F., Ford, H. C., Illingworth, G. D., Sirianni, M., Blakeslee, J. P., Meurer, G. R., Martel, A. R., Benítez, N., Postman, M., Franx, M., Ardila, D. R., Bartko, F., Bouwens, R. J., Broadhurst, T. J., Brown, R. A., Burrows, C. J., Cheng, E. S., Clampin, M., Cross, N. J. G., Feldman, P. D., Golimowski, D. A., Gronwall, C., Hartig, G. F., Infante, L., Kimble, R. A., Krist, J. E., Lesser, M. P., Miley, G. K., Rosati, P., Sparks, W. B., Tran, H. D., Tsvetanov, Z. I., White, R. L., & Zheng, W. 2004. Internal Color Properties of Resolved Spheroids in the Deep Hubble Space Telescope Advanced Camera for Surveys Field of UGC 10214. *ApJ*, **612**(Sept.), 202–214.

Mihos, C., Dubinski, J., & Hernquist, L. 1995. Using Tidal Tails to Probe Dark Matter Halos. *Bulletin of the American Astronomical Society*, **27**(May), 868–+.

Mo, H. J., Mao, S., & White, S. D. M. 1998. The formation of galactic discs. *MNRAS*, **295**(Apr.), 319–336.

Moustakas, L. A., Casertano, S., Conselice, C. J., Dickinson, M. E., Eisenhardt, P., Ferguson, H. C., Giavalisco, M., Grogin, N. A., Koekemoer, A. M., Lucas, R. A., Mobasher, B., Papovich, C., Renzini, A., Somerville, R. S., & Stern, D. 2004. Morphologies and Spectral Energy Distributions of Extremely Red Galaxies in the GOODS-South Field. *ApJ*, **600**(Jan.), L131–L134.

Naab, T., & Burkert, A. 1999. Formation of elliptical galaxies in the merger scenario. *Pages 11–+ of:* Schielicke, R. E. (ed), *Astronomische Gesellschaft Meeting Abstracts*.

Naab, T., & Burkert, A. 2001. The Formation of Disks in Elliptical Galaxies. *ApJ*, **555**(July), L91–L94.

Naab, T., Johansson, P. H., Ostriker, J. P., & Efstathiou, G. 2005. Formation of early-type galaxies from cosmological initial conditions. *ArXiv Astrophysics e-prints*, Dec.

Navarro, J. F., & Steinmetz, M. 2000. Dark Halo and Disk Galaxy Scaling Laws in Hierarchical Universes. *ApJ*, **538**(Aug.), 477–488.

Navarro, J. F., & White, S. D. M. 1994. Simulations of dissipative galaxy formation in hierarchically clustering universes-2. Dynamics of the baryonic component in galactic haloes. *MNRAS*, **267**(Mar.), 401–412.

Norberg, P., Cole, S., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Cross, N. J. G., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Sutherland, W., & Taylor, K. 2002. The 2dF Galaxy Redshift Survey: the b_J-band galaxy luminosity function and survey selection function. *MNRAS*, **336**(Nov.), 907–931.

Patton, D. R., Pritchet, C. J., Carlberg, R. G., Marzke, R. O., Yee, H. K. C., Hall, P. B., Lin, H., Morris, S. L., Sawicki, M., Shepherd, C. W., & Wirth, G. D. 2002. Dynamically Close Galaxy Pairs and Merger Rate Evolution in the CNOC2 Redshift Survey. *Astrophys. J.*, **565**(Jan.), 208–222.

Peebles, P. J. E. 1969. Origin of the Angular Momentum of Galaxies. *ApJ*, **155**(Feb.), 393–+.

Peletier, R. F., Balcells, M., Davies, R. L., Andredakis, Y., Vazdekis, A., Burkert, A., & Prada, F. 1999. Galactic bulges from Hubble Space Telescope NICMOS observations: ages and dust. *MNRAS*, **310**(Dec.), 703–716.

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002. Detailed Structural Decomposition of Galaxy Images. *Astronomical J.*, **124**(July), 266–293.

Percival, W. J., Sutherland, W., Peacock, J. A., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Moody, S., Norberg, P., Peterson, B. A., & Taylor, K. 2002. Parameter constraints for flat cosmologies from cosmic microwave background and 2dFGRS power spectra. *MNRAS*, **337**(Dec.), 1068–1080.

Petrosian, V. 1976. Surface brightness and evolution of galaxies. *ApJ*, **209**(Oct.), L1–L5.

Phleps, S., & Meisenheimer, K. 2003. The evolution of galaxy clustering since z = 1 from the Calar Alto Deep Imaging Survey (CADIS). *A&A*, **407**(Sept.), 855–868.

Pignatelli, E., Fasano, G., & Cassata, P. 2006. GASPHOT: a tool for Galaxy Automatic Surface PHOTometry. *A&A*, **446**(Jan.), 373–388.

Pizagno, J., Prada, F., Weinberg, D. H., Rix, H.-W., Harbeck, D., Grebel, E. K., Bell, E. F., Brinkmann, J., Holtzman, J., & West, A. 2005. Dark Matter and Stellar Mass in the Luminous Regions of Disk Galaxies. *ApJ*, **633**(Nov.), 844–856.

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

Ravindranath, S., Ferguson, H. C., Conselice, C., Giavalisco, M., Dickinson, M., Chatzichristou, E., de Mello, D., Fall, S. M., Gardner, J. P., Grogin, N. A., Hornschemeier, A., Jogee, S., Koekemoer, A., Kretchmer, C., Livio, M., Mobasher, B., & Somerville, R. 2004. The Evolution of Disk Galaxies in the GOODS-South Field: Number Densities and Size Distribution. *ApJ*, **604**(Mar.), L9–L12.

Rix, H.-W., & Rieke, M. J. 1993. Tracing the Stellar Mass in M51. *ApJ*, **418**(Nov.), 123–+.

Rix, H.-W., Barden, M., Beckwith, S. V. W., Bell, E. F., Borch, A., Caldwell, J. A. R., Häussler, B., Jahnke, K., Jogee, S., McIntosh, D. H., Meisenheimer, K., Peng, C. Y., Sanchez, S. F., Somerville, R. S., Wisotzki, L., & Wolf, C. 2004. GEMS: Galaxy Evolution from Morphologies and SEDs. *ApJS*, **152**(June), 163–173.

Roberts, M. S., & Haynes, M. P. 1994. Physical Parameters along the Hubble Sequence. *Annual Review of Astronomy and Astrophysics*, **32**, 115–152.

Rocha-Pinto, H. J., Scalo, J., Maciel, W. J., & Flynn, C. 2000. Chemical enrichment and star formation in the Milky Way disk. II. Star formation history. *A&A*, **358**(June), 869–885.

Rosati, P., Tozzi, P., Giacconi, R., Gilli, R., Hasinger, G., Kewley, L., Mainieri, V., Nonino, M., Norman, C., Szokoly, G., Wang, J. X., Zirm, A., Bergeron, J., Borgani, S., Gilmozzi, R., Grogin, N., Koekemoer, A., Schreier, E., & Zheng, W. 2002. The Chandra Deep Field-South: The 1 Million Second Exposure. *ApJ*, **566**(Feb.), 667–674.

Ryden, B. S. 2006. The Intrinsic Shape of Spiral Galaxies in the 2MASS Large Galaxy Atlas. *ApJ*, **641**(Apr.), 773–784.

Sánchez, S. F., Jahnke, K., Wisotzki, L., McIntosh, D. H., Bell, E. F., Barden, M., Beckwith, S. V. W., Borch, A., Caldwell, J. A. R., Häussler, B., Jogee, S., Meisenheimer, K., Peng, C. Y., Rix, H.-W., Somerville, R. S., & Wolf, C. 2004. Colors of Active Galactic Nucleus Host Galaxies at 0.5¡z¡1.1 from the GEMS Survey. *ApJ*, **614**(Oct.), 586–606.

Sanders, D. B., & Mirabel, I. F. 1996. Luminous Infrared Galaxies. *ARA&A*, **34**, 749–+.

Sargent, M. T., Carollo, C. M., Lilly, S. J., Scarlata, C., Feldmann, R., Kampczyk, P., Koekemoer, A. M., Scoville, N., Kneib, J. ., Leauthaud, A., Massey, R., Rhodes, J., Tasca, L. A. M., Capak, P., McCracken, H. J., Porciani, C., Renzini, A., Taniguchi, Y., Thompson, D. J., Sheth, K., & for the COSMOS collaboration. 2006. The evolution of the number density of large disk galaxies in COSMOS. *ArXiv Astrophysics e-prints*, Sept.

Schade, D., Barrientos, L. F., & Lopez-Cruz, O. 1997. Evolution of Cluster Ellipticals at 0.2$< z <$1.2 from Hubble Space Telescope Imaging. *ApJ*, **477**(Mar.), L17+.

Schade, D., Lilly, S. J., Crampton, D., Ellis, R. S., Le Fèvre, O., Hammer, F., Brinchmann, J., Abraham, R., Colless, M., Glazebrook, K., Tresse, L., & Broadhurst, T. 1999. Hubble Space Telescope Imaging of the CFRS and LDSS Redshift Surveys. III. Field Elliptical Galaxies at 0.2$< z <$1.0. *ApJ*, **525**(Nov.), 31–46.

Scoville, N., Aussel, H., Brusa, M. andCapak, P., Carollo, M., Elvis, C., Giavalisco, M., Guzzo, M. L., Hasinger, G., Impey, C., Kneib, J.-P., LeFevre, O., Lilly, S. J., Mobasher, B., Renzini, A., Rich, R. M., Sanders, D.B., Schinnerer, E., Schminovich, D., Shopbell, P., Taniguchi, Y., & Tyson, N. D. 2007. COSMOS. *submitted to ApJ.*

Sersic, J. L. 1968. *Atlas de galaxias australes*. Cordoba, Argentina: Observatorio Astronomico, 1968.

Shen, S., Mo, H. J., White, S. D. M., Blanton, M. R., Kauffmann, G., Voges, W., Brinkmann, J., & Csabai, I. 2003. The size distribution of galaxies in the Sloan Digital Sky Survey. *MNRAS*, **343**(Aug.), 978–994.

Simard, L. 1998. GIM2D: an IRAF package for the Quantitative Morphology Analysis of Distant Galaxies. *Pages 108–+ of:* Albrecht, R., Hook, R. N., & Bushouse, H. A. (eds), *ASP Conf. Ser. 145: Astronomical Data Analysis Software and Systems VII.*

Simard, L., & Pritchet, C. J. 1998. Internal Kinematics of Field Galaxies at Intermediate Redshifts. *ApJ*, **505**(Sept.), 96–110.

Simard, L., Koo, D. C., Faber, S. M., Sarajedini, V. L., Vogt, N. P., Phillips, A. C., Gebhardt, K., Illingworth, G. D., & Wu, K. L. 1999. The Magnitude-Size Relation of Galaxies out to $z \approx 1$. *ApJ*, **519**(July), 563–579.

Simard, L., Willmer, C. N. A., Vogt, N. P., Sarajedini, V. L., Phillips, A. C., Weiner, B. J., Koo, D. C., Im, M., Illingworth, G. D., & Faber, S. M. 2002. The DEEP Groth Strip Survey. II. Hubble Space Telescope Structural Parameters of Galaxies in the Groth Strip. *ApJS*, **142**(Sept.), 1–33.

Skrutskie, M. F. 1997. *The Impact of Large Scale Near-IR Sky Surveys*. Dordrecht: Kluwer Academic Publishing Company, 1997.

Somerville, R. S., & Primack, J. R. 1999. Semi-analytic modelling of galaxy formation: the local Universe. *MNRAS*, **310**(Dec.), 1087–1110.

Sommer-Larsen, J., & Dolgov, A. 2001. Formation of Disk Galaxies: Warm Dark Matter and the Angular Momentum Problem. *ApJ*, **551**(Apr.), 608–623.

Sommer-Larsen, J., Gelato, S., & Vedel, H. 1999. Formation of Disk Galaxies: Feedback and the Angular Momentum Problem. *ApJ*, **519**(July), 501–512.

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2003. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *ApJS*, **148**(Sept.), 175–194.

Springel, V., & Hernquist, L. 2003. Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. *MNRAS*, **339**(Feb.), 289–311.

Springel, V., Yoshida, N., & White, S. D. M. 2001. GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Astronomy*, **6**(Apr.), 79–117.

Springel, V., Di Matteo, T., & Hernquist, L. 2005. Black Holes in Galaxy Mergers: The Formation of Red Elliptical Galaxies. *ApJ*, **620**(Feb.), L79–L82.

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996. Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts z>3. *ApJ*, **462**(May), L17+.

Steinman, D. *et al.* . 2004. *The DynAPI 2 cross-browser JavaScript library, V2.5.7*. Homepage http://dynapi.sourceforge.net.

Steinmetz, M., & Navarro, J. F. 2002. The hierarchical origin of galaxy morphologies. *New Astronomy*, **7**(June), 155–160.

Stetson, P. B. 1987. DAOPHOT – A computer program for crowded-field stellar photometry. *PASP*, **99**(Mar.), 191–222.

Strateva, I., Ivezić, Ž., Knapp, G. R., Narayanan, V. K., Strauss, M. A., Gunn, J. E., Lupton, R. H., Schlegel, D., Bahcall, N. A., Brinkmann, J., Brunner, R. J., Budavári, T., Csabai, I., Castander, F. J., Doi, M., Fukugita, M., Győry, Z., Hamabe, M., Hennessy, G., Ichikawa, T., Kunszt, P. Z., Lamb, D. Q., McKay, T. A., Okamura, S., Racusin, J., Sekiguchi, M., Schneider, D. P., Shimasaku, K., & York, D. 2001. Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data. *AJ*, **122**(Oct.), 1861–1874.

Thacker, R. J., & Couchman, H. M. P. 2001. Star Formation, Supernova Feedback, and the Angular Momentum Problem in Numerical Cold Dark Matter Cosmogony: Halfway There? *ApJ*, **555**(July), L17–L20.

Toomre, A., & Toomre, J. 1972. Galactic Bridges and Tails. *ApJ*, **178**(Dec.), 623–666.

Trujillo, I., & Aguerri, J. A. L. 2004. Quantitative morphological analysis of the Hubble Deep Field North and Hubble Deep Field South – I. Early- and late-type luminosity-size relations of galaxies out to z˜ 1. *MNRAS*, **355**(Nov.), 82–96.

Trujillo, I., & Pohlen, M. 2005. Stellar Disk Truncations at High z: Probing Inside-Out Galaxy Formation. *ApJ*, **630**(Sept.), L17–L20.

Trujillo, I., Graham, A. W., & Caon, N. 2001. On the estimation of galaxy structural parameters: the Sérsic model. *MNRAS*, **326**(Sept.), 869–876.

Trujillo, I., Rudnick, G., Rix, H.-W., Labbé, I., Franx, M., Daddi, E., van Dokkum, P. G., Förster Schreiber, N. M., Kuijken, K., Moorwood, A., Röttgering, H., van de Wel, A., van der Werf, P., & van Starkenburg, L. 2004. The Luminosity-Size and Mass-Size Relations of Galaxies out to z ˜ 3. *ApJ*, **604**(Apr.), 521–533.

Trujillo, I., Förster Schreiber, N. M., Rudnick, G., Barden, M., Franx, M., Rix, H.-W., Caldwell, J. A. R., McIntosh, D. H., Toft, S., Häussler, B., Zirm, A., van Dokkum, P. G., Labbé, I., Moorwood, A., Röttgering, H., van der Wel, A., van der Werf, P., & van Starkenburg, L. 2006. The Size Evolution of Galaxies since z˜3: Combining SDSS, GEMS, and FIRES. *ApJ*, **650**(Oct.), 18–41.

van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 1998. Luminosity Evolution of Early-Type Galaxies to Z = 0.83: Constraints on Formation Epoch and Omega. *ApJ*, **504**(Sept.), L17+.

Vitvitska, M., Klypin, A. A., Kravtsov, A. V., Wechsler, R. H., Primack, J. R., & Bullock, J. S. 2002. The Origin of Angular Momentum in Dark Matter Halos. *ApJ*, **581**(Dec.), 799–809.

Vogt, N. P., Forbes, D. A., Phillips, A. C., Gronwall, C., Faber, S. M., Illingworth, G. D., & Koo, D. C. 1996. Optical Rotation Curves of Distant Field Galaxies: Keck Results at Reshifts to Z approximately 1. *ApJ*, **465**(July), L15+.

Weil, M. L., Eke, V. R., & Efstathiou, G. 1998. The formation of disc galaxies. *MNRAS*, **300**(Nov.), 773–789.

White, S. D. M., & Rees, M. J. 1978. Core condensation in heavy halos – A two-stage theory for galaxy formation and clustering. *MNRAS*, **183**(May), 341–358.

Williams, R. E., Blacker, B., Dickinson, M., Dixon, W. V. D., Ferguson, H. C., Fruchter, A. S., Giavalisco, M., Gilliland, R. L., Heyer, I., Katsanis, R., Levay, Z., Lucas, R. A., McElroy, D. B., Petro, L., Postman, M., Adorf, H.-M., & Hook, R. 1996. The Hubble Deep Field: Observations, Data Reduction, and Galaxy Photometry. *AJ*, **112**(Oct.), 1335–+.

Wolf, C., Dye, S., Kleinheinrich, M., Meisenheimer, K., Rix, H.-W., & Wisotzki, L. 2001. Deep BVR photometry of the Chandra Deep Field South from the COMBO-17 survey. *A&A*, **377**(Oct.), 442–449.

Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., & Kleinheinrich, M. 2003a. The COMBO-17 survey: Evolution of the galaxy luminosity function from 25 000 galaxies with 0.2< z <1.2. *A&A*, **401**(Apr.), 73–98.

Wolf, C., Wisotzki, L., Borch, A., Dye, S., Kleinheinrich, M., & Meisenheimer, K. 2003b. The evolution of faint AGN between z≈ 1 and z≈5 from the COMBO-17 survey. *A&A*, **408**(Sept.), 499–514.

Wolf, C., Meisenheimer, K., Kleinheinrich, M., Borch, A., Dye, S., Gray, M., Wisotzki, L., Bell, E. F., Rix, H.-W., Cimatti, A., Hasinger, G., & Szokoly, G. 2004. A catalogue of the Chandra Deep Field South with multi-colour classification and photometric redshifts from COMBO-17. *A&A*, **421**(July), 913–936.

Wolf, C., Bell, E. F., McIntosh, D. H., Rix, H.-W., Barden, M., Beckwith, S. V. W., Borch, A., Caldwell, J. A. R., Häussler, B., Heymans, C., Jahnke, K., Jogee, S., Meisenheimer, K., Peng, C. Y., Sánchez, S. F., Somerville, R. S., & Wisotzki, L. 2005. GEMS: Which Galaxies Dominate the z˜0.7 Ultraviolet Luminosity Density? *ApJ*, **630**(Sept.), 771–783.

York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., Castander, F. J., Chen, B., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B. R., Evans, M. L., Fan, X., Federwitz, G. R., Fiscelli, L., Friedman, S., Frieman, J. A., Fukugita, M., Gillespie, B., Gunn, J. E., Gurbani, V. K., de Haas, E., Haldeman, M., Harris, F. H., Hayes, J., Heckman, T. M., Hennessy, G. S., Hindsley, R. B., Holm, S., Holmgren, D. J., Huang, C.-h., Hull, C., Husby, D., Ichikawa, S.-I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R. S. J., Kinney, E., Klaene, M., Kleinman, A. N., Kleinman, S., Knapp, G. R., Korienek, J., Kron, R. G., Kunszt, P. Z., Lamb, D. Q., Lee, B., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R. H., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., Meiksin, A., Merelli, A., Monet, D. G., Munn, J. A., Narayanan, V. K., Nash, T., Neilsen, E., Neswold, R., Newberg, H. J., Nichol, R. C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. L., Petravick, D., Pier, J. R., Pope, A., Pordes, R., Prosapio, A., Rechenmacher, R., Quinn, T. R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Ruthmansdorfer, K., Sandford, D., Schlegel, D. J., Schneider, D. P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W. A., Smee, S., Smith, J. A., Snedden, S., Stone, R., Stoughton, C., Strauss, M. A., Stubbs, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Szokoly, G. P., Thakar, A. R., Tremonti, C., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Wang, S.-i., Watanabe, M., Weinberg, D. H., Yanny, B., & Yasuda, N. 2000. The Sloan Digital Sky Survey: Technical Summary. *AJ*, **120**(Sept.), 1579–1587.

Acknowledgement

Here, unfortunately well hidden on the **very** last page of my thesis, I would like to take the opportunity to thank many many people for indeed very different things that they did for me. I am absolutely sure that I will forget many people, I should thank. To those: Please feel greeted and thanked at this place.

First of all, I would like to thank Hans-Walter for getting me into the GEMS project. It was a great opportunity for me to enter astronomy, to be part of a great collaboration and to do some very interesting work with incredible data. At this place also thanks to all the GEMS-Team. I really enjoyed working in this 'small but fine' collaboration. You all taught me a lot. Thanks also to the very new STAGES-Team for letting me be part of this new collaboration. I hope, I can continue working with you all.

I would like to thank Immo Appenzeller for agreeing to be my second referee and to take part in my thesis defense. I really appreciate this, having already had my diploma thesis read by you.

My probably warmest thanks goes to Eric for his supervision and friendship. I am really grateful about the way that you supervised me. I've never had the feeling to have a 'supervisor', I always experienced you as a friend. Many thanks for guiding me through the years, for answering lots of (sometimes stupid) questions, for proofreading, for being there when I had problems (technical, scientific and personal) and last not least (to not make this an endless list) for cheering me up when it was necessary in between. Greetings to Amy, Glenn and Gloria.

Thanks to Knud Jahnke and Micaela Stumpf for reading parts of my thesis and to Andreas Müller for figure 1.7, that was specially created for this thesis in both content and appearance.

Thanks to everyone at the MPIA, I enjoyed my stay here very much and working with all of you was fun, even, and especially, when it was not exactly scientific, I am thinking about the 'open day' or the busy (indeed **very** busy) 'Lange Nacht der Museen', which ended only early in the morning. I liked these events. Thanks also to Götz Hoeppe for organizing 'Das halbe Universum unter dem Odeonsplatz'. It's still a very fascinating project that we had, and I wished we'd have had more of these ideas. Thanks to the computer department at the MPIA, especially Uli Hiller and Frank Richter, who were always glad to help (well, I don't know whether they were glad to, but they did always help quickly). Thanks also to the soccer and the basketball teams.

Thanks to all the students at MPIA. We have had so many events together and I expe-

rienced a nice atmosphere up on the Königstuhl. I especially think of the great students workshops on Rügen, in Oberau and Brixlegg, to my opinion the best workshops I have been to so far, not only because they were the most fun, but also because I think I learned the most there. At this place, thanks to the Patzer-foundation for supporting these events financially. Special thanks to the tea-Team (ok, or coffee-team) that helped me a lot on bad days with 30 minutes of fun every single day. I will miss that when I'm gone. On behalf of everyone (as I am sure I would forget too many people here), thanks to Sebastiano Ligori who, to best of my knowledge, was one of the people starting the daily coffee-break after lunch. Thanks to the 'students a-capella choir', Mica, Bernhard, Nati, Alex, Sascha, Katharina and Jutta. Thanks to all the people in room 216, I was very glad that I was one of the few lucky people to be 'allowed' to work in this beautiful office, we've had a great time there. Thanks to Johnny for the PSF lunch on wednesdays, it rescued my lunch very often.

Very special thanks to Sascha, Stefan Hanke and Nadine for listening, helping me, cheering me up (or trying to) when I had bad times.

Thanks to my family and all my friends for being there and for making me the person that I am. You made this thesis possible, although you might not know this. Thanks to everybody who was part of my life in the last years, although they might not know.

Thanks to Jürgen Klinsmann and the german soccer team 2006. You made this year very special and changed germany in a very positive way. I enjoyed the world championship very much as it was a really special atmosphere. Please always remember:

Im Prinzip bewahrt ihn nur jemand 4 Jahre für uns auf!

I cannot end this thesis without thanking and hugging Natascha heartily for being my best friend and helping me in countless situations when I needed it. It's great to have friends like you, although, as I know, this is not always easy. Without you, for sure this thesis would look **very** different.

THANK YOU

Erklärung

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den ____________________ ________________________

(Boris Häußler)