The Faint End of the Luminosity Function in the Core of the Coma Cluster

by
Margaret Louise Milne
B.Sc. University of Waterloo 2000
B.Ed. Queen's University 2000

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Physics and Astronomy

Supervisor: Dr. C. J. Pritchet

Abstract

We present optical measurements of the faint end of the luminosity function in the core of the Coma cluster. Dwarf galaxies are detected down to a limiting magnitude of $m_R = 25.75$ in images taken with the Hubble Space Telescope. This represents the faintest determination of the Coma luminosity function to date. Evidence is found for a steep faint end slope with $\alpha \sim -2$. Such a value is expected in theories in which reionization and other feedback effects inhibit dwarf galaxy formation in low density regions, while not affecting higher density regions.

Contents

List of Tables

List of Figures

Glossary

2-D: Two-Dimensional.

3-D: Three-Dimensional.

Å: Angstrom.

arcmin: Arcminute.

arcsec: Arcsecond.

CCD: Charge Coupled Device. A photometric detector, like that used in commercial digital cameras.

CDM: Cold Dark Matter.

CFHT: Canada France Hawai'i Telescope.

deg: Degree.

DN: Data Number. Also sometimes referred to as an "ADU", for Analogue to Digital Unit.

e^-: Electron.

EDCC: Edinburgh-Durham Cluster Catalogue.

F606W: WFPC2 filter, centred on 8269 Å with a width of 1758.0 Å.

F814W: WFPC2 filter, centred on 8269 Å with a width of 1758.0 Å.

FWHM: Full Width Half Maximum.

HDF: Hubble Deep Field.

HST: Hubble Space Telescope.

IC: Index Catalogue.

IRAF: Image Reduction and Analysis Facility. The most commonly used image analysis program in optical astronomy.

kpc: Kiloparsec. A parsec is a measure of distance, equal to 3.262 lightyears

M/L_B: Mass to light ratio in the B band.

$M_\odot/L_\odot$: Solar mass to light ratio.

mag: Magnitude.

Mpc: Megaparsec. See kpc.

NGC: New General Catalogue.

PC1: Planetary Camera, refers to the high resolution chip of the WFPC2.

pix: Pixel.

STScI: Space Telescope Science Institute.

WF: Wide Field, refers to the three lower resolution chips of the WFPC2.

WFPC2: Wide Field Planetary Camera 2. An optical camera on HST.

Acknowledgments

This work was supported in part by a Post-Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada.

I wish to formally thank my supervisor Dr. C. J. Pritchet. Without his support, both financial and scientific, this thesis would not have been possible. Acknowledgments must also be made to G. B. Poole, for sharing his results from early work on this project; Dr. J. J. Kavelaars, for advice regarding the Coma field data; and Dr. S. D. J. Gwyn, for many helpful discussions and Perl scripts.

A number of people also deserve my personal and heartfelt thanks: my friends and especially my family back home, whose faith and love kept me going; Peter, whose support got me out here in the first place; Chris, for his brilliance, tolerance, patience and extreme modesty – but most importantly, for his understanding as I spent research time on public outreach; the staff and manager(s) of the Centre of the Universe, for *their* understanding as I spent public outreach time on research; the Department of Physics and Astronomy secretaries, for their endless knowledge and timely reminders; the English Chamber Orchestra under the direction of Raymond Leppard, whose recording of Bach's Brandenburg Concertos was the soundtrack for much of the writing of this thesis; the other grad students, old, new and Newf, for making it fun; and finally Stephen, for more than space or professionalism will allow me to relate.

Chapter 1

Introduction

Imagine being faced with a grouping of galaxies and no data but a list of their brightnesses. How might one go about analyzing this structure? After exhausting such trifles as determining the brightest galaxy, the faintest galaxy, and the average galaxy brightness, the next obvious step would be to make a histogram. How many galaxies are there in each small range of brightness? What is the distribution of the brightnesses in this galaxy grouping?

This simple statistic – the number of galaxies per unit magnitude[1] per unit area on the sky – is known as a galaxy luminosity function[2]. Although it may appear almost too trivial to be of concern, the luminosity function is in fact a subtle and powerful tool. Its form is closely linked to that of the galaxy *mass* function, the number of galaxies per unit mass per unit volume or area. In standard models of structure formation, today's mass function is

[1] "Magnitude" is a measure of an object's brightness. The magnitude scale is reversed – brighter objects have smaller magnitudes.

[2] More precisely, this is a luminosity *distribution*. A true luminosity function would be measured as the number of galaxies per unit magnitude per unit *volume*. These differences are discussed further in Section 6.2.1; in this work, we follow the convention generally employed in the literature and refer to luminosity distributions as luminosity functions.

directly connected to infinitesimal perturbations in the initial density field of the universe. Therefore, the luminosity function offers a direct observational probe into the fundamental structure of the universe and the conditions of its early history.

The true power of the luminosity function, however, is that it traces not mass but light. Light comes from stars, stars form from gas, and that gas is subject to many processes beyond simple gravity: it can be cooled and heated; polluted and stripped; compressed and expanded. Once stars have formed, they in turn will evolve and die and affect the remaining gas and future generations of stars. Galaxy formation and evolution is a complex and involved process, and its results are summed up in the luminosity function. Any theory attempting to explain how galaxies form and evolve must test its predictions against the observed shape of the luminosity function.

The faint end of the luminosity function is of particular interest. Faint galaxies are generally small galaxies, and so the faint end of the luminosity function probes the smallest, least massive galaxies in the universe. According to the current cold dark matter (CDM) hierarchical clustering model of galaxy formation, all structure in the universe was built up from these small galaxy building blocks. Many small, faint galaxies should still remain today – observationally, the slope of the faint end of the luminosity function should be steep. Studying the faint end of the galaxy luminosity function is an excellent method of testing the cold dark matter paradigm.

Of course, the faint end of the luminosity function is arguably the most difficult part to determine. Although small galaxies are numerous, their faintness makes them hard to detect. It can also be very difficult to determine

the distance, and hence true intrinsic magnitude, of a faint galaxy. The standard technique is to examine a galaxy's spectrum, searching for emission or absorption lines that have been shifted from their characteristic wavelengths due to the expansion of the universe. This, however, requires the light of the galaxy to be separated into a series of wavelength bins so as to construct the spectrum. Faint galaxies simply do not give enough light to produce good spectra.

One way to avoid the task of determining distances for all the galaxies in a luminosity function is to create a luminosity function from galaxies that are all at the same distance. This is the case when the luminosity function of a cluster or group is studied; the galaxies in the group or cluster all lie at approximately the same distance, and so their relative apparent magnitudes are the same as their relative absolute magnitudes. The problem, however, has merely been shifted, not solved. When determining the luminosity function of a cluster or group of galaxies, the important thing is to ensure that no objects in the background (or foreground) of the cluster are included in the luminosity function. It would appear that the distance to each object would still be needed, to make sure that only galaxies at the distance of the group or cluster are used.

There is a way around this problem. Consider what is obtained when all the galaxies in an image of a cluster are used to construct a histogram of magnitudes. The luminosity function – the histogram of the cluster galaxies' magnitudes – is present, but contaminated in each bin by galaxies from the background. How many background galaxies are in each bin? Make a histogram of only background galaxies to find out – image a blank piece

of the sky, free from any known galaxy groups or clusters, and construct a histogram from all the galaxies in that field. Subtract this background-only histogram from the cluster plus background histogram, and the result is a histogram with only cluster galaxies – a luminosity function.

Figure 1.1 shows a schematic overview of this method of "statistical background subtraction" – statistical, because no information is obtained about whether or not any one particular galaxy belongs to the cluster. However, statistically, the background galaxies can still be removed. The method, pioneered by Zwicky (1957), is an enormously powerful tool for determining luminosity functions in rich clusters. Spectra for each galaxy – time consuming to obtain for bright galaxies, next to impossible to obtain for faint ones – are not needed, and so luminosity functions of large areas and to deep magnitudes can be made in a reasonable amount of observing time.

Statistical background subtraction has its drawbacks, however. First of all, it can only be used where there is a high enough concentration of galaxies to create a good statistical contrast against the background – typically, only in clusters of galaxies. Second, and most important, it is vitally important that the detection characteristics of the control image be matched as closely as possible to the detection characteristics of the image containing the cluster. Any slight mismatch in the ease or difficulty with which galaxies are detected, or in the magnitudes that are found for them, can shift one histogram with respect to the other and so render the subtraction meaningless. The advent of digital detectors, such as CCDs, has made matching detection characteristics much simpler, but extreme caution must still be used.

However the faint end of the luminosity function is determined, the steep-

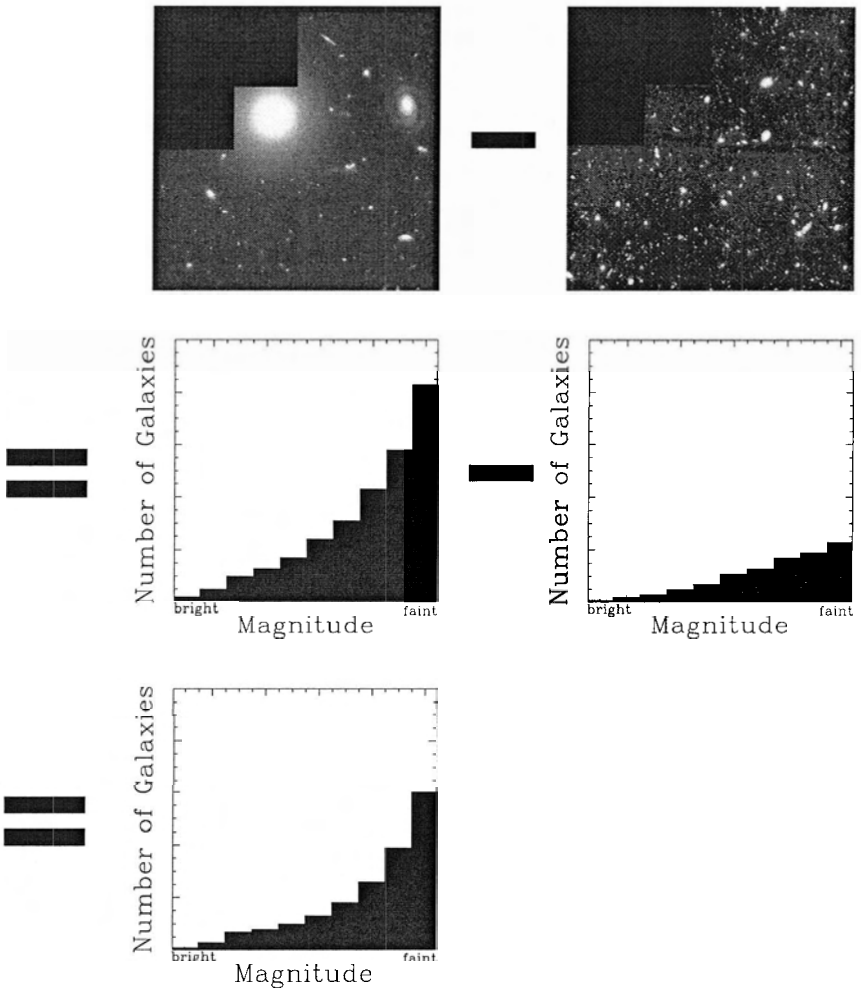

Figure 1.1: A schematic overview of how to use statistical background subtraction to find the luminosity function of a cluster. Two images are taken: one of the cluster (plus the inevitable background objects), and one of a control field containing only background objects. A histogram is made of the number of objects in each magnitude bin for each image. When the background-only histogram is subtracted from the cluster+background histogram, the result is a cluster-only histogram – a luminosity function.

ness of its slope is the primary interest. As mentioned earlier, CDM models predict the slope should be steep. Although this has been seen for some clusters, most studies have found flatter faint end slopes than CDM predicts. Does a true steep faint end slope exist anywhere in the universe? There is accumulating evidence that the slope of the faint end of the luminosity function increases in environments of increasing density (see, e.g., Trentham & Hodgkin, 2002). Therefore, a good place to look for a steep faint end slope would be a region of extremely high galactic density.

The Coma cluster is just such a region. Nearby ($z \sim 0.023$, equivalent to a distance of ~ 108 Mpc for $H_o = 65$ km/s/Mpc), rich in galaxies (Abell class 2), and in a region of the sky far from the light-blocking dust of our galactic plane ($b_{II} = 88°$), the Coma cluster is one of the most studied clusters in the sky. Many attempts have been made at determining the slope of the faint end of the luminosity function in Coma, using various combinations of cameras and telescopes, looking in several bands of the electromagnetic spectrum, and covering different areas around the cluster core. A wide range of slopes have been reported, ranging from values approaching that required by CDM, right down to values more like those found in the galaxy-poor field.

In this study, we propose to determine the slope of the faint end of the luminosity function in the core of the Coma cluster. Using the method of statistical background subtraction, we will obtain both our Coma cluster images and our control field images from Hubble Space Telescope data. This marks the first time that space-based optical data has been used to study Coma's luminosity function. Using this data, we will be able to study the luminosity function to fainter magnitudes and with greater resolution than

any previous survey. We hope to determine whether or not the Coma cluster marks an environment where the CDM prediction of steep faint end slopes will hold.

In Chapter Two, we detail the observations we obtained from the Hubble Space Telescope archive and discuss the image processing and data reduction steps that we carried out. The process by which the detection characteristics of the control image were matched to those of the cluster image are also described in full.

In Chapter Three, we begin by highlighting some of the more subtle aspects of SExtractor, the detection and photometry software used in this work. We then discuss the "add-galaxy" experiments used to determine the best procedures and parameters for detection and photometry, and the tests that were carried out to ensure the match in detection characteristics between the data and control images. Finally, we explain how the initial catalogues were created and compared to published catalogues of the same fields.

In Chapter Four, we describe the creation of globular cluster masks, used to prevent globular clusters and their blends from contaminating our galaxy catalogues. We go through the steps of creating the final catalogue, including detection and photometry, omitting objects from suspect regions, calculating colours, separating stars from galaxies, and transforming magnitudes into a standard system. We end by comparing our catalogue of faint Coma cluster galaxies to the dwarf galaxies found in other surveys.

In Chapter Five, the luminosity function is presented. We start by determining the limiting magnitude of our survey through two independent methods. We estimate the error in our counts due to cosmic variance, and

finally construct the luminosity function. A parametric fit is applied to the luminosity function to determine the faint end slope.

In Chapter Six, we discuss our results, beginning by comparing them to other studies of the Coma cluster luminosity function. We highlight some possible systematic explanations for our faint end slope and detail how such arguments do not apply in this case. We then discuss the implications of our result.

Finally, in Chapter Seven, we provide a summary of this work. We end by suggesting future directions for the study of Coma's luminosity function, and luminosity function research in general.

Chapter 2

Observations And Data Reduction

2.1 Coma Cluster

2.1.1 Observations

Our Coma cluster observations consist of data obtained from the Hubble Space Telescope archive. On 16 and 24 August 1997, F606W and F814W images were taken with the WFPC2 camera[1], with the centre of the PC1 CCD placed on the nucleus of one of Coma's central galaxies, NGC 4874 (program GO-5905). The data from this program used in this study consists of 16 F606W exposures, totaling 20 400 s, and 6 F814W exposures totaling 7800 s. For details on all the exposures from this program, please refer

[1]The WFPC2 Camera is a four CCD mosaic camera on HST. Three chips are "Wide Field" (WF) chips, measuring 800 × 800 pix with a resolution of $\sim 0.1''$/pix. The fourth chip is the Planetary Camera (PC1) chip, also measuring 800 × 800 pix but with a resolution of $\sim 0.046''$/pix. F606W and F814W are wideband filters that can be used with the WFPC2 Camera. F606W is centred on 5843 Å with a width of 1578.8 Å (approximately equivalent to the Johnson-Cousins V or R filter), and F814W is centred on 8269 Å with a width of 1758.0 Å (approximately equivalent to the Johnson-Cousins I filter).

to Table 2.1. To aid in the eventual removal of cosmic rays and bad-pixel artifacts, the long exposures in each filter were dithered by fractional pixel shifts in a pentagonal pattern. Due to light contamination from the nucleus of NGC 4874, the PC1 chip was not used in this work. Therefore, the total area (before trimming) covered by these observations equals 19 200 arcsec^2 on the sky.

Table 2.1: HST Observing Log for GO-5905[a]

R.A. (J2000)	Decl. (J2000)	Filter	Exposure time (seconds)
$12^h59^m33^s.43$	$+27°57'43''.3$	F606W	3×180 [b]
			2×1100
			14×1300
		F814W	4×230 [b]
			6×1300

[a] Taken from Kavelaars et al. (2000)

[b] While part of program GO-5905, these exposures were not used in this project.

The raw data was processed with the standard HST pre-processing pipeline. The next two sections describe in detail the additional processing steps that we applied to the data after retrieving it from the archive. These steps are: registering the frames to the same coordinate system; removing cosmic rays; coadding the frames; modelling and subtracting large elliptical galaxies; trimming the images; and creating and subtracting a model of the background light.

2.1.2 Registration And Coadding

The NGC 4874 images used in this project were registered and coadded by G. B. Poole. The following steps were performed on the data.

Approximately 20 stellar objects were chosen on each chip. One object, reasonably bright and free from crowding from other objects or cosmic rays, was chosen from this sample. The xy coordinates of this object on the first (i.e., reference) frame were compared with its xy coordinates on the other frames to determine an initial guess at the shifts between the frames.

Each frame in turn was then fed into the IRAF task `imcentroid`, along with the initial guess at that frame's shift, the reference frame, and the list of the xy coordinates of the other stellar objects on the reference frame. The task `imcentroid` then found the shift of each object relative to its position on the reference frame using a "marginal" centroid algorithm. The IRAF defaults for the centering algorithm parameters were used, with the exception of the parameter `boxsize`. This is the size in pixels of the box that is used in the final centering. A size of 5 pix was chosen for this parameter.

For each frame, the average of the approximately 20 shifts was taken as the true shift between the frames. The IRAF task `blkrep` was then used to block replicate each frame by a factor of 7. The block-replicated image was divided by the square of the block replication factor (i.e., 49) to normalize the flux. The IRAF task `imshift` was then used to shift each image (except the reference image) by its average shift times the block replication factor.

After all frames had been shifted, they were coadded using the IRAF task `imcombine`. This task first scaled each frame by multiplying it by the

reciprocal of its exposure time (as recorded in the image header). Cosmic rays were then removed from each frame by setting the `reject` parameter to `crreject`. For each pixel, this algorithm takes the median or unweighted average (excluding the minimum and maximum value) of the pixel values from each frame. The expected standard deviation of these values is calculated using the CCD noise parameters that were input to the algorithm (in this case, we specified a readnoise of 5.32 e^- and a gain of 7.0 e^-/DN). Any pixel that deviates more than 3 times this standard deviation from the median is then rejected. The process is repeated until no more pixels are rejected.

After cosmic rays were rejected, the images were combined using the median pixel value. This resulted in a final coadded image with an exposure time of 1300 s.

Finally, the coadded image pixels were returned to their normal size by using the IRAF task `blkavg` to block average the image by a factor of 7. The image was multiplied by the square of the block average factor (i.e., 49) to renormalize the flux.

2.1.3 Background Subtraction And Trimming

The first step in subtracting the background light from the NGC 4874 images was to fit and remove the large elliptical galaxies in the images. Two large elliptical galaxies were identified on each WF chip. The xy coordinates of their approximate centres were fed into the IRAF task `ellipse`. This task was used to interactively fit isophotes to the elliptical galaxies.

These isophotes were then fed into the IRAF task `bmodel`. For each galaxy, `bmodel` created an image with zero background containing a noiseless

photometric model of the galaxy. The fitting algorithm in `bmodel` does not extrapolate the galaxy model out to the point where its signal drops to zero; therefore, the models created in `bmodel` have a discontinuous step at their boundaries. The amount of this step was determined and `bmodel` was rerun with the `background` parameter set to the amount of the step. This created an image where every background pixel had the value of the step, and so the galaxy model had no discontinuity at its boundary. The value of the step was then subtracted from every pixel on the image to bring the background level back to zero. This, of course, slightly changes the shape of the galaxy model. However, as Figure 2.1 shows, this procedure was found to result in better background subtraction results, without edge effects around the bright ellipticals.

Figure 2.1 also shows that the process of modelling and subtracting the elliptical galaxies does not cleanly remove all the light from these galaxies. Specifically, strange residuals were left at the centre of each subtracted galaxy. Detections in these regions obviously cannot be trusted and must be omitted from the catalogue. This process is discussed briefly in Section 3.2.1, and in more detail in Section 4.2.3.

After models were made for both bright galaxies on a chip, the two model images were added together to create a master image of the large ellipticals on that chip. This image was subtracted from the coadded image to create an image with the large ellipticals removed.

The elliptical-subtracted images were then trimmed. On each of the images, the region free from edge defects was determined by eye. The regions for each chip as determined from the F606W and from the F814W images

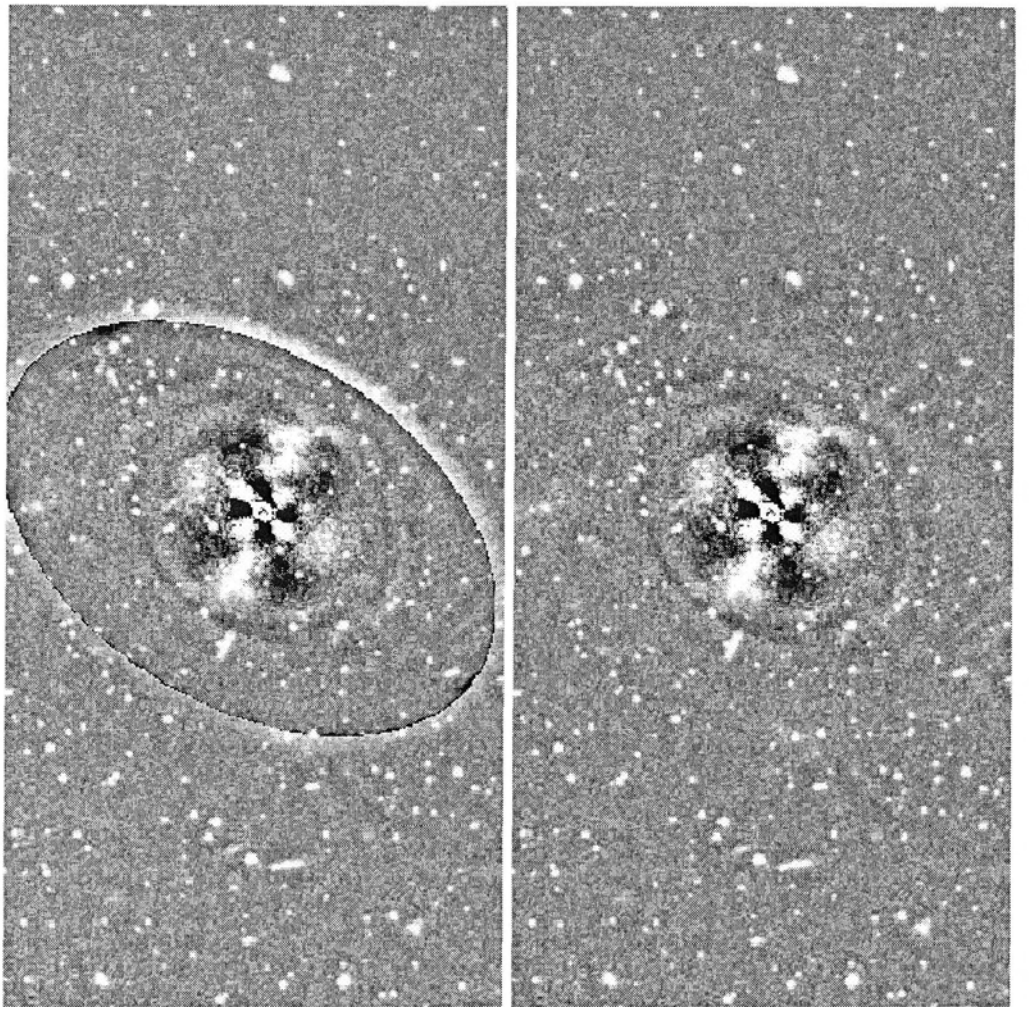

Figure 2.1: The results of fitting and subtracting one of the large elliptical galaxies in the Coma field. In the left panel, a model with a discontinuous step was subtracted. In the right panel, the discontinuous step was removed from the model before subtracting.

were compared, and the smaller of the two regions adopted for that chip, to ensure the images in both filters would be of equal size and both free from edge defects. The IRAF task `imcopy` was then used to trim the images to the determined regions. The final dimensions of the trimmed images are given in Table 2.2. The master ellipse images for each chip were also trimmed to match the corresponding data image.

Table 2.2: Dimensions of the coadded and trimmed NGC 4874 images

Chip	Dimensions (pix)
2	741 × 757
3	756 × 736
4	739 × 741

The elliptical-subtracted and trimmed images were then ring median filtered using the IRAF task `rmedian`. This algorithm slides a circular annulus over each pixel in the image. The centre pixel of the circle is replaced by the median of the pixels in the annulus. The result is that objects with scale-length equal to the inner radius of the annulus are removed and replaced with an estimate of the local background value. The radius of the inner annulus was therefore set to 7 pix to match the typical object diameter on the chips. The radius of the outer annulus was set to 10 pix to be just smaller than the typical separation of objects on the chips.

The resulting map of the background light on each chip was subtracted from the elliptical-subtracted and trimmed image of that chip to create a final background-subtracted image. The background light map was also added to

the trimmed master image of the ellipticals on that chip to create a final model of the background.

2.2 Control Field

2.2.1 Observations

For control field data, we downloaded F606W and F814W images of the Hubble Deep Field North (HDF) from the STScI website[2]. The HDF was a Director's Discretionary program on HST in Cycle 5. Its goal was "to image a typical field at high galactic latitude in four wavelength passbands as deeply as reasonably possible" (from the website). From 18 - 30 December 1995, the field centred at 12h 36m 49.4s +62d 12m 58.0s (J2000) was imaged using the WFPC-2 camera. The data we used were the Version 2 final reduced images – registered, sky-subtracted, and normalized to an exposure time of one second. The number of frames used in these images and total exposure time in each filter can be found in Table 2.3. The exposures in each filter were dithered. Images at the different dither positions were then registered to a small fraction of an original pixel, corrected for geometric distortion, and "drizzled" onto a final image with a sampling of 0.04″/pix. As a result, the three WF chips effectively measure 2048 × 2048 pix, and so cover $\sim$ 20 133 arcsec^2 on the sky.

The next three sections describe the processing steps we used to match the detection characteristics of the HDF to those of the Coma image. These steps are: rebinning the HDF to degrade resolution; increasing the effective

[2]http://www.stsci.edu/ftp/science/hdf/hdf.html

Table 2.3: HDF Version 2 images [a]

Filter	Number of frames	Total exposure time (seconds)
F606W	103	109 050
F814W	58	123 600

[a] Taken from http://www.stsci.edu/ftp/science/hdf/hdf.html

exposure time of the coadded image; trimming the HDF to the area of the Coma field; modelling the noise of Coma and adding it to the HDF; and matching extinction.

2.2.2 Matching Resolution, Exposure Time And Size

In the method of differential counts, it is imperative to recreate the detection characteristics of the data field as closely as possible in the control field. It was therefore necessary to degrade the HDF. The first step was to match its resolution, exposure time and size to that of the Coma image.

As mentioned above, the HDF was drizzled to achieve a pixel scale of 0.04″/pix. The NGC 4874 images have the normal WFPC2 pixel scale of 0.1″/pix. To give the HDF the same resolution, the images had to be re-binned by a factor of 2.5. The IRAF task `magnify` was used to do this. The parameters `xmag` and `ymag` were set to -2.5, while `interpolant` was set to `drizzle` with the drizzle pixel fraction set to 1.0. These settings caused the task to sum the flux from each 2.5 × 2.5 pix block and use that value for the corresponding single pixel in a lower resolution image. The original HDF

chips were each 2048 × 2048 pix; after rebinning, the lower resolution HDF chips were 819 × 819 pix.

It is very important that this process of rebinning conserve flux. An object should have the same magnitude on both the high resolution and the low resolution image. To check this, aperture photometry was performed on the high resolution and the low resolution images for a variety of aperture radii: 0.2″, 0.4″, 0.6″, 0.8″ and 1.0″. These sizes corresponds to radii of 2, 4, 6, 8 and 10 pix for the low resolution image, and 5, 10, 15, 20 and 25 pix for the high resolution image. The catalogues were matched, and the aperture magnitudes obtained from the high resolution image and the low resolution image were compared. As can be seen in Figure 2.2, the magnitudes match very well at every aperture. The high resolution and the low resolution images produce the same magnitude for any given object; that is, flux has been conserved.

As mentioned above, the HDF Version 2 images are normalized to an exposure time of one second. To match the Coma images' effective exposure time of 1300 s, the IRAF task `imarith` was used to multiply each HDF image by 1300.

Finally, to match the Coma image size, each HDF chip was trimmed such that its dimensions matched that of the corresponding Coma chip. The trimming regions were chosen by eye to also exclude pixels affected by edge effects.

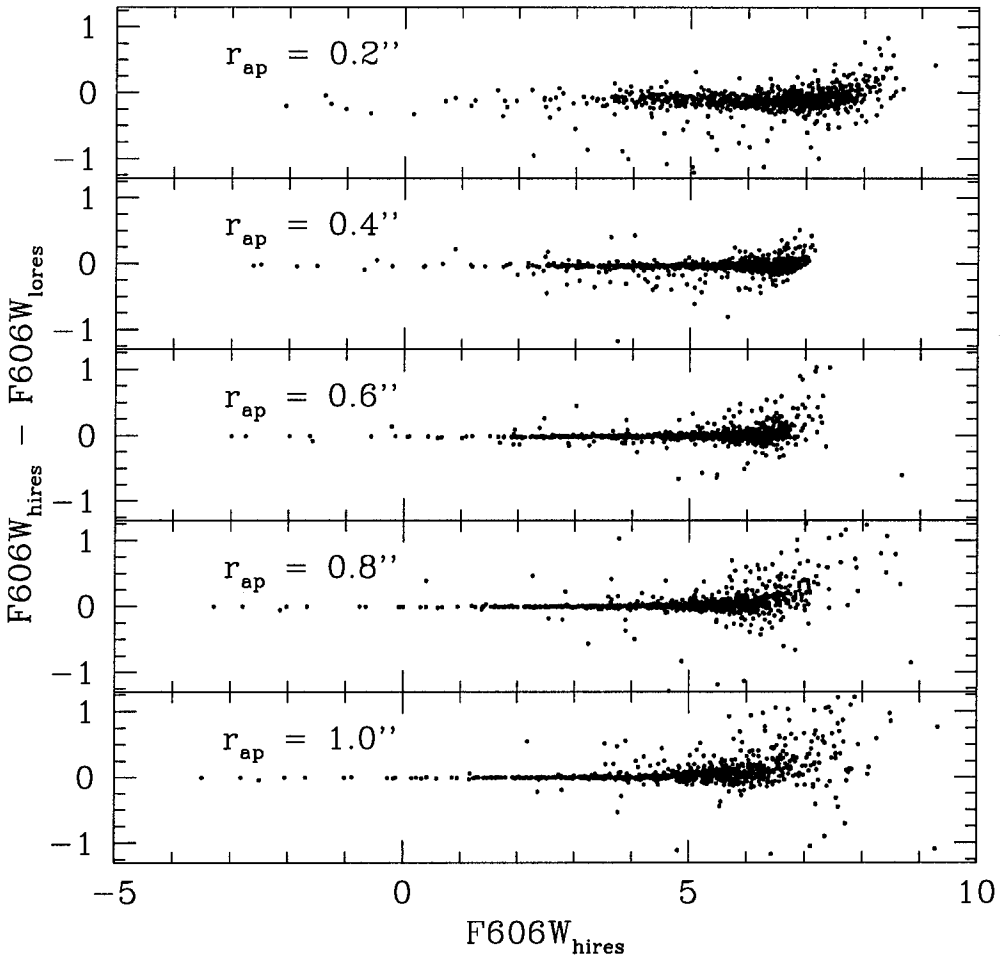

Figure 2.2: Difference between aperture magnitudes from the high resolution HDF image and aperture magnitudes from the low resolution (i.e., rebinned) HDF image versus aperture magnitudes on the high resolution image. Each plot shows the results of using a different sized aperture; from the top, the aperture radii were 0.2″, 0.4″, 0.6″, 0.8″ and 1.0″. All magnitudes are instrumental in the F606W band.

2.2.3 Matching Noise

The second important step in matching the detection characteristics of the HDF to those of Coma is to match the noise. The total exposure time of the HDF is much greater than that of Coma, and so the HDF is much less affected by Poisson noise. To see this, consider a pixel emitting n e^-/s. The flux of that pixel on each of the images, and the error in the flux, is then

$$F_{Coma} = 20400n \pm \sqrt{20400n} \quad (2.1)$$

$$F_{HDF} = 109050n \pm \sqrt{109050n} \quad (2.2)$$

for a Coma exposure time of 20 400 s and a HDF exposure time of 109 050 s. After normalizing to a one second exposure, the fluxes and errors become

$$F_{Coma} = n \pm \sqrt{\frac{n}{20400}} \quad (2.3)$$

$$F_{HDF} = n \pm \sqrt{\frac{n}{109050}} \quad (2.4)$$

The relative noise in the HDF as compared to the noise in Coma is then

$$\frac{\sqrt{\frac{n}{109050}}}{\sqrt{\frac{n}{20400}}} = 0.432 \quad (2.5)$$

For equal signals, the HDF will have less than half the noise of Coma. The result is the same if the two images are normalized to a 1300 s exposure, as was done in this work. If detection characteristics are to be matched, noise must be added to the HDF.

A standard technique for generating noise is to draw numbers at random from a Poisson distribution. In practice, this is usually done by using the Gaussian approximation to the Poisson distribution. To create noise for the

HDF that resembles the noise in Coma, one could select numbers from a Gaussian distribution centred on zero with standard deviation equal to the typical standard deviation of the Coma background.

The background noise in the Coma images is not very constant, however. The large elliptical galaxies on each chip and the light from NGC 4874 on the PC1 chip lead to varying background levels across each Coma chip. If noise for the HDF was generated using a single "typical" noise value, the detection characteristics of the two fields would not be very well matched. In particular, the effective area over which detections could be made would not be matched. Determining the proper scaling factor to account for this would be very difficult.

To avoid these problems, we chose to create a spatially dependent model of the noise in each Coma chip and superimpose these models on the HDF chips. A model was made of the Coma background on each chip as part of the background subtraction procedure. This can be used to generate a noise image: at each pixel of the noise image, the expected noise value of the corresponding background model pixel is calculated based on the background level at that pixel, the read noise, and the number of frames that were coadded to create the image. A value for the noise pixel is then drawn at random from a Gaussian distribution centred on zero with standard deviation equal to that expected noise value. Mathematically, we have

$$N_{i,j} = \frac{\sqrt{B_{i,j}g + r^2}}{\sqrt{n_{frames}}} \frac{G}{g} \tag{2.6}$$

where $N_{i,j}$ is the signal in DN of the (i,j)th pixel in the noise model, $B_{i,j}$ is the signal in DN of the (i,j)th pixel of the background model, g is the

gain, r is the readnoise, n_{frames} is the number of frames that went into the coadded Coma image, and G is the Gaussian deviate.

This method will produce a good model of the Coma noise if the Coma noise is truly "white" – that is, if the noise in the background in due entirely to Poisson variation. To check if this was the case, one of the background-subtracted Coma chips was compared to its background model. 165 positions were identified on the background-subtracted Coma image that were free from objects. In a 5 × 5 pix box at each of the 165 positions, the IRAF task `imexamine` was used to determine the standard deviation of the background-subtracted image and the mean of the background model. The mean was used to generate a measure of the expected Poisson noise, using the expression given in Equation 2.6 (omitting the multiplication by a deviate). The gain was set to 7 e^-/s, and the read noise to 5 e^-.

The actual standard deviation at each location was then compared to the modelled Poisson noise at that location. Figure 2.3a shows the results. As can be seen, the actual noise – the standard deviation of the background-subtracted image – is consistently less than would be predicted from a Poisson noise model based on the mean of the background model. This is explained by noting that the noise in the coadded Coma image cannot in fact be completely "white": the coadding process naturally produces correlations between the pixels, and so reduces the noise.

To add the proper amount of noise to the HDF image, it is necessary to somewhat reduce the noise generated from the background model. Multiplying the generated noise by 0.85 *a priori* produces a better match between the actual noise and the generated noise (Figure 2.3b). Therefore, the noise-

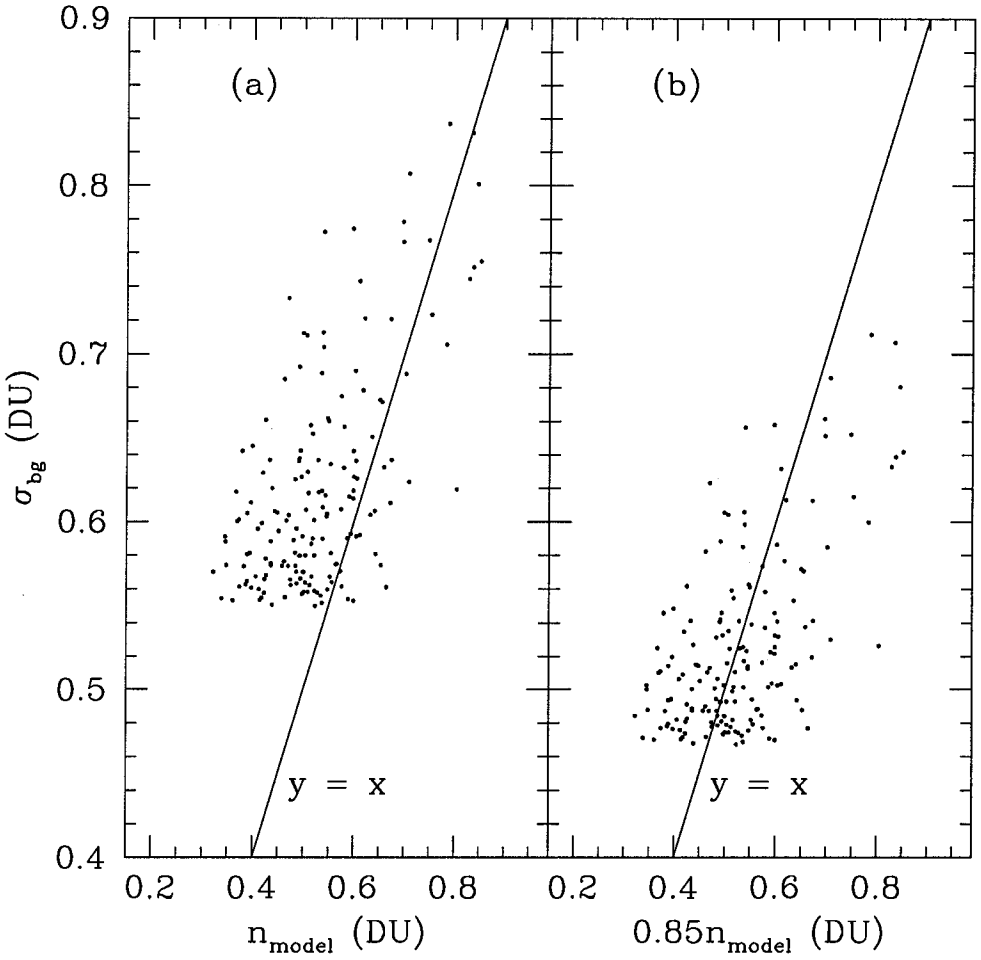

Figure 2.3: Standard deviation of background regions on the background-subtracted Coma image versus modelled Poisson noise for those same regions. In (a), Equation 2.6 is used to model the noise (omitting the deviate). In (b), 0.85 of the previous value is used (i.e., Equation 2.7, omitting the deviate).

generating formula given in Equation 2.6 is modified to

$$N_{i,j} = 0.85 \frac{\sqrt{B_{i,j}g + r^2}}{\sqrt{n_{frames}}} \frac{G}{g} \tag{2.7}$$

This formula was applied to the background model images of each of the Coma chips to generate noise model images. As a check, the noise model for one chip was compared to the background-subtracted image of that chip. The IRAF task `imexamine` was used to examine the standard deviation of the noise model and the background-subtracted image at the 165 background regions defined earlier. Figure 2.4 shows the results: the noise in the model is a good representation of the true noise in the image.

The noise models were then added to the corresponding rebinned, scaled and trimmed HDF chips. Normally, noise must be added in quadrature. However, using the results of Equation 2.5, we have

$$\begin{aligned} & \sqrt{noise^2_{Coma} + noise^2_{HDF}} \\ = \; & \sqrt{noise^2_{Coma} + 0.432\ noise^2_{Coma}} \\ = \; & 1.20\ noise_{Coma} \end{aligned} \tag{2.8}$$

The noise from the HDF only adds $\sim$ 20% to the total noise, and so can be neglected. The IRAF task `imarith` was used to add the noise model images directly to the HDF images. To check that the HDF noise levels were now similar to those in Coma, a background-subtracted Coma image was compared to a degraded HDF image. The IRAF task `imexamine` was used to find the standard deviation of the Coma image in the 165 regions defined earlier. A new set of 165 object-free 5 $\times$ 5 pix regions were defined on the

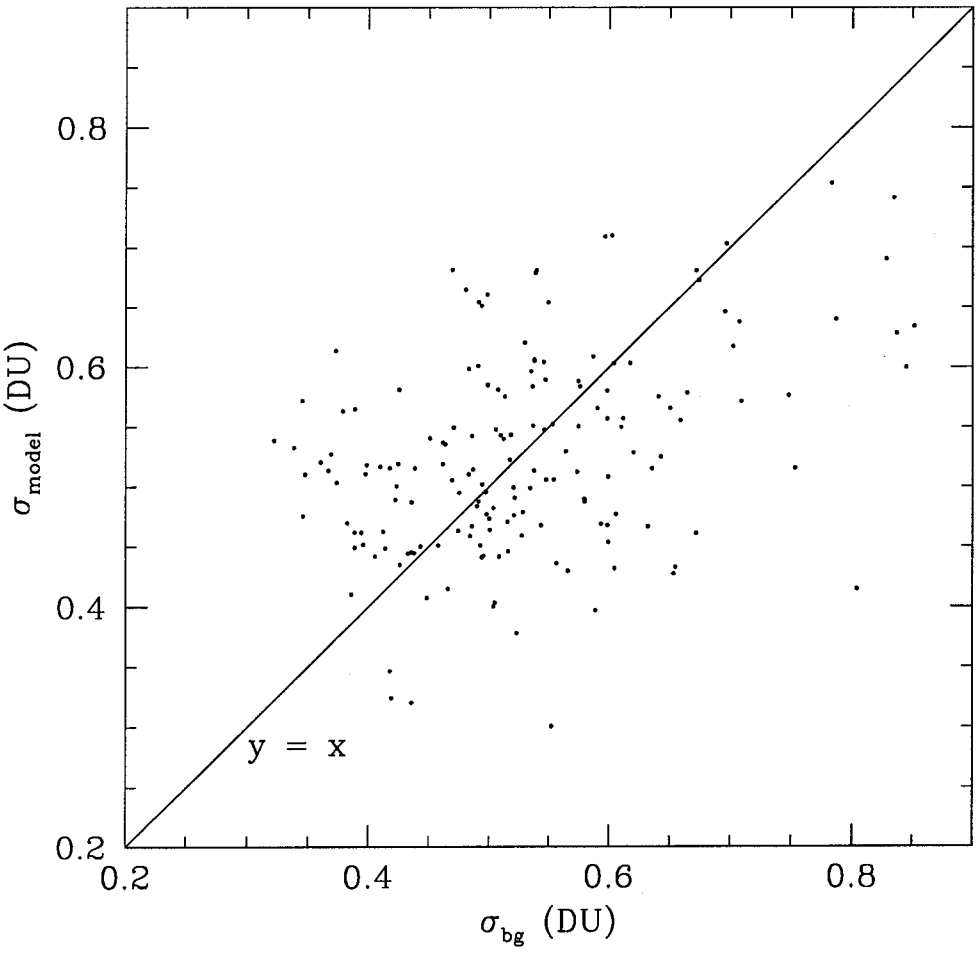

Figure 2.4: The standard deviation of the noise model versus the standard deviation of the Coma image background.

HDF image, and the standard deviation in these was also determined. The mean noise level and standard deviation in that mean are shown for Coma and the HDF in Table 2.4. As can be seen, the noise in the HDF is now a good match to the noise in Coma.

Table 2.4: Comparing noise in Coma and the degraded HDF

Field	Mean noise (DN)	Standard deviation (DN)
Coma	0.530	0.112
Degraded HDF	0.583	0.175

2.2.4 Matching Extinction

The final step in matching the detection characteristics of Coma and the HDF is to match extinction. Table 2.5 gives the coordinates, reddening, and extinction in F606W and F814W for Coma and the HDF. The reddening values are from Schlegel et al. (1998). They also give the conversion from reddening to extinction in F606W and F814W:

$$A_{F606W} = 2.889E(B-V) \quad (2.9)$$

$$A_{F814W} = 1.948E(B-V) \quad (2.10)$$

As can be seen in Table 2.5, the extinction in both bands for both Coma and the HDF is negligible. The extinction in Coma and the HDF can be considered matched.

Table 2.5: Extinction for Coma and the HDF

Field	R.A. (J2000)	Decl. (J2000)	E(B-V) (mag)	A_{F606W} (mag)	A_{F814W} (mag)
HDF	$12^h36^m49^s$	$+62°12'58''$	0.012	0.035	0.023
Coma	$12^h59^m48^s$	$27°58'48''$	0.008	0.023	0.016

Chapter 3

Initial Catalogue

This chapter and the one following discuss in great detail the processing steps that were followed to create galaxy catalogues. In this chapter, we focus on the preliminary steps leading to an initial catalogue. Specifically, we discuss:

- Important features of SExtractor, the detection and photometry package used
- "Add-galaxy" experiments, used to determine the best detection and photometry parameters
- Tests to compare the detection characteristics of the data field and the control field
- Comparisons between our initial catalogues and published catalogues of the same fields

A summary of Chapter 4 can be found at the beginning of that chapter. The reader who is less interested in such technical details may wish to skip directly to Chapter 5, where the luminosity function is presented.

3.1 The SExtractor Package

The SExtractor (**S**ource **Extractor**) package of Bertin & Arnouts (1996) was used for detection and photometry in this work. SExtractor takes an astronomical image and produces a catalogue of sources on that image, through a process of seven steps: estimation of the sky background; thresholding; deblending; filtering of detections; photometry; classification; and catalogue output. More information on the basic workings of SExtractor can be found in Bertin & Arnouts (1996) and the various SExtractor User Guides available on the Terapix website[1].

SExtractor has a variety of features and aspects that bear special mention. These are discussed in more detail below.

3.1.1 The Link Between Detection And Photometry

Through the testing and use of SExtractor during this work, an important fact came to light that must be stressed: *in SExtractor, detection and photometry are linked.*

Detection and photometry are of course always linked: photometry is only performed on objects that are detected. However, with the SExtractor package, the link is stronger. During the detection phase of the algorithm, object centres and isophotes are determined from the *convolved* image. Both the centres *and the isophotes* are then passed to the photometry portion of the program: the isophotes are used to determine the first light moment r_1, which is used to define the aperture for automatic aperture photometry

[1] http://terapix.iap.fr/rubrique_2.php?id_rubrique=91

(MAG_AUTO).

This means that the magnitudes of objects depend on the convolution kernel with which the image was filtered. A different convolution kernel will produce a different convolved image, which will lead to different isophotes, different r_1 values, different apertures, and so different magnitudes. This is an unexpected and counter-intuitive result, of which future users of SExtractor should be aware.

3.1.2 Convolution And The Detection Threshold

As stated in the SExtractor User's Guide, the user must enter the detection threshold in units of the background's standard deviation. A more subtle point is that the standard deviation indicated is that of the *unconvolved* image, whereas the detection of sources actually occurs on the *convolved* image.

This is important because convolving an image reduces the background noise in that image. Consider a kernel defined by a series of pixel weights,

$$\{w_1, w_2, w_3, ...w_n\} \tag{3.1}$$

where

$$\sum_{i=1}^{n} w_i = 1 \tag{3.2}$$

Applying this kernel to an image will replace each pixel in the image with the weighted sum of the pixels around it: w_1 of the first pixel's value, plus w_2 of the second pixel's value, and so on.

The noise N' in the convolved pixel will then be the weighted sum of the noise values in the surrounding pixels, added in quadrature. If N_i is the

noise in the i^{th} pixel, we have

$$N' = \sqrt{\sum_{i=1}^{n} (w_i \; N_i)^2} \tag{3.3}$$

Assuming the surrounding pixels contain only background light, the noise in each will be approximately equal:

$$N_1 \sim N_2 \sim ... \sim N_n \sim N \tag{3.4}$$

Therefore, Equation 3.3 becomes

$$\begin{aligned} N' &= \sqrt{\sum_{i=1}^{n} (w_i \; N)^2} \\ &= N\sqrt{S} \end{aligned} \tag{3.5}$$

where S is the sum of the squares of the kernel weights. Therefore, convolving with a kernel reduces the background noise by a factor of $\sqrt{S}$.

This means that if a detection threshold of 3.5σ is desired, the value that should be entered into SExtractor is actually $3.5\sqrt{S}$, where $\sqrt{S}$ is the noise reduction factor of the chosen convolution kernel.

The noise reduction factors of some common SExtractor convolution kernels are given in Table 3.1, along with the scaled 2.5σ, 3.5σ and 4.0σ thresholds for each kernel.

3.1.3 Weight Images And Thresholds

SExtractor allows the user to specify a weight image – an image the same size as the data image which describes the noise intensity at each pixel. SExtractor uses these weight images to adjust the detection threshold at each

Table 3.1: Noise reduction factors and scaled detection thresholds for SExtractor convolution kernels

Kernel	Noise reduction factor	2.5σ	3.5σ	4.0σ
gauss_1.5_3x3.conv	0.423	1.06	1.48	1.69
gauss_2.0_3x3.conv	0.366	0.92	1.28	1.46
gauss_2.0_5x5.conv	0.316	0.79	1.11	1.26
gauss_2.5_5x5.conv	0.266	0.66	0.93	1.06
gauss_3.0_5x5.conv	0.238	0.59	0.83	0.95
gauss_3.0_7x7.conv	0.218	0.54	0.76	0.87
gauss_4.0_7x7.conv	0.177	0.44	0.62	0.71
gauss_5.0_9x9.conv	0.140	0.35	0.49	0.56
gauss_7.0_11x11.conv	0.107	0.27	0.38	0.43

pixel, based on the local level of background noise. This helps the detection process in images where the background noise is varying over the frame: in regions of high noise, the detection threshold will be raised, preventing the detection of noise spikes; in regions of low noise, the detection threshold will be lowered to allow the detection of faint sources.

SExtractor will accept weight images of a variety of types. A common type is `MAP_VAR`, where the weight image is read in units of relative variance. SExtractor then scales to the appropriate absolute level by comparing the input variance map to an internally generated one. A model of the background light of an image is a good variance-type weight image: if background noise σ goes as $\sqrt{S}$, where S is the background value, then a map of the background value S at each pixel is actually a map of σ^2, the variance.

Throughout this work, unless otherwise specified, SExtractor was always run on the Coma images with a model of the background used as a variance-

type weight image. These background models were constructed as part of the background subtraction procedure. For detection and photometry on the degraded HDF images, unless otherwise specified, SExtractor was given the Coma background models as variance-type weight images. This is appropriate, as the noise in the degraded HDF was generated from the Coma background models.

Another common type of weight image is `MAP_WEIGHT`, where the weight image is read in units of relative weights. By definition,

$$variance \propto \frac{1}{weight} \tag{3.6}$$

Therefore, it should be equivalent to give SExtractor a background model as a variance type weight image, or the inverse of that model as a weight-type weight image. This was tested, and indeed found to be the case.

SExtractor also offers the user the option of imposing a weight threshold. In theory, this should mean that pixels with weights below the threshold (or variances above the threshold) will not be detected. However, in practice, it appears that SExtractor has difficulty with weights close to the weight threshold. For example, a simple weight-type weight image was tested on the Coma chip 2 image. The weight image had half its pixels set equal to 5 and the other half set equal to zero. With the weight threshold set to zero, this should mean that nothing in the weight = 0 region could be detected. However, this was not found to be the case: as Figure 3.1 shows, a few spurious detections are found in the weight = 0 region.

Further tests were performed with other two-valued weight images. Similar results were found for a weight image with weights of -100 and 5: with

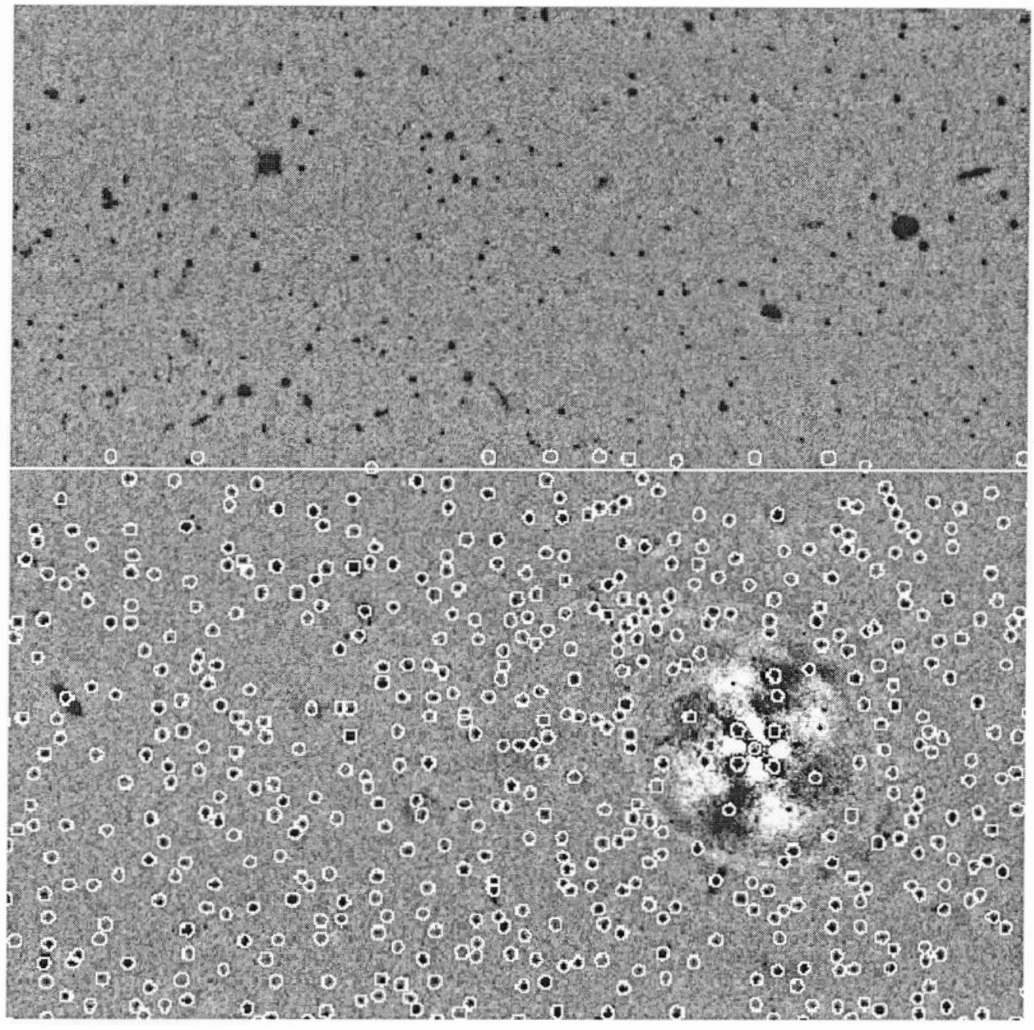

Figure 3.1: A subsection of the Coma chip 2 background-subtracted image, showing the results running SExtractor with a simple weight-type weight image and a weight threshold. The weight image had the lower half of its pixels set to 5 and the upper half set to zero; the boundary between the two regions is shown with a white line in this figure. The weight threshold was set to zero. The white circles show the detections made by SExtractor.

the threshold set to zero, a few detections were found in the weight = -100 region. However, when testing a weight image with weights of 0 and 9999, we found that setting the threshold to 100 resulted in no detections in the weight = 0 region.

This suggested a way to work around this problem. Recall that SExtractor accepts its weight images in units of *relative* weight or variance. If weight = 0 is used to flag pixels that should not be detected, the weight image can be multiplied by a constant to make the next smallest weight much larger than zero. Then, a threshold can be chosen at a sufficiently high value that no pixels with weight = 0 are detected.

Another "feature" of using weight thresholds is that they seem to make it difficult for SExtractor to determine aperture magnitudes. Setting a weight threshold, whether or not any pixels actually fall below the threshold, results in a large number of the aperture magnitudes being set to -99. This problem, and a way around it, are discussed in more detail in Section 4.1.4

3.1.4 Star-Galaxy Separation

The SExtractor package provides a neural-network-based method of star-galaxy separation. The network was trained on over 10^6 images of stars and galaxies, and will assign a "stellarity index" to an object based on 8 isophotal areas, the peak intensity, and the entered FWHM value.

As stated in the SExtractor User's Guide, the star/galaxy classifier is only experimental and should be used with caution. Another caveat is that the network was trained on ground-based images of stars and galaxies. This makes its use with space-based data, such as that from the HST, somewhat

ill-advised.

An alternative method for separating stars from galaxies is to compare their central concentration. Stellar objects are more centrally concentrated than galaxies. As well, all stellar objects have the same profile – the point spread function – which differs from star to star only in terms of the brightness scale. A plot of central concentration versus magnitude will therefore show a clear, magnitude-independent stellar trend. Objects in this locus can then be removed to leave a catalogue with only extended objects.

There are various measures of central concentration. A common one is the half-light radius $r_{0.5}$, defined as the radius which contains half of an object's total light. SExtractor will output the half-light radius of objects as `FLUX_RADIUS` if the parameter `PHOT_FLUXFRAC` is set to 0.5.

Other measures of central concentration are not calculated by SExtractor. The r_{-2} image moment (Kron, 1980) is defined by

$$\left(\frac{1}{r_{-2}}\right)^2 \equiv \frac{\int_1^\infty \frac{1}{x^2} g(x)dx}{\int_1^\infty g(x)dx} \tag{3.7}$$

where x is the radial distance from the object centre and $g(x)$ is the light distribution function. When calculating r_{-2}, the upper limit of integration is in practice set to some reasonable finite value.

Another measure of central concentration is the simplified Petrosian radius r_{petros} (Pritchet, 2003, private communication), defined as the radius at which the function

$$\eta' = \frac{F(<r)}{r} \tag{3.8}$$

reaches its maximum, where $F(<r)$ is the flux within an aperture with radius r. Note that η' is proportional to signal-to-noise when an image is sky-noise

dominated. This measure is loosely based on the Petrosian (1976) radius, which has been used to determine central concentration for various purposes in numerous studies (see, for example, Strauss et al., 2002).

To determine which star-galaxy separation method to use, SExtractor was run with a nominal set of input parameters on the background-subtracted Coma images to create a test catalogue. The half-light radius, r_{-2} image moment and simplified Petrosian radius were determined for each object. SExtractor was used to determine the half-light radius, and small additional programs were written to use the SExtractor outputs to calculate the other two measures.

Each measure of central concentration was plotted against magnitude (Figure 3.2). The stellar locus was defined by eye in each plot as those objects with $r_{0.5} \leq 1.44$, $r_{-2} \leq 1.13$ and $r_{petros} \leq 1.7$ respectively. These criteria were applied to create a galaxy catalogue for each method. These catalogues were then matched to find the objects that were classified as galaxies by all three selection methods: 79% of the $r_{0.5}$ selected galaxies, 68% of the r_{-2} selected galaxies, and 75% of the r_{petros} selected galaxies appeared in all three catalogues.

The objects that appeared in only one or two of the galaxy catalogues were then examined. First, the chip 2 objects that were classified as galaxies by the $r_{0.5}$ measure but not by the r_{-2} measure were examined by eye to determine if these objects should truly have been classified as galaxies. Next, the chip 2 objects classified as galaxies by the r_{-2} measure but not by the $r_{0.5}$ measure were examined in the same way. Based on these empirical observations, the $r_{0.5}$ measure seemed to create the better galaxy catalogue.

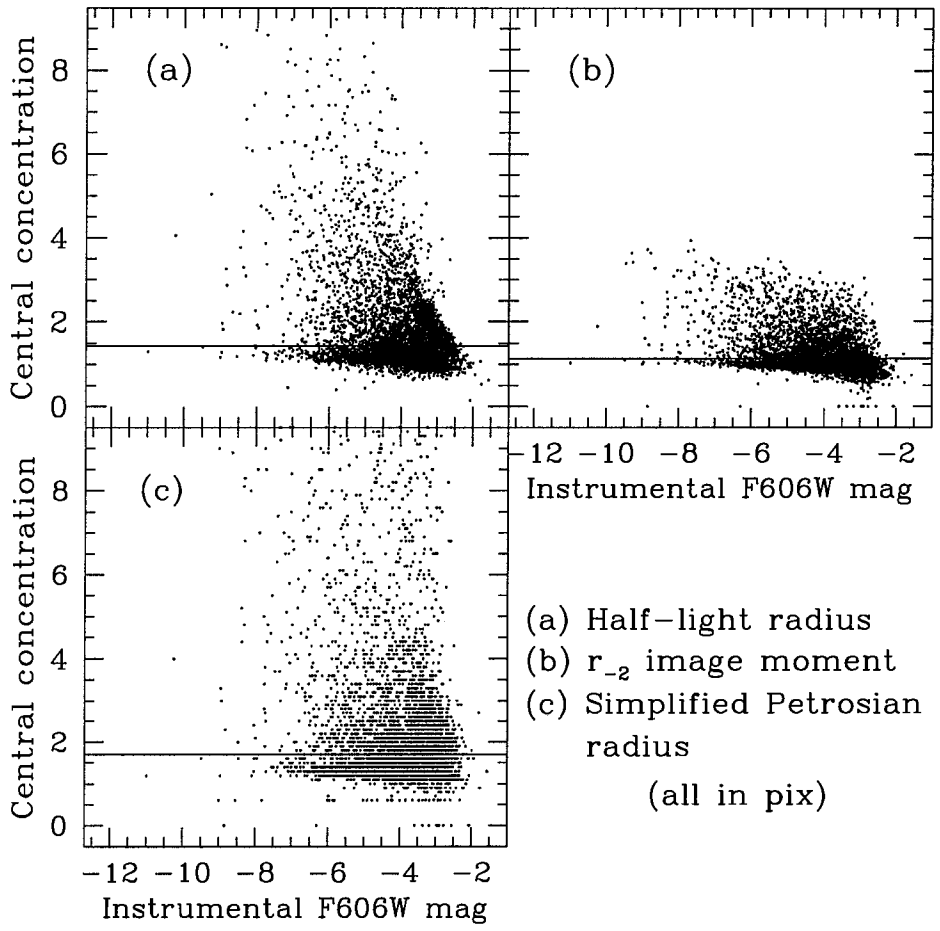

Figure 3.2: Central concentration versus instrumental F606W magnitude for objects in the Coma field. Three measures of central concentration are shown: (a) half-light radius; (b) the r_{-2} image moment; and (c) the simplified Petrosian radius. Shown on each plot is the line chosen to separate stars from galaxies: $r_{0.5} = 1.44$; $r_{-2} = 1.13$; and $r_{petros} = 1.7$. In each case, stellar objects fall below the line.

Finally, the r_{petros} galaxy catalogue was compared to the $r_{0.5}$ galaxy catalogue in the same way. Using the same criteria, the r_{petros} measure was found to create the best galaxy catalogue. Therefore, we decided to use r_{petros} to separate stars from galaxies in this work. The small program for calculating r_{petros} based on the SExtractor generated image centres was always run concurrently with SExtractor, and the calculated values appended as an extra column to the SExtractor output file.

It should be noted that SExtractor is an open source program that allows users to add their own functions and parameters to the code (more details are available in the Version 1.0a User's Guide). Future users may wish to incorporate the calculation of r_{petros} directly into the SExtractor code. In fact, the author of the SExtractor code has stated he hopes to include some measure of the Petrosian radius in a future version of SExtractor[2].

3.2 Choosing Optimum Parameters

SExtractor offers the user a great deal of control over the detection and photometry process through a wide range of user-controlled parameters. To determine the optimum set of these parameters, "add-galaxy" experiments were used: fake galaxies were created and added to the image, then SExtractor was run on the added-object image with a nominal set of parameters. The resulting catalogue was examined to determine how many of the added galaxies were recovered, and to compare their recovered magnitudes to their input magnitudes. Then, the process was repeated with different sets of pa-

[2]http://melimelo.iap.fr/mailserv_archive/sextractor/msg00253.html

rameters to see if the number and magnitudes of the recovered objects could be improved upon.

3.2.1 Generating The Fake Galaxies

The first step in "add-galaxy" experiments is to generate a list of the fake galaxies to be added. The IRAF task `gallist` was used for this. This task will produce a list of fake galaxies, indicating their xy position, magnitude, morphological type, half-power radius, axial ratio and position angle.

For the most part, the default settings for `gallist` were used. The parameters `xmax` and `ymax` were set to 741 and 757, the dimensions of chip 2. The luminosity distribution was changed to a uniform distribution between `minmag` and `maxmag`, and the absorption coefficient for edge-on spirals was set to zero magnitudes. The random seeds for generating xy coordinates and magnitudes were also changed, to be set to the clock time at execution.

The other parameters that needed to be changed were `minmag` and `maxmag`, the minimum and maximum magnitudes for the generated galaxies, and `eradius`, the maximum elliptical galaxy half-flux semi-major scale radius. To determine these values, the magnitudes and sizes of galaxies in the HDF were examined. SExtractor was run on the degraded HDF images using a nominal set of input parameters. From the resulting catalogue, a plot was made of the objects' Petrosian radius versus their magnitude (Figure 3.3). On the basis of that plot, objects with $r_{petros} \leq 1.7$ were removed as potentially stellar objects.

The remaining extended objects were divided into 12 0.5 mag bins, and the average half-light radius was determined for each bin. Table 3.2 shows

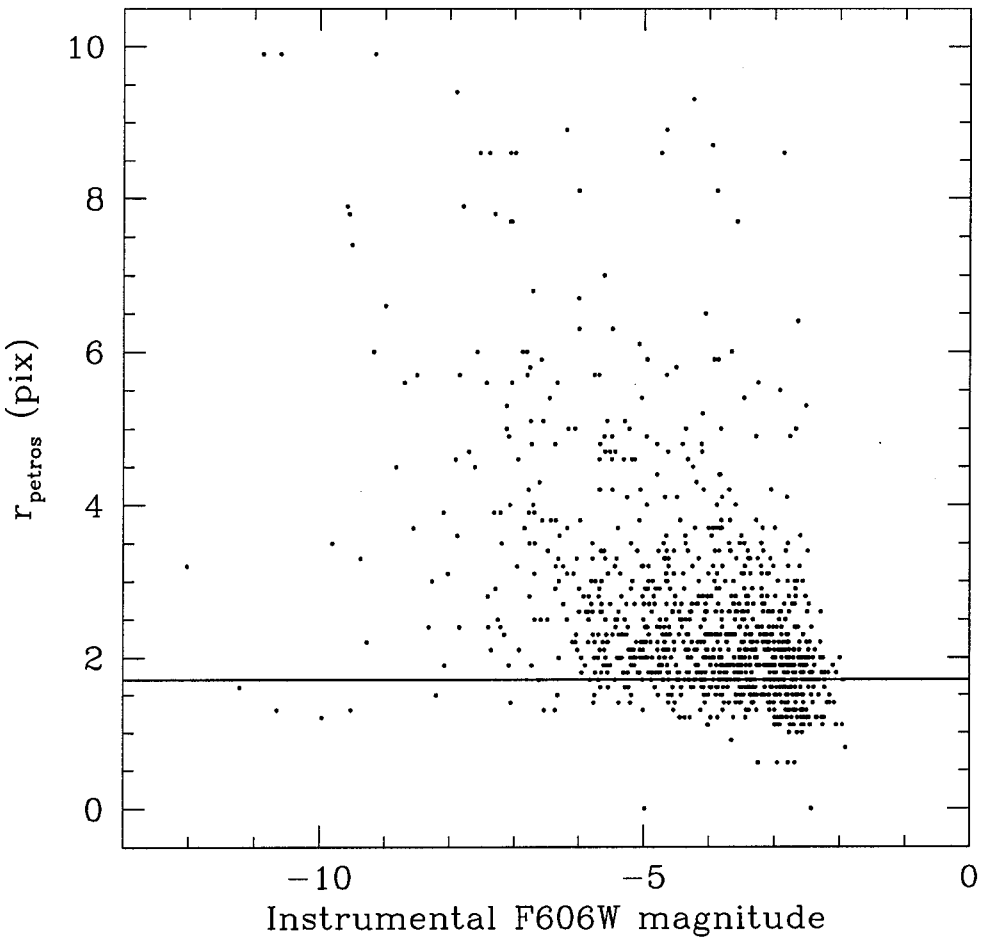

Figure 3.3: Petrosian radius versus magnitude for objects in the HDF. The few stellar objects fall under the line $r_{petros} = 1.7$.

the results. These values were then entered into `gallist`: `gallist` was run once for every 0.5 mag bin, with `minmag` set to the lower bound of that bin and `maxmag` set to the upper bound, `eradius` set to the typical half-light radius, and other parameters set as indicated above.

Table 3.2: Typical half-light radii of galaxies in the HDF

Magnitude range (instrumental mags)	Average half-light radius (pix)	Number of galaxies in bin
$-9.5 \leq m \leq -9.0$	5.73	4
$-9.0 \leq m \leq -8.5$	5.03	5
$-8.5 \leq m \leq -8.0$	3.31	5
$-8.0 \leq m \leq -7.5$	5.40	10
$-7.5 \leq m \leq -7.0$	4.31	22
$-7.0 \leq m \leq -6.5$	4.17	28
$-6.5 \leq m \leq -6.0$	3.43	30
$-6.0 \leq m \leq -5.5$	2.85	64
$-5.5 \leq m \leq -5.0$	2.80	69
$-5.0 \leq m \leq -4.5$	2.45	99
$-4.5 \leq m \leq -4.0$	2.31	83
$-4.0 \leq m \leq -3.5$	2.20	117

Before using the `gallist`-produced lists to generate the fake galaxies, the randomly-generated xy positions had to be slightly modified. The presence of large elliptical galaxies on the Coma image results in regions in which detections cannot be trusted. Even though the large elliptical galaxies have been removed, the regions they previously occupied have residual light and strange noise. This also affects the degraded HDF images; although they never had large elliptical galaxies, the noise which was added to them was

based on the Coma background models, which did include these galaxies. Therefore, a region must be defined around each of the large elliptical galaxies; detections falling in these regions, from either the Coma or HDF images, must be discarded.

If the point of "add-galaxy" experiments is to examine how many added objects are recovered, it makes no sense to add objects to regions from which all detections will be omitted. Therefore, a region was defined around each of the large elliptical galaxies, slightly larger than the fitted ellipse. The lists of fake galaxies were run through a small program which checked if the randomly generated xy positions fell in one of these omit regions. If so, a new xy position was randomly generated for that object. This process was repeated until none of the xy positions in the galaxy lists corresponded to omit regions on the chips.

The 12 modified galaxy lists were then used in the IRAF task `mkobjects`. This task takes a list of objects – such as that produced by `gallist` – and generates an image consisting of a constant background with the fake objects superimposed. The constant background was set to the value zero, and the type of star and point spread function parameter `star` was set to `gaussian`.

A note must be made about the galaxies produced by `mkobjects`. Two types of galaxies can be generated: `elliptical` galaxies, represented by de Vaucouleurs surface brightness profiles; and `spiral` galaxies, represented by exponential disk surface brightness profiles. The functional forms used by `mkobjects` for these two profiles are given by

$$I_{devauc} = e^{-7.67(\frac{r}{r_e})^{\frac{1}{4}}} \tag{3.9}$$

$$I_{expdisk} = e^{-1.6783(\frac{r}{r_e})} \tag{3.10}$$

where I is intensity, r_e is the half-flux semi-major scale radius, and r is the circularly symmetric radius. These two profiles are plotted in Figure 3.4. As can be seen, the de Vaucouleurs profile is much more centrally concentrated than the exponential disk profile.

This presents a problem for the task `mkobjects`. When it creates a fake galaxy, the innermost pixels will necessarily depart somewhat from the strictly mathematically correct values, due to pixel quantization. For a profile as centrally-concentrated as the de Vaucouleurs profile, these small errors can lead to a substantial fraction of the flux being lost. This can be seen by constructing an image with a single fake galaxy and performing aperture photometry with a range of aperture radii, so as to produce a growth curve for the fake galaxy. Figure 3.5 shows the result of this process for an exponential disk profile galaxy and a de Vaucouleurs profile galaxy with the same total flux. The less centrally-concentrated exponential disk galaxy has a growth curve that approaches the total input flux. The more centrally-concentrated de Vaucouleurs galaxy has a growth curve that levels off at $\sim$ 90% of the input flux.

This result was found to be independent of the magnitude of the de Vaucouleurs galaxy. A series of de Vaucouleurs galaxies with varying magnitudes were created and aperture photometry was used to obtain a growth curve for each. Table 3.3 shows the results: de Vaucouleurs galaxies at all magnitudes are created by `mkobjects` with only 89% of their input flux.

A loss of 11% of flux corresponds to a gain of 0.126 mag. Therefore, one

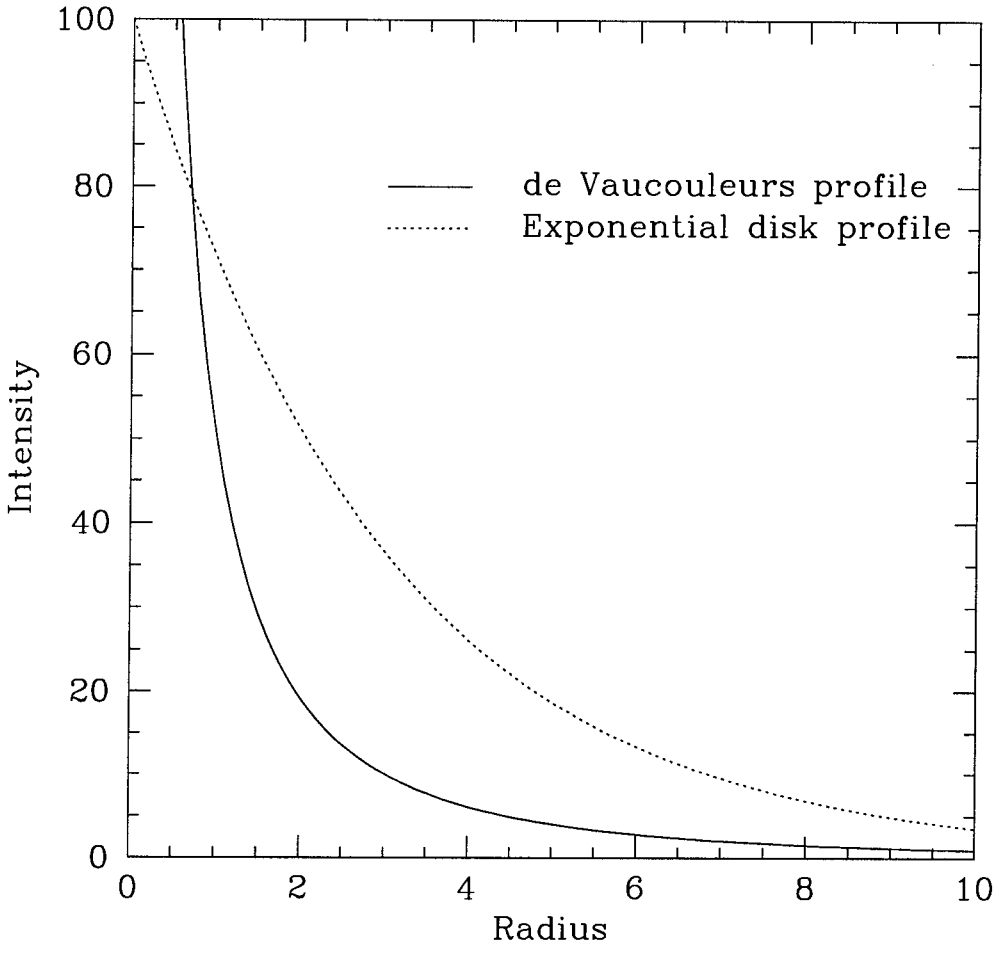

Figure 3.4: Intensity versus radius for the de Vaucouleurs profile (solid line) and the exponential disk profile (dotted line) as used by `mkobjects`. The units of the axes are arbitrary. The two curves have been constructed with the same scale radius, and the area under the curves is equal when integrated to infinity.

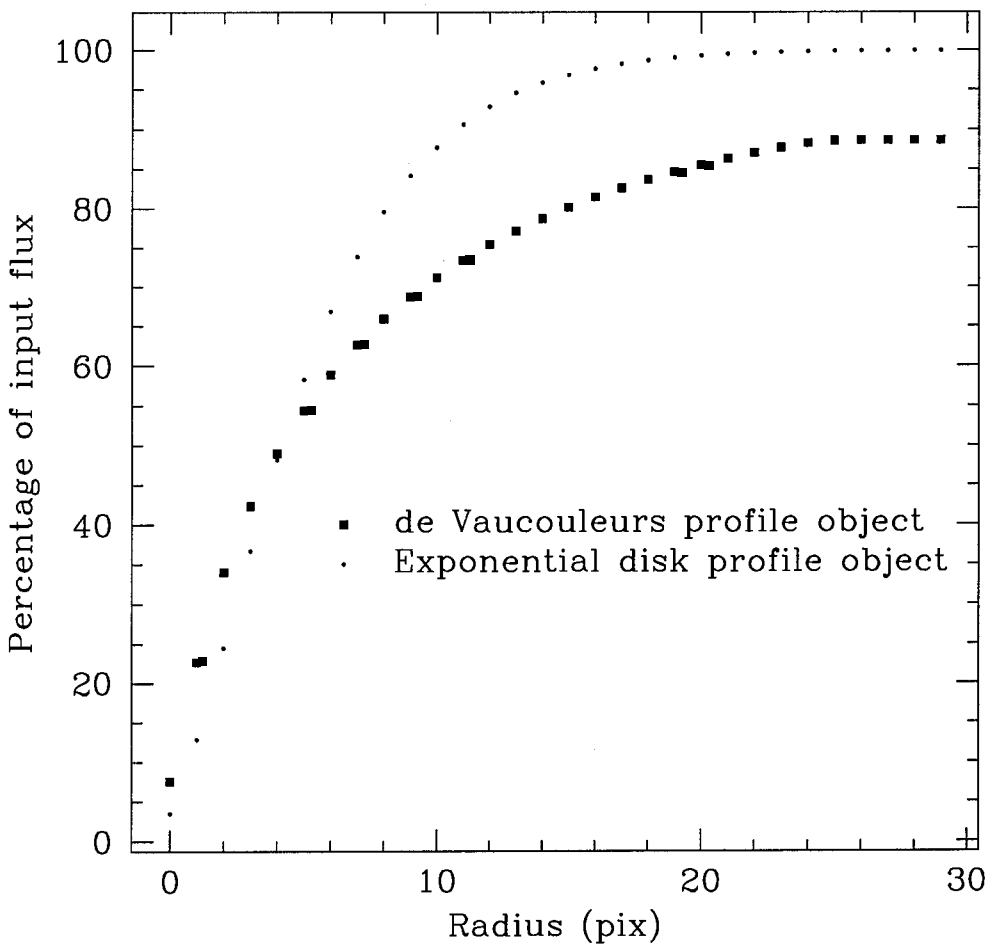

Figure 3.5: Growth curves of a de Vaucouleurs profile object (squares) and an exponential disk profile object (circles), obtained by performing aperture photometry. Both objects were generated with the IRAF task `mkobjects` and set to have the same total flux.

Table 3.3: Input and output flux of de Vaucouleurs galaxies created by `mkobjects`

Input Magnitude (instrumental mags)	Input Flux (e^-)	Output Flux (e^-)	Input Flux / Output Flux
-1	2.51	2.23	0.887
-2	6.31	5.60	0.887
-3	15.8	14.0	0.887
-4	39.8	35.3	0.887
-5	1.00×10^2	88.7	0.887
-6	2.51×10^2	2.23×10^2	0.887
-7	6.31×10^2	5.60×10^2	0.887
-8	1.58×10^3	1.40×10^3	0.887
-9	3.98×10^3	3.53×10^3	0.887
-10	1.00×10^4	8.87×10^4	0.887

must consider a de Vaucouleurs profile galaxy created by `mkobjects` with a magnitude M to actually have a magnitude $M + 0.126$.

After running `mkobjects`, the final result was 12 images, each with 100 fake galaxies, with magnitudes drawn randomly from a 0.5 mag range, and sizes typical of HDF galaxies in that magnitude range. These images were each then added to Coma and HDF chip 2 data images to produce 12 different added-object images for both Coma and the HDF. The specific chip 2 data image to which the fake galaxies were added depended on the field. For the HDF, the fake galaxy images were added to the chip 2 degraded HDF image. For Coma, the fake galaxy images were added to the ellipse-subtracted, but not background-subtracted, chip 2 image. The objects were added at this point so that the fake galaxies would undergo the same background sub-

traction process as the real objects. This allows the effects of background subtraction on detection and photometry to be explored.

3.2.2 Background Subtraction Method

The first thing that was explored with "add-galaxy" experiments was the question of which background subtraction method to use. In creating the original background-subtracted Coma images, a ring median filter with inner radius 7 pix and outer radius 10 pix was applied to the ellipse-subtracted images to create models of Coma's background. These models were subtracted from the ellipse-subtracted images to create the background-subtracted images (Section 2.1.3). The results of this process, however, left a strange "pebbly" looking background (Figure 3.6a). Could this be improved upon?

The ring median and subtraction process was repeated on the Coma chip 2 image a number of times, trying a range of different values for the inner and outer radii of the ring median. The most satisfactory looking image resulted from an inner radius of 10 pix and an outer radius of 20 pix (Figure 3.6b).

Another background subtraction method was also explored: SExtractor was run on the ellipse-subtracted image, with the check image set to `BACKGROUND`. The result was an image of the interpolated background SExtractor had estimated for the data image. This was then subtracted from the ellipse-subtracted image to create the background-subtracted image. This process was repeated a number of times, trying a range of different values for the background estimation parameters `BACK_SIZE` and `BACK_FILTERSIZE`. The best looking image from this technique was found to be generated with `BACK_SIZE` set to 20 pix, and `BACK_FILTERSIZE` set to 3 (Figure 3.6c).

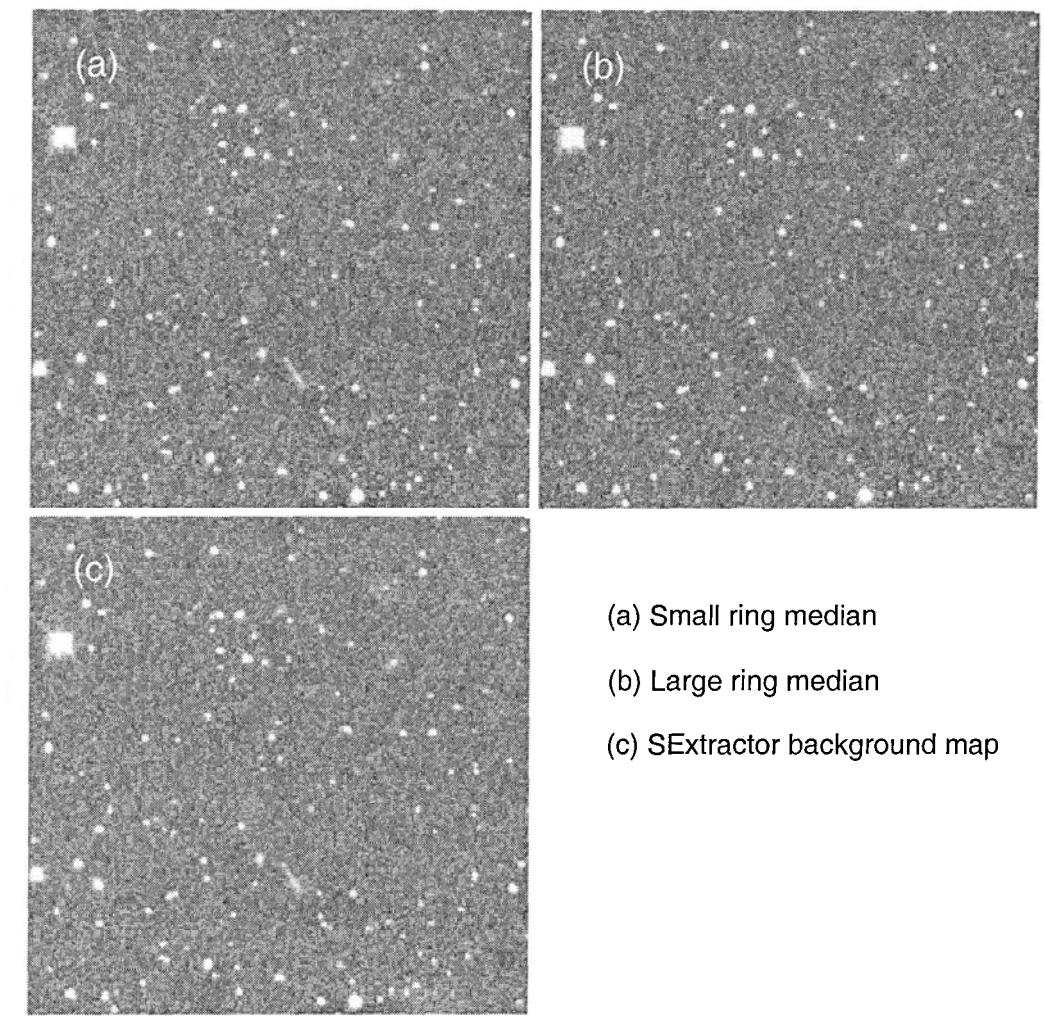

Figure 3.6: A subsection of the chip 2 background-subtracted Coma image. Three different methods of creating a background model from the original coadded image were used: (a) a small ring median filter with an inner radius of 7 pix and an outer radius of 10 pix; (b) a large ring median filter with radii of 10 and 20 pix; and (c) the SExtractor `BACKGROUND` check image. In each case, the background model was subtracted from the coadded image to produce the image shown.

"Add-galaxy" experiments were used to determine which of these two methods produced the best background-subtracted image. For each of the 12 0.5 mag bins, a list of 100 fake galaxies was generated with `gallist`. The randomly generated xy coordinates were modified such that no object centres would fall in an omit region. The task `mkobjects` was used to generate an image of the fake galaxies, and this was added to the chip 2 Coma ellipse-subtracted image. The background subtraction method being tested was then used on the added-object image. SExtractor was run on the background-subtracted image using a nominal set of parameters. The resulting catalogue was matched to the input list of galaxies, and the number of added galaxies that was retrieved was recorded. The difference between input magnitude and found magnitude for each of the retrieved galaxies was also recorded, taking into account the 0.126 mag difference between the input magnitude and the generated magnitude for de Vaucouleurs profile galaxies. This process was repeated 20 times for each bin, and the results combined.

The results from the ring median subtraction method and the SExtractor subtraction method were compared. Figure 3.7 shows the average percentage of added objects that were recovered in each of the 0.5 mag bins. The ring median method and the SExtractor method show very similar results. Figure 3.8 shows the average difference between the input magnitude and the found magnitude for objects in each of the 0.5 mag bins. Again, the ring median and SExtractor methods show very similar results.

Background models created using a ring median filter or using SExtractor produce very similar background-subtracted images, both in the aesthetic "look" of the image, and in its detection and photometric properties. For

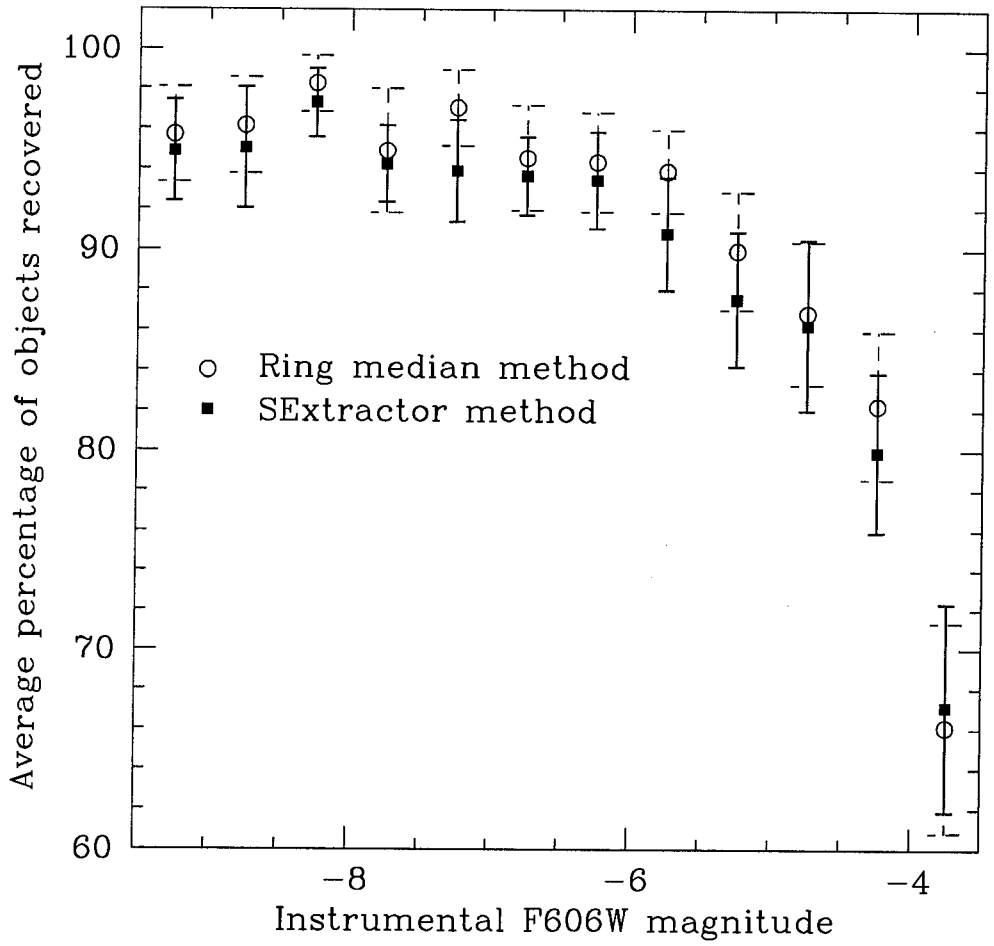

Figure 3.7: Average percentage of added objects recovered in each magnitude bin. Open circles represent the results when the added-objects images are background subtracted using the ring median method; solid squares show the results when background subtraction is done by SExtractor. Each point represents the average of 20 runs. Error bars are the standard deviation in the mean.

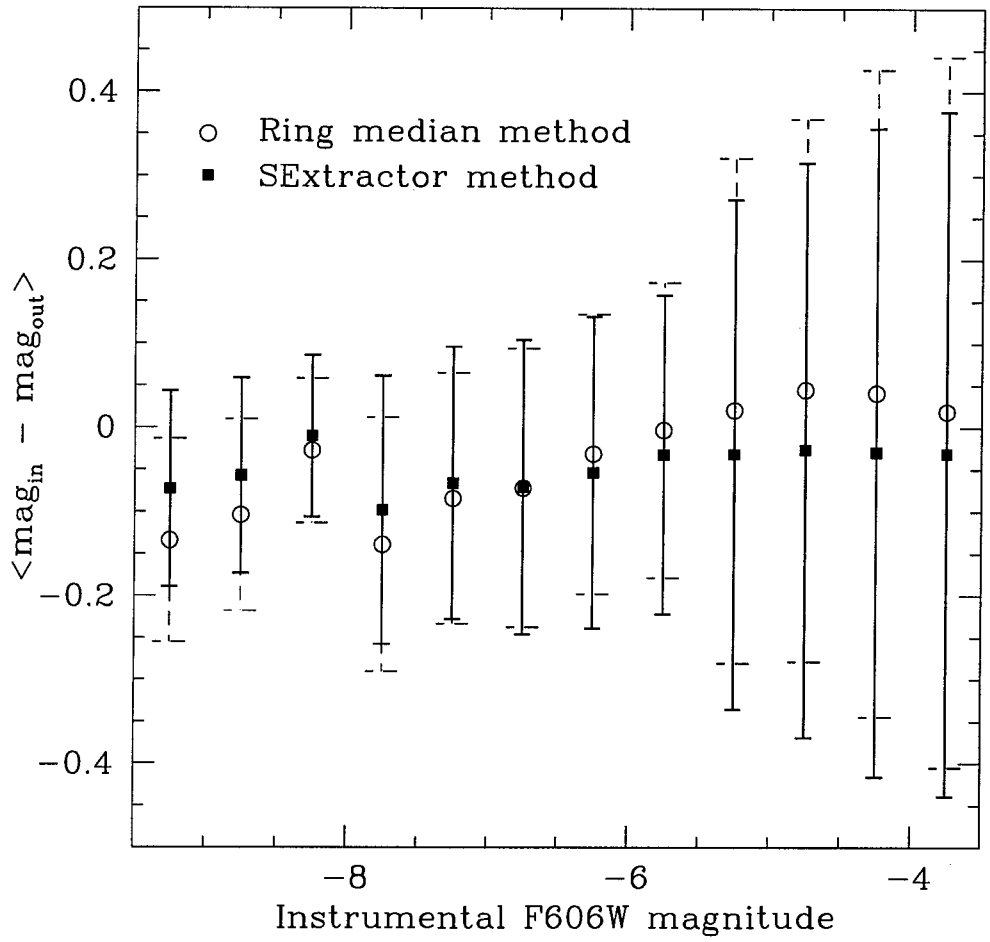

Figure 3.8: Average change in magnitude between added and recovered objects in each magnitude bin. Open circles represent the results when the added-object images are background subtracted using the ring median method; solid squares show the results when background subtraction is done by SExtractor. Each point represents the average of 20 runs of 100 added objects. Error bars are the standard deviation in the mean.

this work, the SExtractor method was chosen for its substantially greater speed.

The old background-subtracted Coma images were discarded, and the new SExtractor-based subtraction method was used to create new background-subtracted Coma images[3]. The old degraded HDF images were also discarded, and new ones generated based on the new Coma background models. All the tests outlined in Section 2.2.3 were repeated for the new background and noise models. It was again found that a better model of Coma's noise could be produced by using 85% of the noise generated from the background model. It was also found that the noise model was again a good representation of the true noise in the Coma image, and that the noise of the degraded HDF was again a good match to the noise in Coma.

3.2.3 Other Parameters

"Add-galaxy" experiments were then used to determine the best set of SExtractor parameters for detection and photometry. The Coma image was examined first. A set of 12 added-object images were created using the procedure described in Section 3.2.1. These were background subtracted using the SExtractor-based method chosen in Section 3.2.2. SExtractor was then run on each of the 12 added-object images with a nominal "best guess" set

[3]The inquisitive reader may wonder why the initial background subtraction was described in Section 2.1.3, if only to be discarded here. The initial background subtraction was necessary to generate background maps of the Coma image; these were necessary to create noise maps with which to degrade the HDF. The HDF had to be degraded so that the typical sizes of background galaxies in a Coma-like image could be found; these were required to generate realistic fake galaxies for "add-galaxy" experiments; which in turn were finally used here to determine the proper background subtraction method.

of parameters. The resulting catalogues were compared to the input galaxy lists. The number of input galaxies that were found in each magnitude bin was recorded, as was the difference between the input magnitude and the found magnitude for each of the retrieved galaxies in that bin.

The process was repeated a number of times, each time changing one SExtractor parameter while holding the others fixed. The results from each run were compared to the initial nominal results: if the new set of parameters showed improvements in the detection and photometry of the added galaxies, those results were kept as the new "best guess" and future runs compared to them.

Table 3.4 (at the end of this chapter) shows the SExtractor parameters that were tested, and the values that were tried for each. The final best set of parameters determined by these trials can be found in Table 3.5 (also at end).

The HDF was then examined. Fake galaxies were generated in the same manner as for Coma, and added to the degraded HDF chip 2 image. This added-object image was then background subtracted using the same method employed on the Coma image. The HDF is of course already background-subtracted. However, this additional step was performed to keep the processing steps as identical as possible between the data image and the control image.

The initial "best guess" set of SExtractor parameters was set to be the best set found for the Coma data. The same procedure of varying the parameters was carried out, examining if any set of parameters could improve upon the results from the best set of Coma parameters. No better set of

parameters was found, and so the same best set of parameters was chosen for both Coma and the HDF.

3.3 Comparing Detection Characteristics

Once the best set of detection parameters had been chosen, the detection characteristics of Coma and the HDF could be compared. Recall that in the method of differential counts, it is vitally important that the data field and the control field have identical detection characteristics. It should be just as hard (or as easy) to detect objects of all magnitudes in one field as in the other, and those objects should have equally accurate (or inaccurate) magnitudes in both fields.

To ensure that this was the case, "add-galaxy" experiments were performed on Coma and the HDF. The procedure was much the same as that used in Section 3.2.2 to determine the best background subtraction method. Again, a list of 100 fake galaxies was generated for each of the 12 0.5 mag bins. The xy coordinates were edited until no objects fell in omit regions, and the fake galaxy images were generated. These were added to the data image, which was then background subtracted. SExtractor was run on the background-subtracted added-object image using the best set of parameters. The catalogue was matched to the input list of fake galaxies, and the number of objects recovered and their input and found magnitudes were recorded. This process was repeated 20 times for both Coma and the HDF.

The results from Coma and the HDF were compared. Figure 3.9 shows the average difference between the input and found magnitude for objects

in each of the 0.5 mag bins. As can be seen, there is a slight difference in the recovered magnitudes at the faint end: faint HDF objects are found to be fainter than Coma objects of similar magnitudes. The greatest difference, however, is $\lesssim 0.07$ mag, a negligible amount compared to the typical bin size of a luminosity function.

Figure 3.10 shows a slightly more worrying result. Comparing the average percentage of added objects that were recovered in each of the 0.5 mag bins, it appears that it is easier to find faint objects in the HDF than it is in Coma.

3.3.1 The Effects Of Crowding

Crowding is one possible explanation for why it is easier to find faint objects in the HDF than in Coma. The Coma fields have many more objects – cluster galaxies and globular clusters – than the HDF fields. An added object therefore has a greater chance of landing on or near a real object. If this occurs, a blend of the two objects will probably be detected as a single object by SExtractor. If the added object is faint, the centre of the blend will remain close to the centre of the original object, and so the faint added object will be lost in the blend. If the added object is bright, however, the centre of the blend will move much closer to the centre of the added object, and the added object will still be found.

A way to test this hypothesis is to apply another correction to the randomly generated xy coordinates of the added objects. If the centre of an added object falls on a pixel already associated with a real object, a new set of coordinates should be generated. If added objects are not allowed to fall

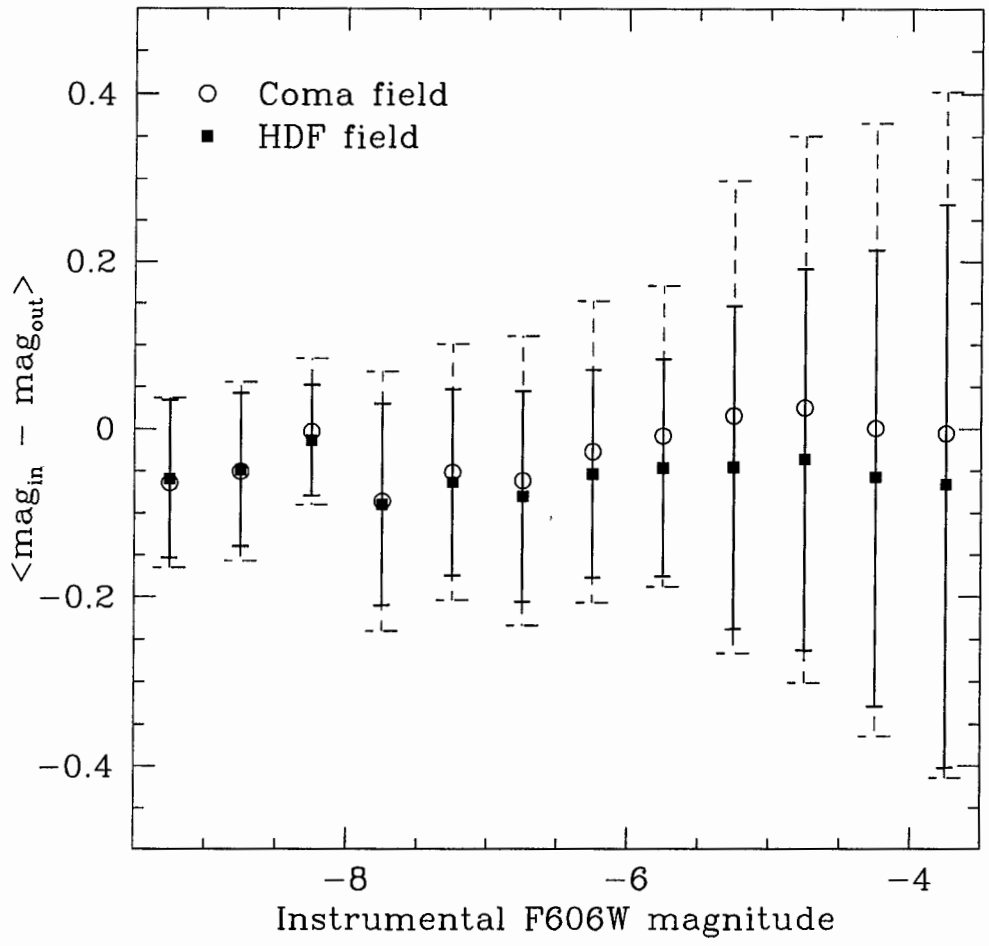

Figure 3.9: Average change in magnitude between added and recovered objects in each magnitude bin. Open circles represent the results of objects added to the Coma field; filled squares are the results of objects added to the HDF field. Each point represents the average of 20 runs of 100 added objects. Error bars are the standard deviation in the mean.

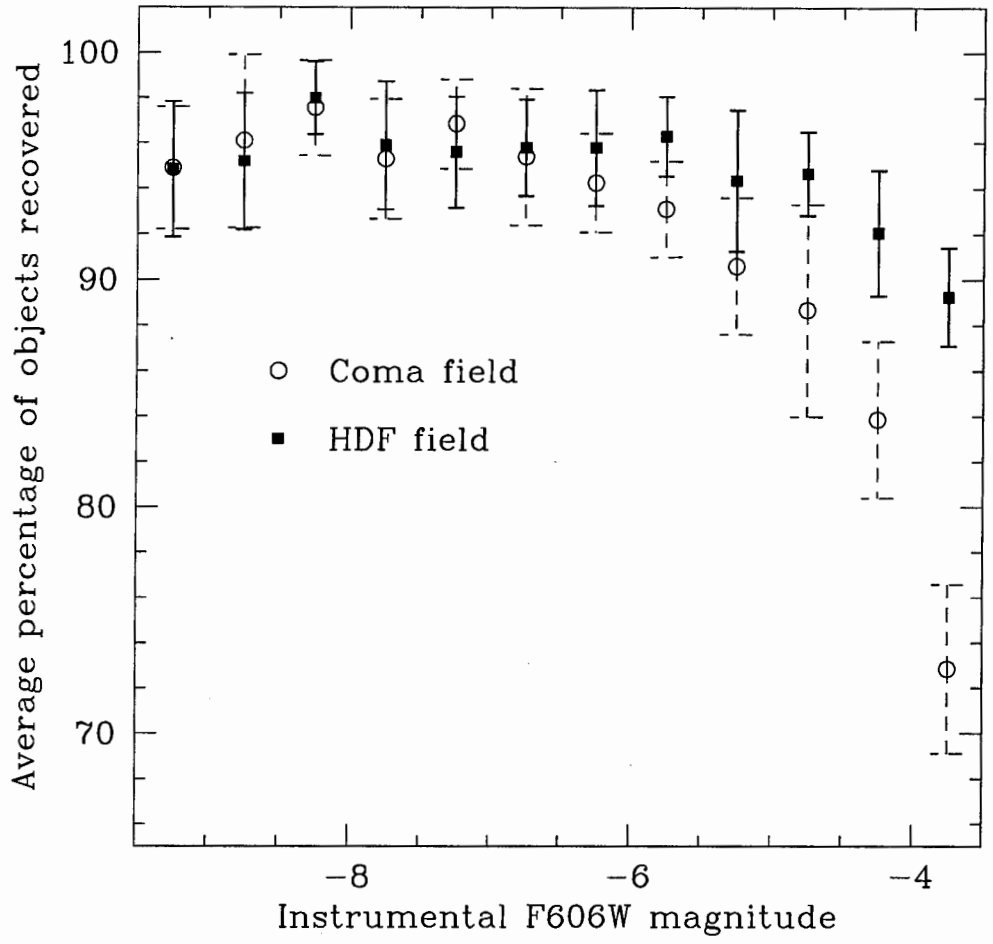

Figure 3.10: Average percentage of added objects recovered in each magnitude bin. Open circles represent the results of objects added to the Coma field; filled squares are the results of objects added to the HDF field. Each point represents the average of 20 runs. Error bars are the standard deviation in the mean.

on existing objects, the number of blends will be greatly reduced, and the effects of crowding basically negated.

To determine which pixels were associated with real objects, SExtractor segmentation images were used. A SExtractor segmentation image is an image the same size as the data image, with each pixel set to the identification number of the object to which it belongs. Pixels not belonging to any object are assigned a value of zero.

SExtractor was run once on a background-subtracted Coma image with no added objects, with the check image set to `SEGMENTATION`. The resulting segmentation image was used in a small program that checked the pixel value of the segmentation image at the location corresponding to the xy coordinates of objects to be added. If the pixel value was non-zero – if that pixel was associated with a real object – a new centre was randomly generated for that object. The program also checked if the object fell in an omit region, and continued generating new coordinates and checking them until neither criterion was met.

With this new algorithm to modify the xy coordinates, the same "add-galaxy" experiment was performed. Figure 3.11 shows the new comparison of the average percentage of added objects that were recovered. The agreement between Coma and the HDF has improved to some degree: crowding in Coma does make it more difficult to find faint objects. However, the overall results still disagree. There must be another factor making it easier to find faint objects in the HDF than in Coma.

The nature of this factor is unknown. However, the effects of it on the luminosity function will be small and can be stated: in faint magnitude bins,

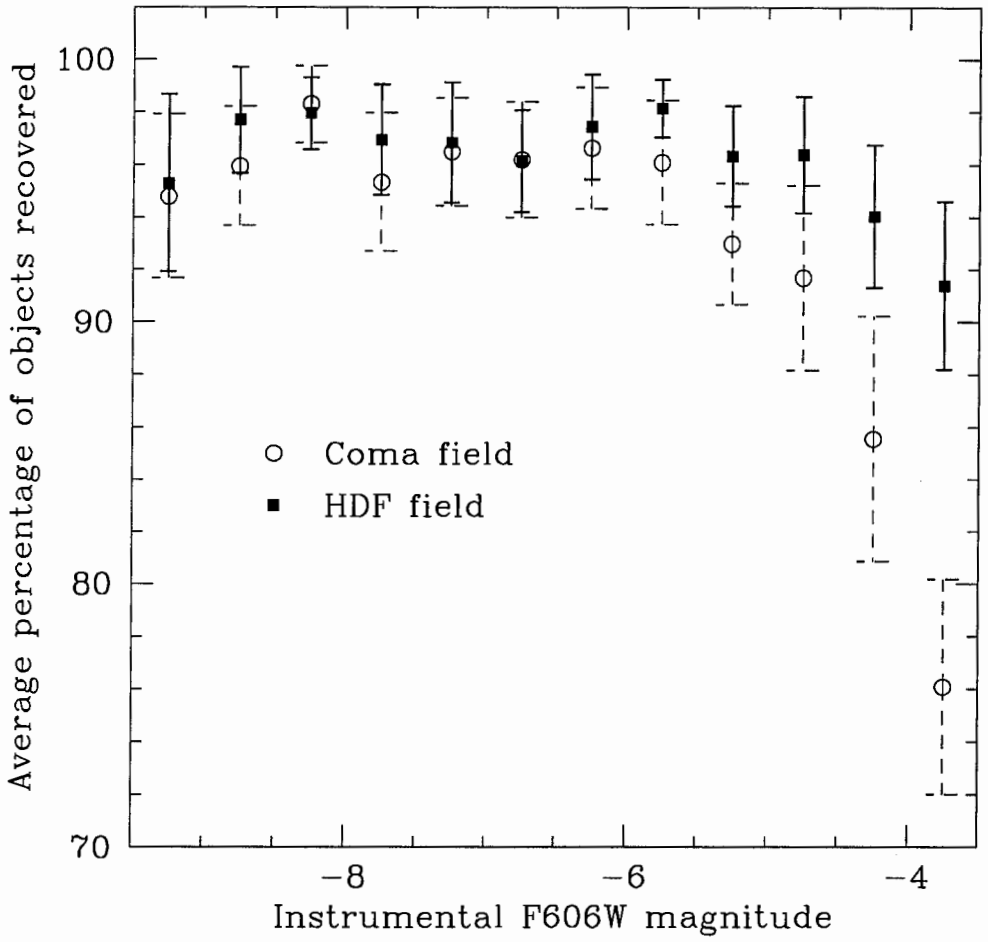

Figure 3.11: Average percentage of added objects recovered in each magnitude bin. Open circles represent the results of adding objects to the Coma field; filled squares are the results of adding objects to the HDF field. In both cases, added object centres were not allowed to fall on pixels associated with real objects. Each point represents the average of 20 runs. Error bars are the standard deviation in the mean.

the background will be over-subtracted, leading to an artificial depression of the slope of the luminosity function[4]. Whatever slope is found for the luminosity function must be considered a lower limit to the true slope.

3.4 Initial Catalogues

Initial catalogues of Coma and HDF objects were finally made. SExtractor was run using the best set of parameters on the 3 background-subtracted chips of each field. Photometry was done in both filters: for the F814W images, SExtractor's double image mode was used. The F606W image was used to determine object centres and isophotes, which were then applied to the F814W image to calculate magnitudes. This greatly simplified the process of using the F814W photometry to calculate colours for the detected objects. Also to aid in the calculation of colours, SExtractor was set up to do aperture photometry as well as Kron photometry. The diameter of the aperture was set to 12 pix.

3.4.1 Comparing Magnitude Errors

As a final check that the detection characteristics of Coma and the HDF were well-matched, the magnitude errors of the objects in each initial catalogue were compared. SExtractor calculates magnitude error as

$$\Delta m = 1.0857 \frac{\sqrt{A\sigma^2 + \frac{F}{g}}}{F} \tag{3.11}$$

[4]The luminosity function is plotted as the number of galaxies as a function of magnitude in this case.

where A is the area in pixels over which the flux F (in DN) is determined, σ is the estimated standard deviation of the background noise (again in DN), and g is the gain. As pointed out in the SExtractor documentation, this error does not include any uncertainty in the estimate of the local background, and so must be considered only a lower limit to the true magnitude error. Nevertheless, it is sufficient for our purposes.

Figures 3.12 and 3.13 show the magnitude errors in F606W and F814W as a function of magnitude. As can be seen, the magnitude errors of Coma objects and HDF objects are very similar at all magnitudes, a good indication that the fields are well-matched.

3.4.2 Comparison To Published Catalogues

The initial Coma and HDF catalogues can be compared to published catalogues made from the same field. This is a good check that photometry is being carried out correctly: the magnitudes and colours in the initial catalogues should closely match those published.

Coma

The initial Coma catalogue was compared to that of Kavelaars et al. (2000). The data on which that paper was based were kindly provided by the authors, as was the script they used to transform their instrumental F606W and F814W magnitudes to their published Vega V and I magnitudes. This script was used in reverse to transform their catalogue back into instrumental magnitudes, which were then further transformed into the instrumental system used in our initial catalogue.

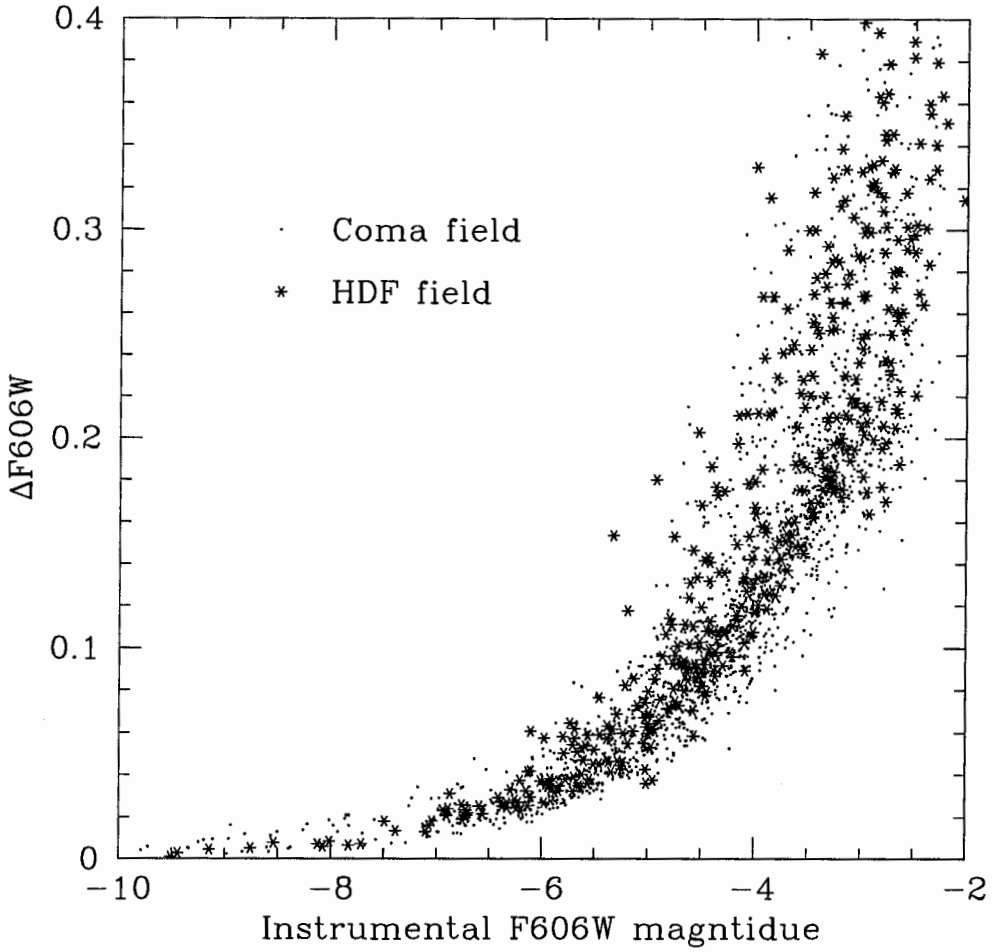

Figure 3.12: Magnitude error versus magnitude for the F606W passband. Dots represent objects in the Coma catalogue; stars represent HDF catalogue objects. Magnitude and magnitude error were calculated by SExtractor.

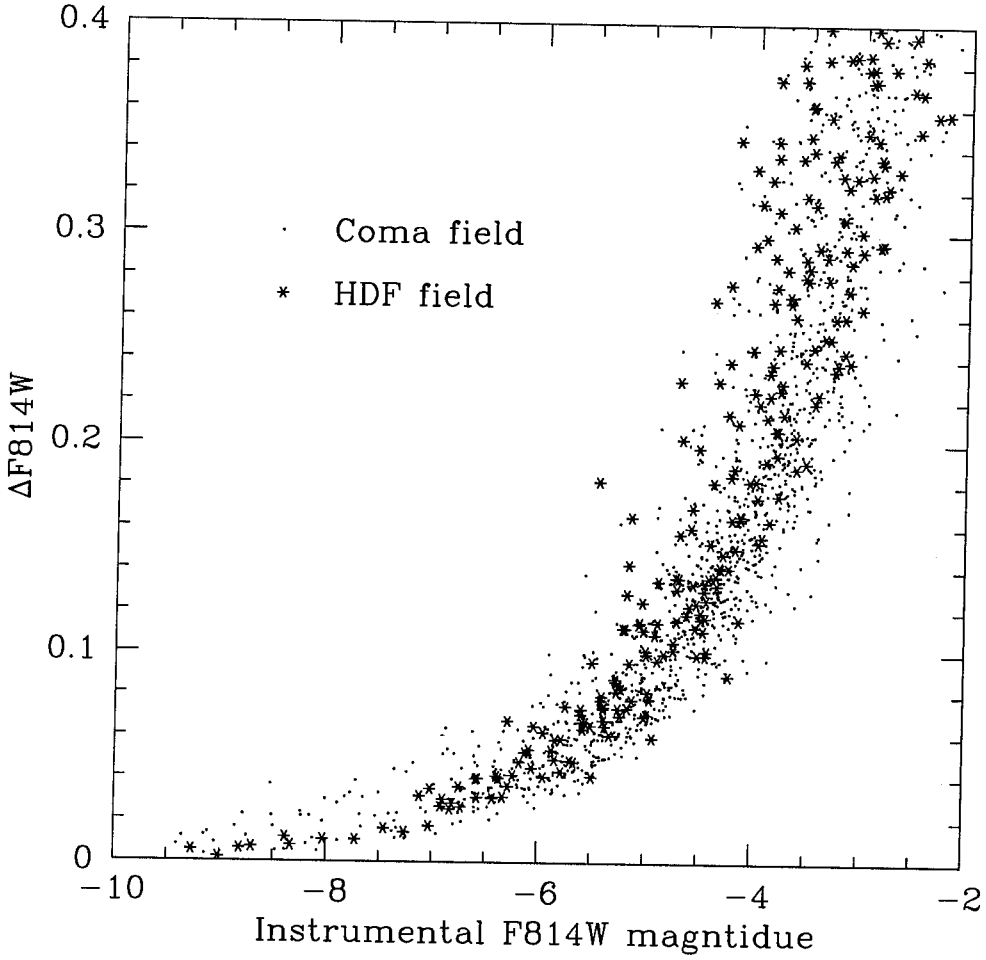

Figure 3.13: As Figure 3.12, for the F814W passband.

A program based on the Groth pattern matching algorithm (Groth, 1986) was used to align the coordinate system of Kavelaars et al.'s catalogue to that of the initial Coma catalogue. The transformed, aligned catalogue was matched to the initial Coma catalogue, and the objects in common were compared.

Figure 3.14 shows the difference in F606W magnitude between the published magnitudes and these magnitudes, as a function of the published magnitude[5]. An offset is definitely present: at all magnitudes, the objects in the initial Coma catalogue were found to be brighter than their published magnitudes.

This is a somewhat worrisome result. The offset can probably be partially explained by the different photometry algorithms employed for the two catalogues. Kavelaars et al. were studying globular clusters in the Coma cluster, and so their catalogue contained only stellar objects. For these objects, point spread function photometry can be used to obtain good estimations of the total magnitude. In this work, however, we are interested in extended objects and so cannot use a photometry algorithm based on the point spread function; instead, we chose to use Kron magnitudes. This difference in techniques may result in part of the difference seen in the magnitudes.

However, an offset of ~ 0.2 mag cannot be totally attributed to different photometry routines. As a further check of the extent of this discrepancy, colour magnitude diagrams made from the two catalogues were also com-

[5]The F814W magnitudes were not compared to published values; as these magnitudes were calculated in double image mode, comparison to any form of "total" magnitude would be difficult.

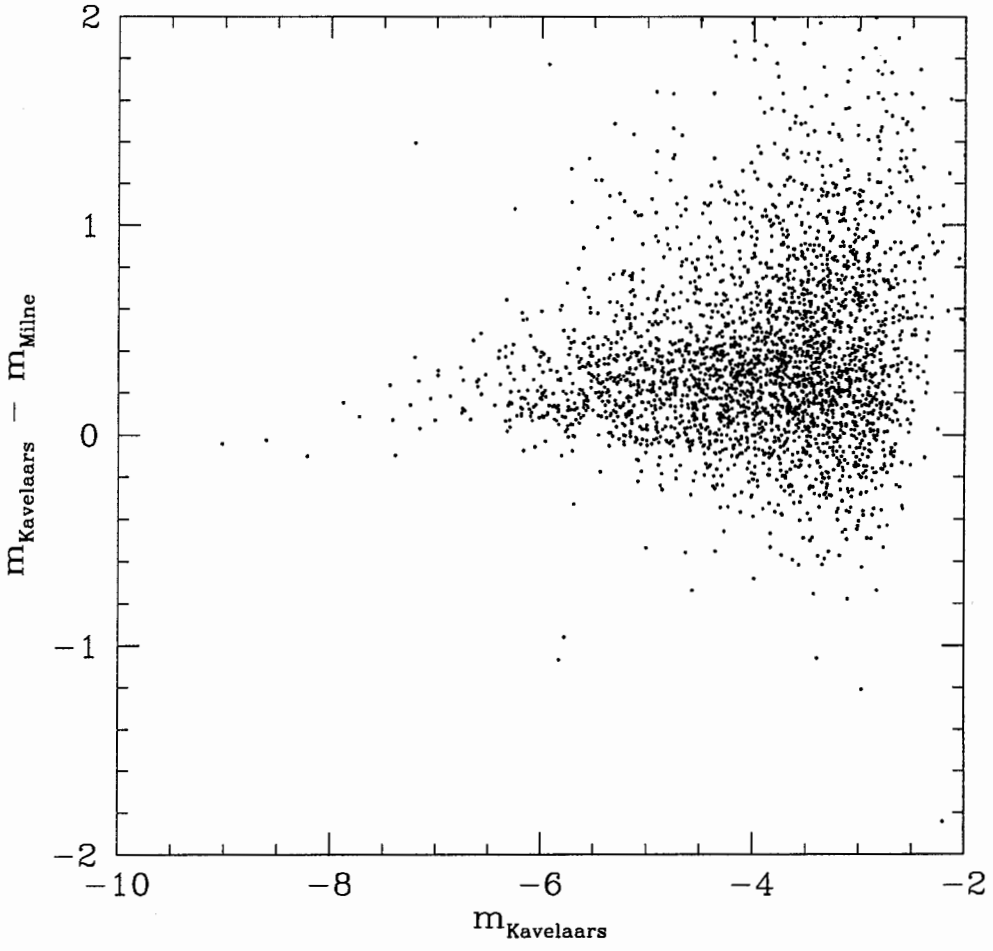

Figure 3.14: Difference in instrumental F606W magnitude as given in the Coma catalogue of Kavelaars et al. (2000) and as determined here, versus magnitude in the Kavelaars et al. catalogue. The magnitudes in the Kavelaars et al. catalogue have been transformed to the instrumental system of this work.

pared. Colour was calculated for the initial Coma catalogue by subtracting the F814W instrumental aperture magnitude from the F606W instrumental aperture magnitude for each object. For Kavelaars et al.'s catalogue, only the "total" magnitude was supplied, so this was used in the colour calculation.

Figure 3.15 shows the two colour magnitude diagrams superimposed. The initial Coma catalogue produces a colour magnitude diagram with the same typical colour as that found from Kavelaars et al.'s catalogue. This is a good indication that the systematic offset found when comparing magnitudes is due more to different photometry techniques than to an error in determining zeropoint.

The initial Coma catalogue is not a perfect match to the published Coma catalogue. However, taking into account the differences in the photometry algorithms, and the encouraging match between colour magnitude diagrams, we are satisfied with the comparison.

HDF

The initial HDF catalogue was compared to that of Couch, available from the Hubble Deep Field site at the Space Telescope European Coordinating Facility[6]. As indicated in the documentation at that site, the magnitudes are expressed as F606W and F814W ST magnitudes. Using the HDF photometric zeropoints on the STScI webpage[7], Couch's catalogue was transformed into the instrumental magnitude system of the initial HDF catalogue.

The coordinate system of the transformed catalogue was aligned to that

[6] http://www.stecf.org/hstprogrammes/hdf/catalogs/
[7] http://www.stsci.edu/ftp/science/hdf/logs/zeropoints.txt

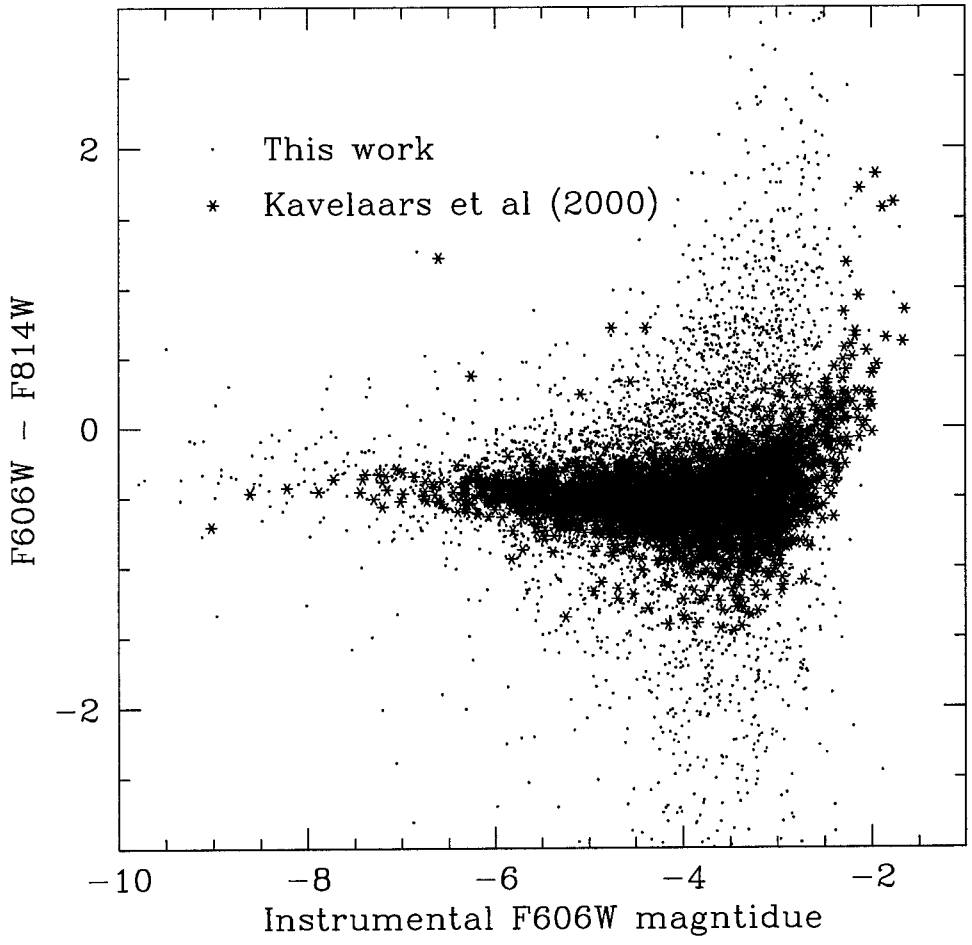

Figure 3.15: Colour magnitude diagrams as constructed with the Coma catalogue of Kavelaars et al. (2000) (stars) and with this catalogue (points). All magnitudes are instrumental in the system used in this work.

of the initial HDF catalogue. The catalogues were matched, and the objects in common were compared.

The difference in F606W magnitude between the published magnitudes and these magnitudes as a function of the published magnitudes is shown in Figure 3.16. The magnitudes agree much better here than they did for Coma: likely, part of the reason why is that Couch's catalogue was also generated using the SExtractor package, and so the effects of different photometry routines are removed.

Colour magnitude diagrams were also used to compare the initial HDF catalogue to Couch's catalogue. Figure 3.17 shows the two colour magnitude diagrams superimposed. The colours from the two catalogues are very similar. The initial HDF catalogue does show more scatter in colour at the faint end than the published catalogue: this is due to the noise that has been added to degrade the HDF. Couch's catalogue was of course made on the true, undegraded HDF, and so has much less noisy magnitudes at the faint end.

Based on the good correspondence seen between our catalogue and Couch's for both magnitudes and colours, we conclude that the initial HDF catalogue is a good match to a published catalogue of the same field.

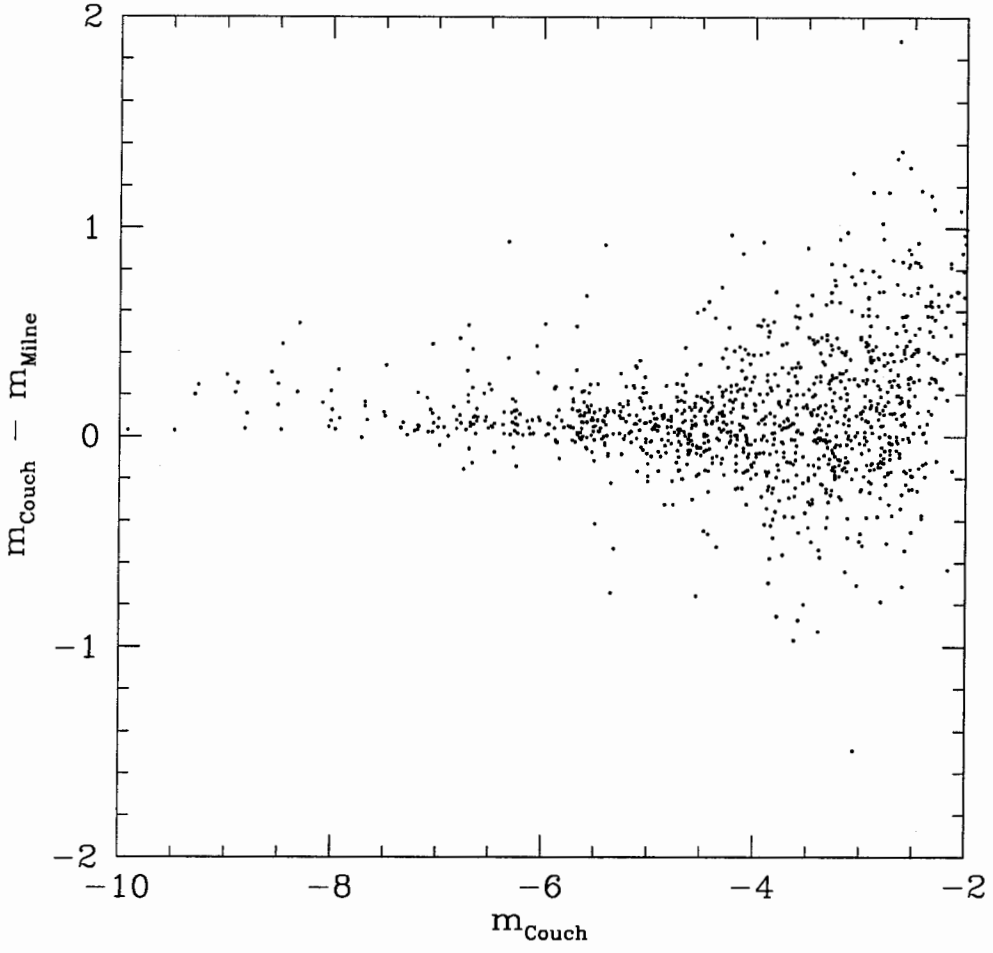

Figure 3.16: Difference in instrumental F606W magnitude as given in the HDF catalogue of Couch and as determined here, versus magnitude in the Couch catalogue. The magnitudes in the Couch catalogue have been transformed to the instrumental system of this work.

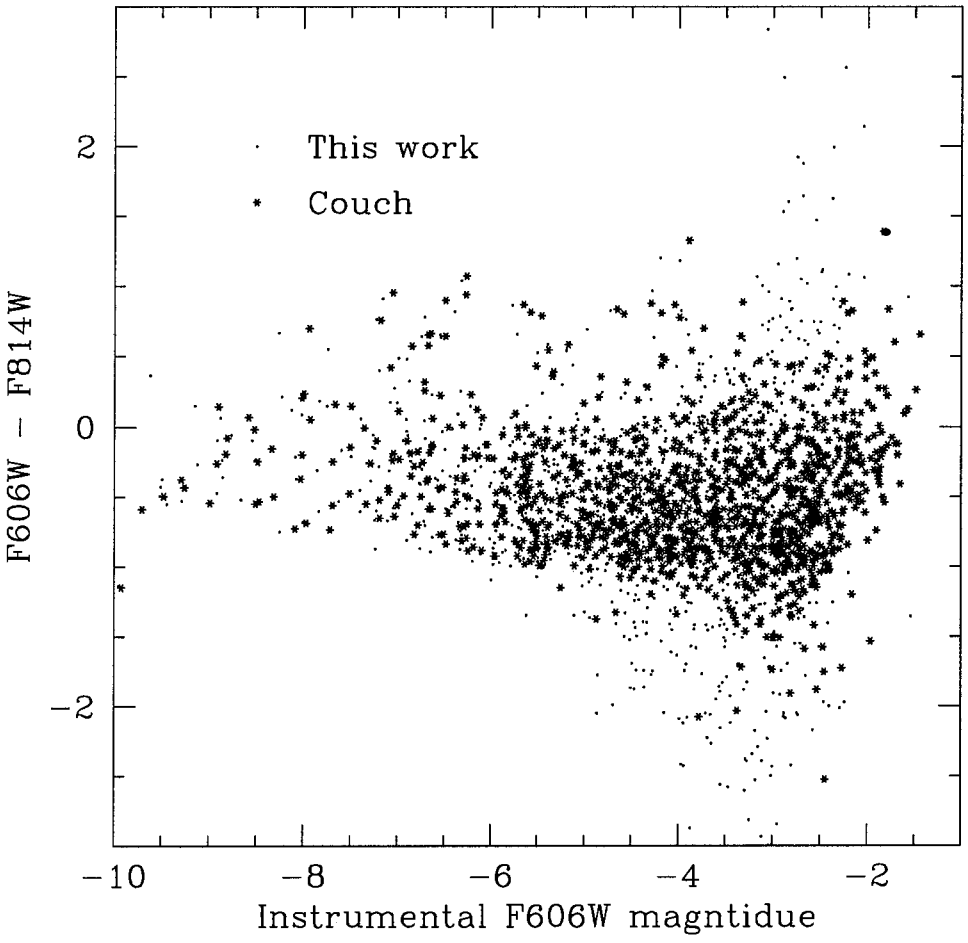

Figure 3.17: Colour magnitude diagrams as constructed with the HDF catalogue of Couch (stars) and with this catalogue (points). All magnitudes are instrumental in the system used in this work.

Table 3.4: SExtractor parameters tested with "add-galaxy" experiments

Parameter	Tested Values
DETECT_MINAREA	2 3 5
DEBLEND_MINCONT	0.002 0.005 0.008 0.05
DEBLEND_NTHRESH	16 32 64
CLEAN / CLEAN_PARAM	N / n/a Y / 0.1 Y / 0.5 Y / 1.0
WEIGHT_TYPE	MAP_VAR MAP_WEIGHT NONE
FILTER_NAME[a]	gauss_1.5_3x3 gauss_2.0_5x5 gauss_3.0_7x7 gauss_5.0_9x9
DETECT_THRESH[b]	2.5σ 3.5σ 4.0σ
BACKPHOTO_TYPE / BACKPHOTO_THICK	GLOBAL / n/a LOCAL / 5 LOCAL / 10 LOCAL / 15 LOCAL / 20
PHOT_AUTOPARAMS	2, 3 2, 4 3, 3 3, 4

[a] "gauss_N_1_N_2xN_2" is a $N_2 \times N_2$ pix Gaussian kernel with FWHM = N_1

[b] Each of these thresholds was tried with each of the convolution kernels above. For each kernel, the threshold was scaled by the appropriate noise reduction factor (see Section 3.1.2)

Table 3.5: Final set of SExtractor parameters

Parameter	Final Value
CATALOG_NAME	out.cat
PARAMETERS_NAME	out.param
FILTER_NAME	gauss_3.0_7x7.conv
DETECT_MINAREA	2
THRESH_TYPE	RELATIVE
DETECT_THRESH	0.86
ANALYSIS_THRESH	0.86
SEEING_FWHM	.2
DEBLEND_MINCONT	0.005
BACK_SIZE	20
BACK_FILTERSIZE	3
BACKPHOTO_TYPE	LOCAL
BACKPHOTO_THICK	10
PHOT_AUTOPARAMS	2.5, 3.5
PHOT_APERTURES	12.
PHOT_AUTOAPERS	0.0, 0.0
MAG_ZEROPOINT	0.
GAIN	7.1
PIXEL_SCALE	0.1
SATUR_LEVEL	50000.0
CHECKIMAGE_TYPE	SEGMENTATION
CHECKIMAGE_NAME	seg.fits
CATALOG_TYPE	ASCII_HEAD
DETECT_TYPE	CCD
FLAG_IMAGE	flag.fits
FILTER	Y
DEBLEND_NTHRESH	32
CLEAN	Y
CLEAN_PARAM	1.
MASK_TYPE	CORRECT
MAG_GAMMA	4.0
STARNNW_NAME	default.nnw
MEMORY_OBJSTACK	2000
MEMORY_PIXSTACK	100000
MEMORY_BUFSIZE	2048
VERBOSE_TYPE	NORMAL
WEIGHT_TYPE	MAP_WEIGHT, MAP_VAR
WEIGHT_IMAGE	weight.fits, var.fits
WEIGHT_THRESH	100, 99999
WEIGHT_GAIN	N

Chapter 4

Final Catalogue

In this chapter, we focus on the steps that led to the final galaxy catalogues. Specifically, we discuss:

- The dangers of globular cluster contamination and the creation, testing and use of a globular cluster mask
- The final detection and photometry process
- The omission of objects which fall in suspect regions of the images
- Star-galaxy separation
- Magnitude transformation, from the instrumental system used in this work to the standard Vega system
- Calculation of object colours, and their use as a further test of the globular cluster mask
- Comparison of the shape of final catalogue objects to known dwarf galaxies

Readers who are uninterested in these technical details may wish to skip directly to Chapter 5, where the luminosity function is presented.

4.1 Globular Cluster Contamination

The Coma images used in this study contain many globular clusters – in fact, it was for globular cluster studies that the data were originally obtained. Globular clusters are point sources at the distance of Coma, but two or more globular clusters (or other stellar objects) can appear to be a single extended object when close enough together, and thus can be misclassified as a cluster galaxy. This effect was discussed by Andreon & Cuillandre (2002), who noted that a large number of faint, extended objects in their CFHT images of the Coma cluster were in fact resolved into separate point sources in corresponding HST images. Globular cluster contamination must be dealt with to ensure a galaxy catalogue as free from contaminants as possible.

4.1.1 Globular Cluster Blends In The Initial Catalogue

This study uses HST images, and so should be much less affected by globular cluster contamination. To check this, the initial catalogue of Coma's chip 2 was marked on a display of the background-subtracted data image. Extended objects (those with $r_{petros} > 1.7$) were marked in a different colour than stellar ($r_{petros} \leq 1.7$) objects. 10 boxes of 100×100 pix (totaling 22.5% of the total chip area) were placed at random locations around the image. The objects in these representative regions were examined. The number of times two or more objects were detected as a single object was recorded, as well as the

number of times that blended objects appeared as extended objects.

Table 4.1 shows the results. The number of blends and blends that appear extended are scaled to the number expected on the whole chip, based on the relative areas of the 10 boxes and the total chip. The numbers can also be expressed as percentages: 6.7% of the objects in the Coma chip 2 catalogue are actually blends of 2 or more objects, and 10.2% of the "galaxies" in the catalogue are actually blends.

Table 4.1: Blended objects in the initial Coma chip 2 catalogue

	Blends	Blends that are "galaxies"
Number in boxes	21	13
Predicted number on whole chip	93	58
Percentage of total objects	6.7%	10.2%

This represents a considerable fraction of the total objects in the Coma catalogue. Although HST images do resolve globular clusters as separate point sources, the detection process employed here reblended them. The process had to be changed.

4.1.2 The Effects Of Filtering On Blending

The step in the detection process that is most likely to blend objects is the filtering step. Using a convolution kernel to enhance the detectability of objects also diminishes the contrast between objects. To test if the filtering step was indeed blending the globular clusters, SExtractor was run with a smaller convolution kernel (FWHM = 1.5 pix) to produce a new catalogue.

This new catalogue was marked on the background-subtracted data image in the same manner as the initial catalogue. Objects were again examined in 10 representative regions to determine how many blends were still present, and how many of those blends still appeared as extended objects. The results are in Table 4.2. Comparing these results to those in Table 4.1, it is obvious that using a small convolution kernel removes much of the globular cluster blend problem.

Table 4.2: Blended objects in the small kernel Coma chip 2 catalogue

	Blends	Blends that are "galaxies"
Number in boxes	5	4
Predicted number on whole chip	22	18
Percentage of total objects	1.3%	2.6%

However, convolving with a small kernel could introduce new problems. Small, centrally-concentrated galaxies – such as nucleated dwarfs – may be detected as stellar objects when convolved with a small kernel. The 10 representative regions were searched for galaxies resembling nucleated dwarfs. Four such objects were found, corresponding to 18 objects on the entire chip. All the objects were detected as extended objects. Using a small kernel does not negatively affect the detection of nucleated dwarf galaxies.

Using a small kernel could also deblend extended objects, possibly causing them to be detected as a series of stellar objects. To check this, the 10 regions were searched for instances of extended objects being split into multiple smaller objects by the small kernel. Three cases of this were found in

the representative regions, corresponding to 13 cases on the entire chip. In every case, all of the objects into which the extended object was split were also found as extended objects. This is to be expected: the type of extended object most likely to be split by a small kernel is a patchy, irregular galaxy. In such a galaxy, each of the patches still appears as an extended object, with a central peak surrounded by diffuse light. Convolving with a small kernel does cause some extended objects to be split, although they tend to be split into additional extended objects.

There is also the fact that "add-galaxy" experiments had indicated that a kernel with FWHM = 3.0 pix was the best choice for the detection and photometry of galaxies (Section 3.2.3). On the basis of this, and the three previous tests, it is obvious that the best detection strategy would be one that combined the use of a small kernel for separating blends of stellar objects, and a large kernel for the detection and photometry of galaxies.

4.1.3 The Globular Cluster Mask

To combine the best traits of the small and large convolution kernels, we decided to employ a two-pass detection and photometry procedure. The images would be convolved once with a small kernel to detect every point source as a separate object. The resulting catalogue would be separated into stellar and extended objects, and a stellar mask would be created. This mask would be applied in the second pass, done with a larger kernel, such that pixels associated with stellar objects in the first pass could not be detected on the second pass. The results of the second pass would then be the catalogue.

For this procedure to work, it is very important that the magnitude

and morphology of objects in the SExtractor catalogue do not depend too strongly on the convolution kernel used to create the catalogue. In the two-pass detection and photometry procedure, objects are defined as stellar on the basis of their magnitude and shape in the small kernel catalogue. If these objects would be found as extended in the large kernel catalogue, it is incorrect to block them from detection on the large kernel run.

To check that magnitude and morphology were sufficiently independent of the convolution kernel used, objects in the initial Coma catalogue were matched to those in the small kernel catalogue. Figure 4.1 shows a comparison of the r_{petros} values found using the small and large kernel. For instrumental magnitudes brighter than $\sim$ -6, there is good agreement between the r_{petros} values from the two catalogues.

Figure 4.2 shows a comparison of magnitudes between the two catalogues. The agreement is not as good: as discussed in Section 3.1.1, SExtractor magnitudes are dependent on the choice of convolution kernel. However, out to a magnitude of -6, there is a reasonable agreement between the magnitude as found by the large kernel and the magnitude as found by the small kernel. Therefore, the separation of the small kernel catalogue into stellar and extended objects should be done based on data from objects brighter than $\sim$ -6.

To construct a test globular cluster mask, SExtractor was run on the Coma chip 2 image using a convolution kernel with FWHM = 1.5 pix. The check image type was set to `SEGMENTATION`. The resulting catalogue was divided into stellar and extended objects based on the objects' r_{petros} values. As it is better to miss a few globular clusters than to misclassify extended ob-

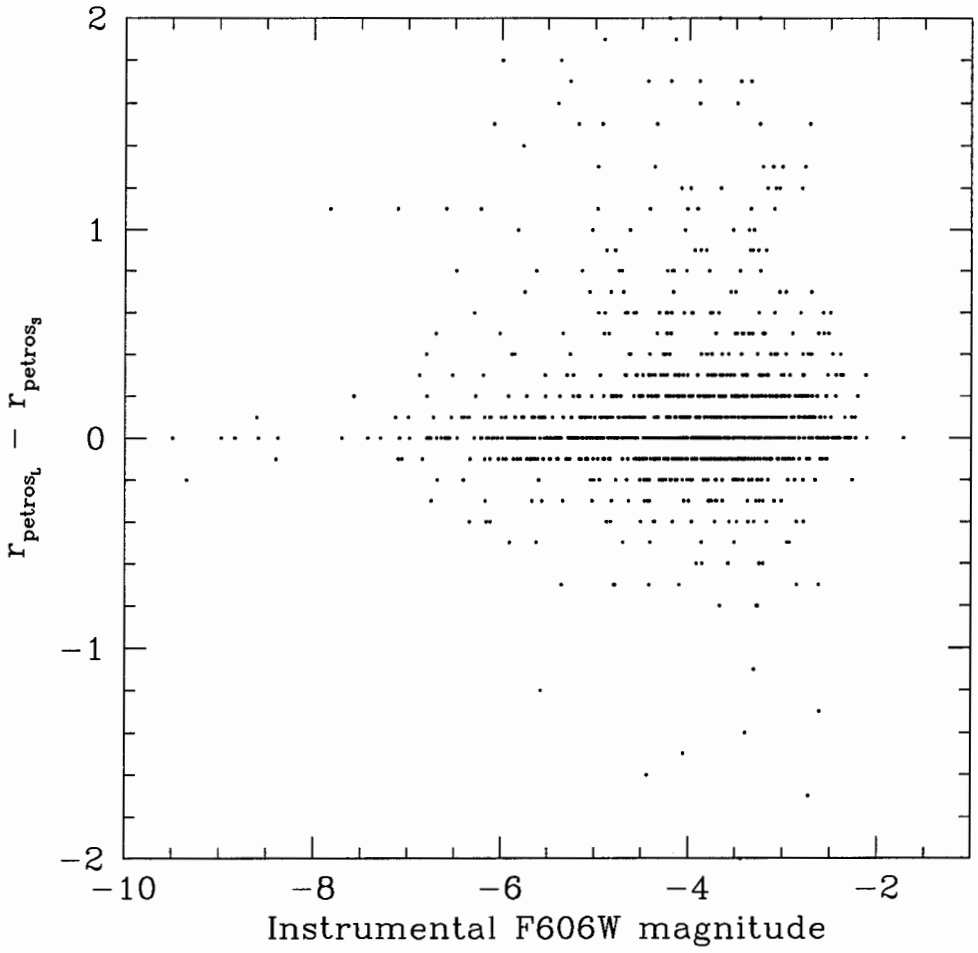

Figure 4.1: The difference between objects' Petrosian radii when detected by convolving with a large kernel and when detected by convolving with a small kernel, versus the objects' magnitudes. The large kernel was a Gaussian with FWHM = 3.0 pix; the small kernel was a Gaussian with FWHM = 1.5 pix. The magnitude is that from the catalogue made with the large kernel.

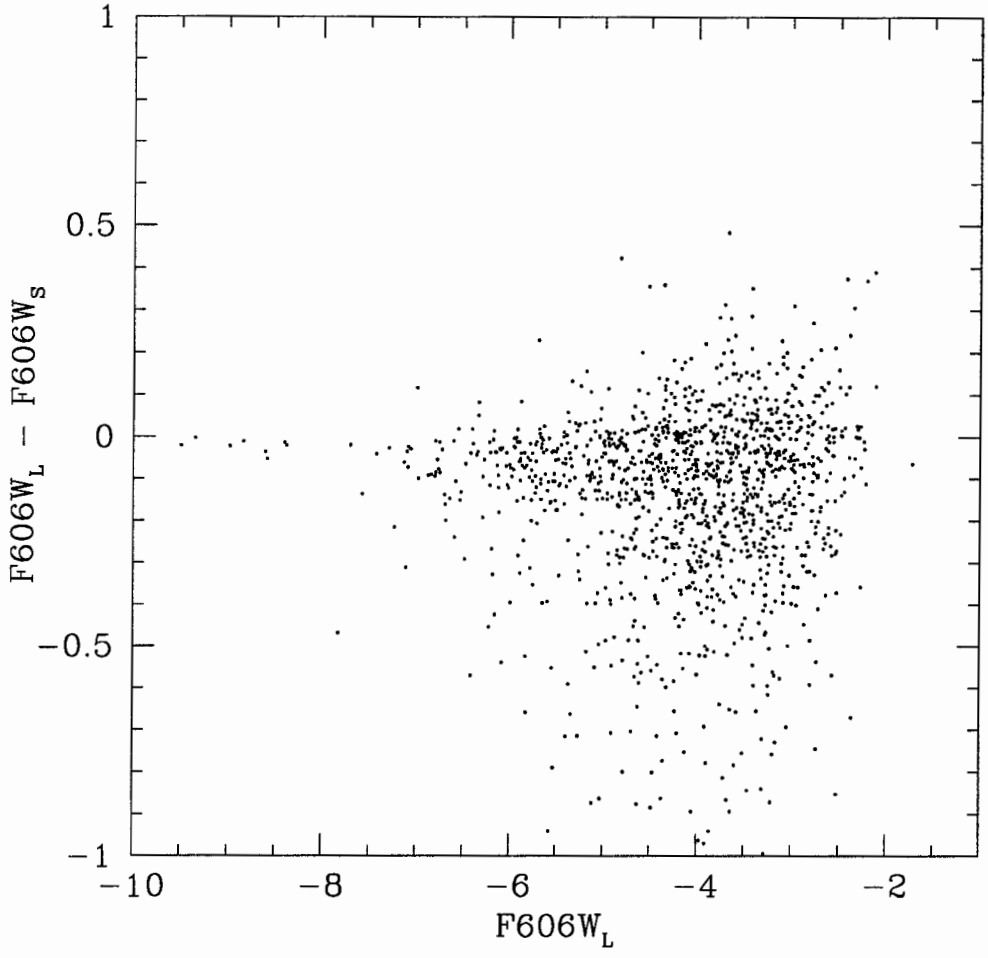

Figure 4.2: The difference between objects' magnitudes when detected by convolving with a large kernel and when detected by convolving with a small kernel, versus the objects' magnitudes. The large kernel was a Gaussian with FWHM = 3.0 pix; the small kernel was a Gaussian with FWHM = 1.5 pix. The magnitude on the y-axis is that from the catalogue made with the large kernel.

jects as stellar, the criterion for choosing stellar objects was relaxed slightly. Only those objects with $r_{petros} \leq 1.5$ were classified as stellar (Figure 4.3).

The list of stellar objects, the segmentation image, and the variance map used for detection – the background model – were then used to create the globular cluster mask. A small program cycled through the coordinates of the chip. At each xy location, it checked the segmentation image and the list of stellar objects to see if the pixel was associated with a stellar object. If so, it placed a value of zero at that pixel in the globular cluster mask. If the pixel was not associated with a stellar object, it was assigned a value equal to the inverse of the corresponding variance map pixel. The result is a weight image, where stellar pixels have a weight of zero, and all other pixels have a weight corresponding to their background value.

As discussed in Section 3.1.3, the SExtractor weight threshold feature does not work well for pixels with weights that are near the threshold. The globular cluster mask was therefore slightly modified. The image was multiplied by a constant such that the next smallest weight (after the weight of zero given to stellar pixels) was greater than 9999. Recall that this does not affect the workings of the weight image: SExtractor accepts weight images in units of *relative* weights and then scales them to the appropriate absolute level. However, it does mean that the weight threshold could be set to a value of 100. This value ensured that pixels associated with stellar objects (i.e., with weight = 0) would always be masked, while other pixels (i.e., with weight > 9999) would never be masked.

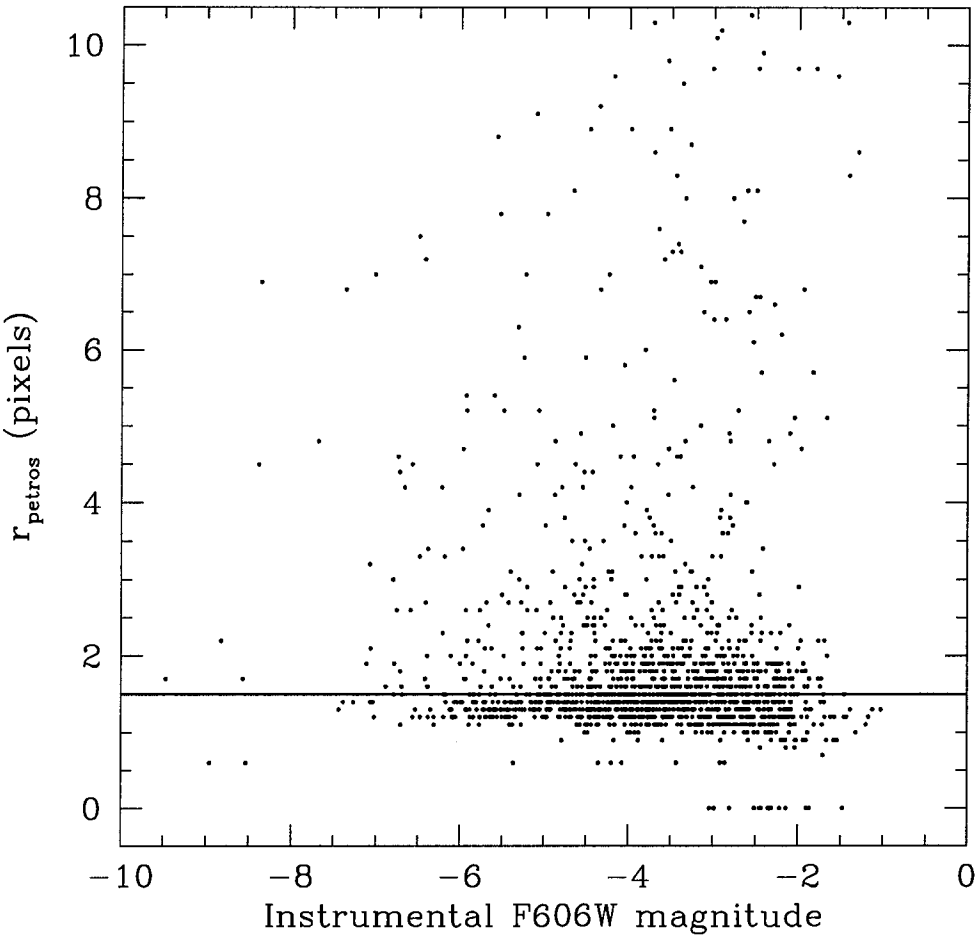

Figure 4.3: Petrosian radius versus instrumental F606W magnitude for objects detected by convolving with a Gaussian kernel of FWHM = 1.5 pix. Objects with $r_{petros} \leq 1.5$ are classified as stellar for the purposes of constructing the globular cluster mask.

4.1.4 Testing The Globular Cluster Mask

The sample globular cluster mask could now be tested. The first step was to find a way that the mask could be used with SExtractor. As mentioned in Section 3.1.3, the use of a weight threshold causes SExtractor to produce a large number of strange aperture magnitudes. Running SExtractor with the weight threshold set to 100 and the globular cluster mask as the weight image, 43% of the objects were found to have aperture magnitudes of -99.

To work around this problem, double image mode was used. The same Coma chip 2 image was used as the data image for both the detection and the photometry pass, but the weight image was changed: for the detection pass, the globular cluster mask was used with the weight threshold set to 100. For the photometry pass, the original variance map was used with no threshold. Using this technique, only 3% of the objects received aperture magnitudes of -99. Tests also showed that the Kron magnitudes found using this method matched those found when using the globular cluster mask and a weight threshold in single image mode (Figure 4.4).

The catalogue resulting from the double image mode run was marked onto the background-subtracted image, and the 10 representative regions were examined to determine how many blends were still present. In the 10 regions, 5 blends were found, corresponding to 22 remaining blends on the entire chip. Comparing this to the numbers in Table 4.1, $\sim$ 75% of the blends have been removed by using the globular cluster mask.

As a final check, the catalogue created with the globular cluster mask was matched to the initial Coma catalogue. The aperture magnitudes, Kron

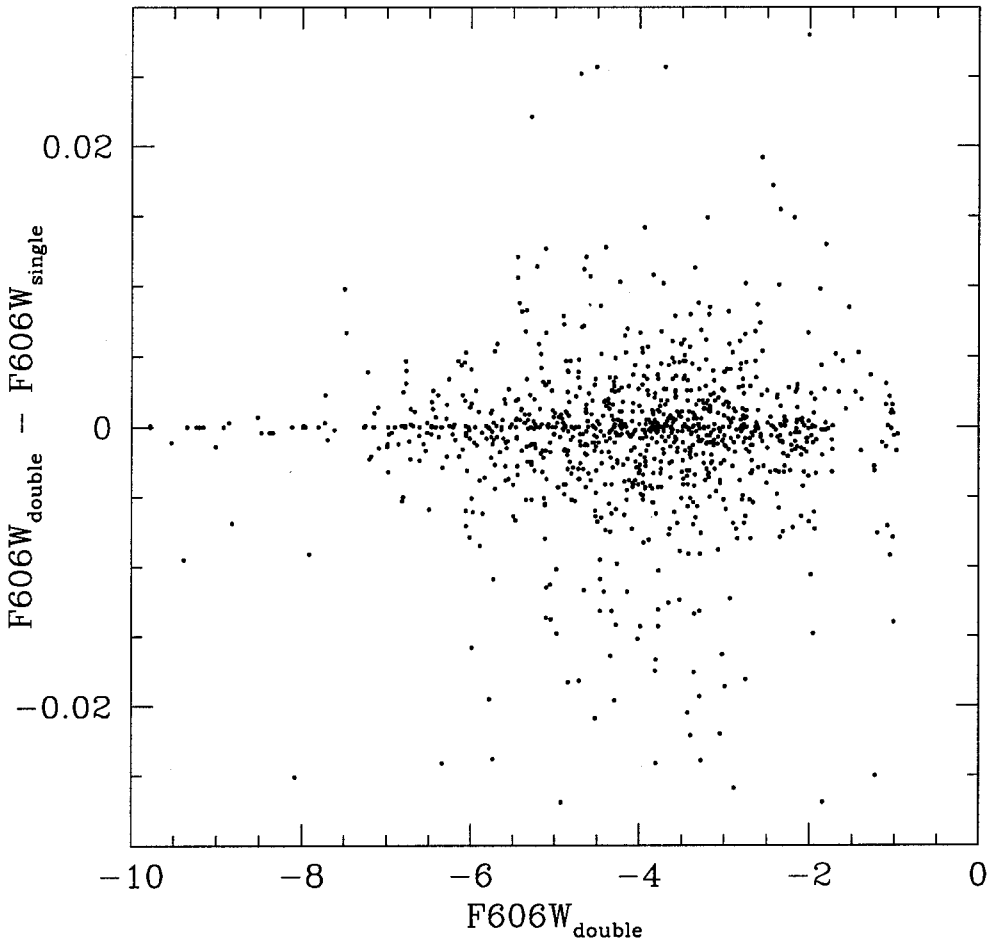

Figure 4.4: The difference between objects' magnitudes when detected in double image mode and when detected in single image mode, versus the objects' double image mode magnitudes. All magnitudes are instrumental Kron magnitudes.

magnitudes and r_{petros} values of matching objects were compared. Figures 4.5, 4.6 and 4.7 show the results: using the globular cluster mask does not substantially change the magnitude or morphology of detected objects.

4.2 Creating The Galaxy Catalogue

4.2.1 Globular Cluster Masks

Globular cluster masks were constructed for all three Coma F606W chips using the procedure outlined in Section 4.1.3. The number of pixels assigned a weight of zero on each chip was recorded.

Globular cluster masks were also created for the three HDF F606W chips. Of course, globular cluster contamination is not a problem in the HDF. However, this step was necessary to keep the processing steps identical between the data and control fields. The Coma background models for each chip were used to construct the globular cluster masks, as these were the variance maps that had been used as the HDF weight images in making the initial catalogue. The same selection criteria of $r_{petros} > 1.5$ was used to separate stellar and extended objects.

4.2.2 Detection And Photometry

SExtractor was run on the Coma images using the best set of parameters determined in Section 3.2. Double image mode was used for both F606W and F814W images. For F606W images, both the detection and the photometry image was set to the F606W image. For the detection pass, the globular cluster mask was used as a weight-type weight image and the weight threshold

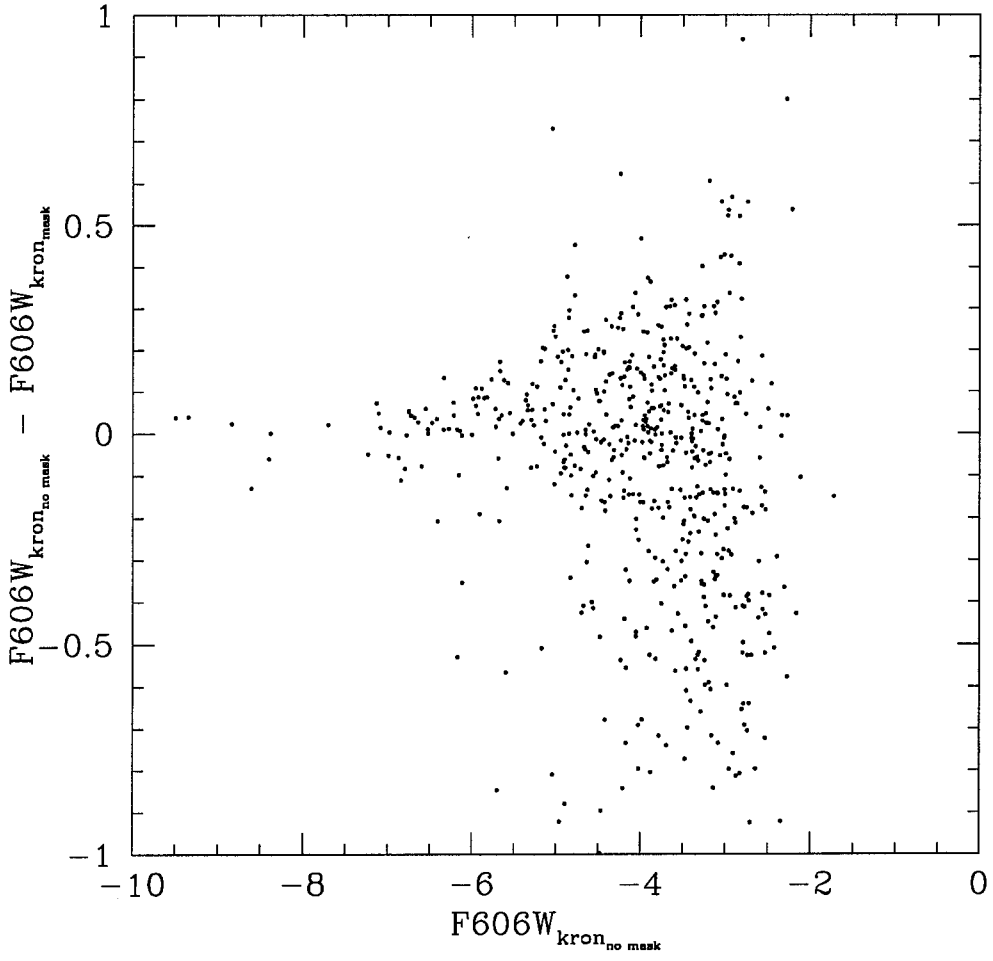

Figure 4.5: The difference between objects' Kron magnitudes when detected without a globular cluster mask and when detected with a globular cluster mask, versus the objects' no-mask magnitudes. All magnitudes are instrumental.

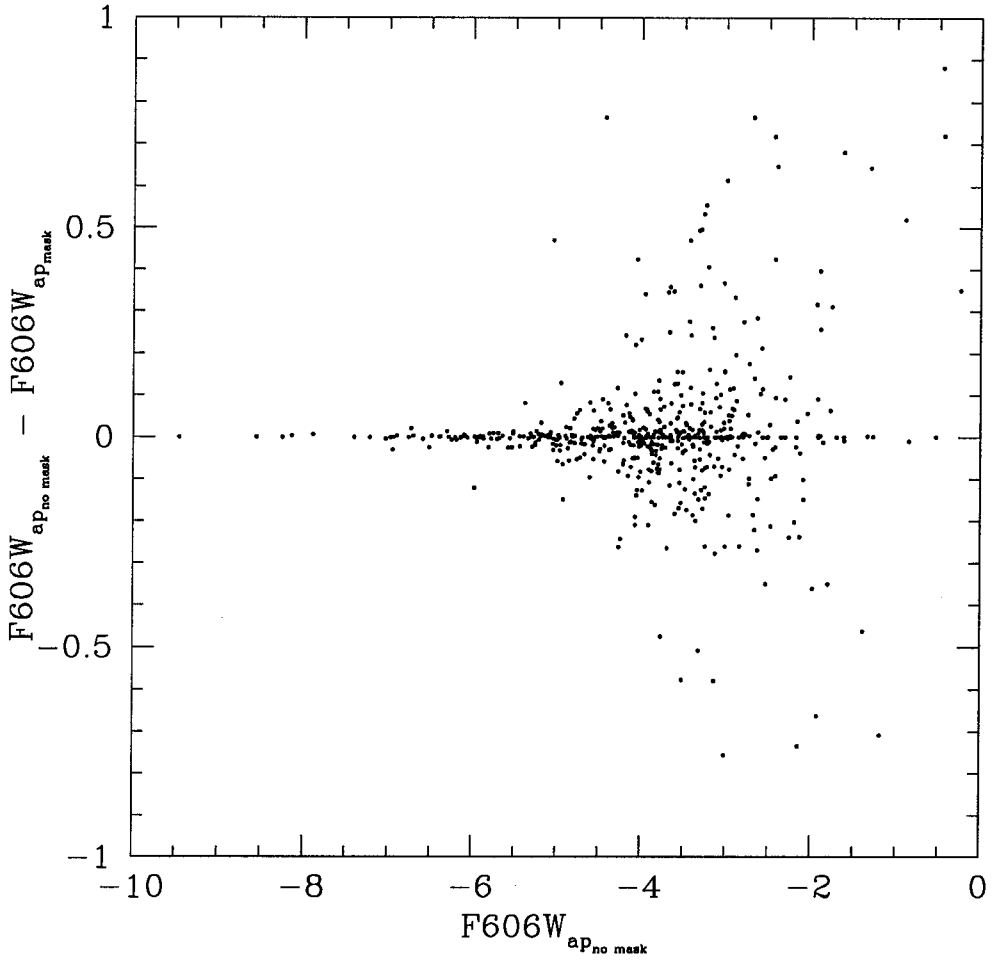

Figure 4.6: As Figure 4.5, but for aperture magnitudes.

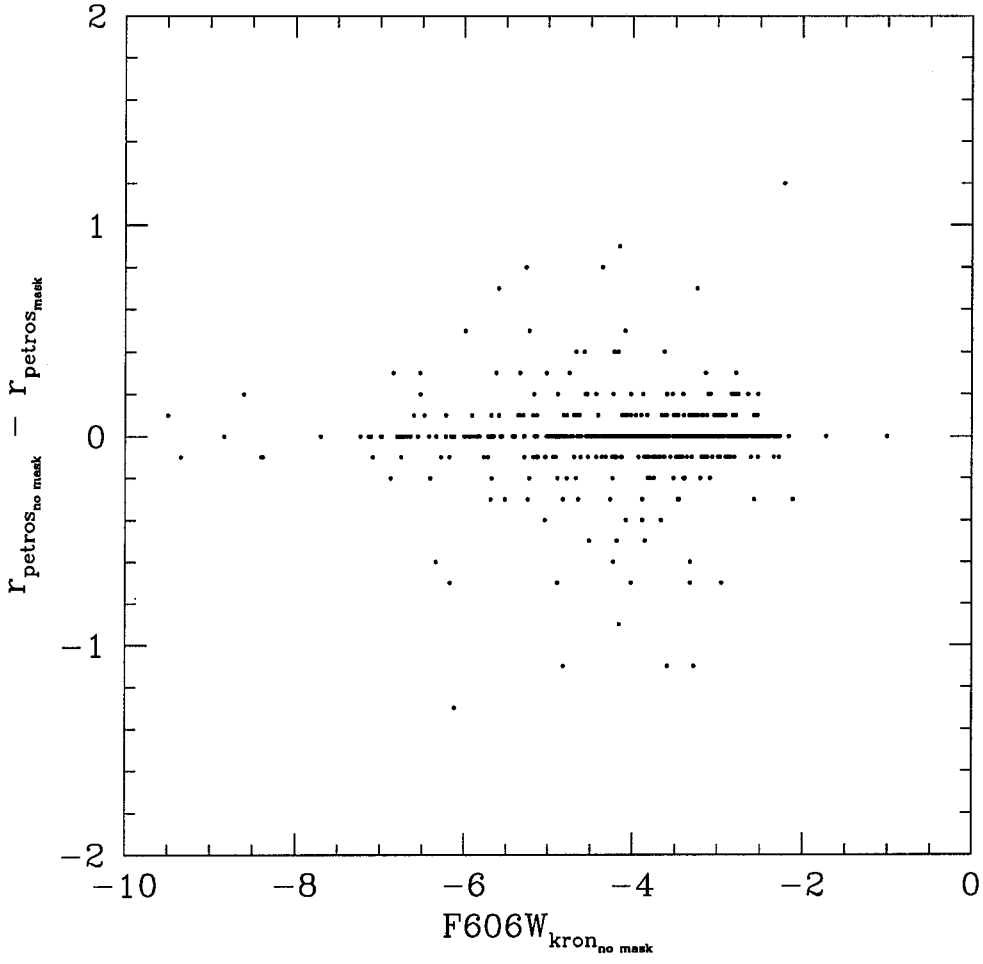

Figure 4.7: As Figure 4.5, but for Petrosian radii.

was set to 100. For the photometry pass, the background model was used as a variance type weight image and no weight threshold was set. For the F814W images, the F606W image and globular cluster mask were used with the weight threshold set to 100 for the detection pass. For the photometry pass, the F814W image and variance map were used with no weight threshold.

The degraded HDF images were background subtracted, to keep the processing steps identical between Coma and the HDF. SExtractor was run on the degraded, background-subtracted HDF fields using the same parameters and procedures as were used for Coma.

4.2.3 Omitting Objects

The Coma F606W catalogues were marked onto the background-subtracted images of each chip. Regions from which detections should not be trusted were defined around each of the large elliptical galaxies, based on the cosmetic appearance of the background-subtracted image, and on the presence of spurious or strangely-clustered detections. These omit regions were used on both the F606W and F814W catalogues to remove objects that fell in the bad regions.

The same omit regions were used to remove detections from the HDF catalogues.

4.2.4 Calculating And Comparing Colours

The aperture magnitudes from the F606W and F814W Coma images were combined to calculate colours for objects in the Coma F606W catalogues. If either the F606W or F814W aperture magnitude was missing, the colour

was assigned a value of -99.

These colours provided an opportunity to check that the globular cluster mask had in fact removed globular clusters. Globular clusters have a typical colour. As the Coma images contain mostly globular colours, the most common colour of objects in the initial Coma catalogue is that of a typical globular cluster. If the globular cluster mask removed random objects, it would remove objects of all colours. A histogram of object colour in the final catalogue would simply be a scaled-down version of the initial catalogue colour histogram. However, if the globular cluster mask truly removed globular clusters, objects of that peak colour should be preferentially missing from the final catalogue.

Histograms were made from the initial and final Coma catalogues of the instrumental colour of objects. To compare, the final catalogue histogram was scaled up by 1.72 to contain the same total number of objects as the initial catalogue histogram. Figure 4.8 shows the scaled final histogram compared to the initial histogram. The two curves do not match. The colour histogram of final catalogue objects is not simply a scaled down version of the colour histogram of initial catalogue objects – objects of the peak colour have been preferentially removed. The globular cluster mask is removing objects with the colour of globular clusters.

Note that this is a completely independent test of the globular cluster mask. The mask was constructed by looking at the shape of objects and masking all objects with the typical stellar shape of globular clusters. This test has shown that the masked objects also have the typical colour of globular clusters. The fact that both the shape and the colour of the masked

objects match those of globular clusters is a good indication that we are in fact removing globular clusters from the catalogue.

Colours were also calculated for objects in the F606W HDF catalogue, using aperture F606W and F814W magnitudes. Again, if a colour could not be computed, the colour was set to -99.

4.2.5 Star-Galaxy Separation

The Coma catalogue was then divided into stellar and extended objects. Because the globular cluster mask had removed so many stellar objects, a plot of r_{petros} versus magnitude did not show a strong enough stellar trend to decide on a criterion with which to separate stars and galaxies. Therefore, a plot was made of r_{petros} versus magnitude for the objects in the initial Coma catalogue (Figure 4.9). On the basis of this plot, objects with $r_{petros} > 1.55$ were kept as extended objects.

The HDF catalogue was separated into stars and galaxies based on the same criteria.

4.2.6 Magnitude Transformation

Finally, the catalogues were transformed from instrumental to standard magnitudes. To match the majority of the published literature on dwarf galaxies, we decided to transform magnitudes to the Vega system. Using the effective exposure time of the coadded Coma images (1300 s), and the HDF photometric zeropoints on the STScI webpage[1], the transformation from instrumental

[1]http://www.stsci.edu/ftp/science/hdf/logs/zeropoints.txt

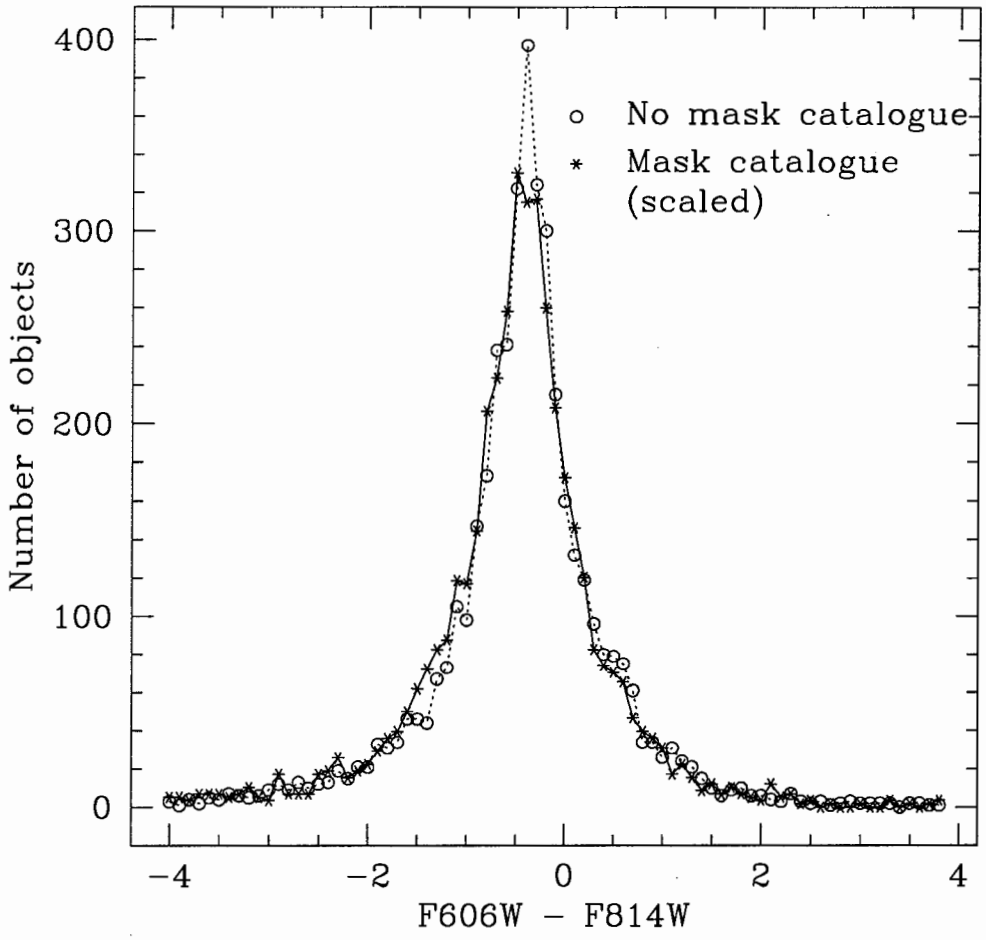

Figure 4.8: Histograms of the instrumental F606W-F814W colour of objects found without the globular cluster mask (open circles) and with the mask (stars). The histogram of objects found with the mask has been scaled up by 1.72 to have the same total counts as the no-mask histogram.

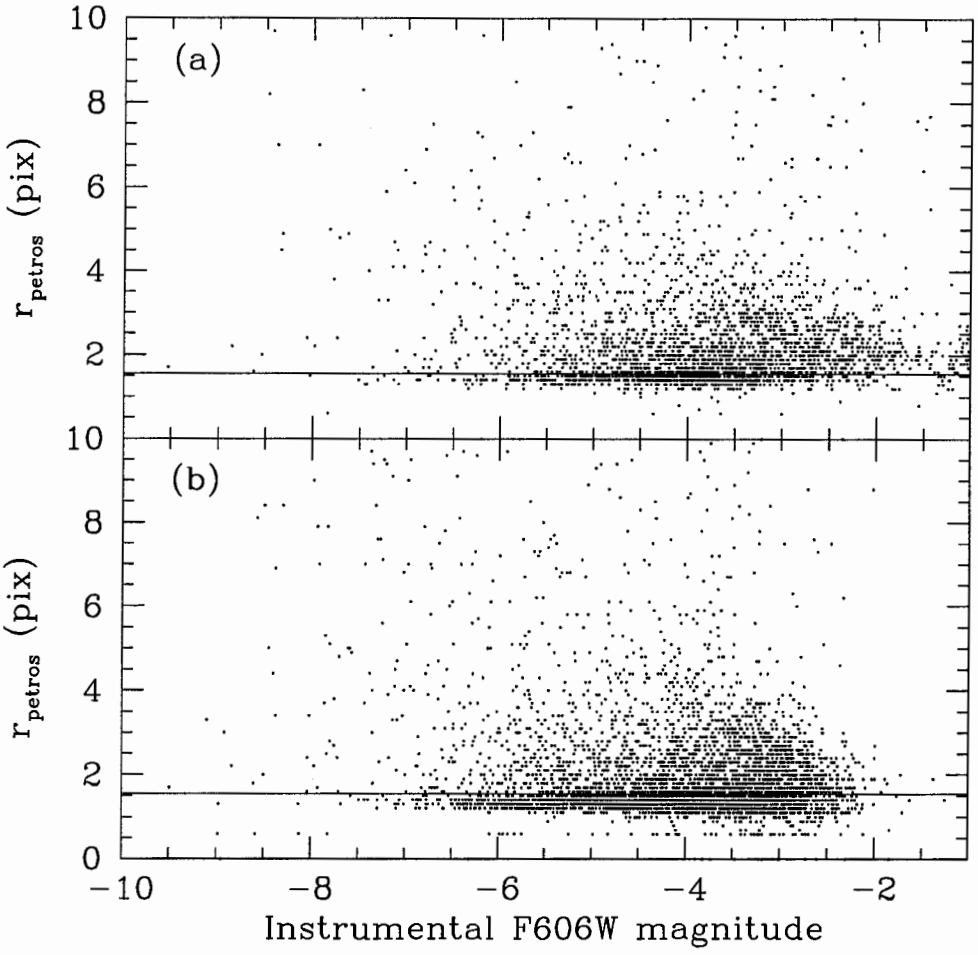

Figure 4.9: Simplified Petrosian radius versus instrumental F606W magnitude for objects in the Coma catalogue. Two different catalogues are shown: (a) the catalogue constructed with a globular cluster mask; and (b) the catalogue constructed without a globular cluster mask. As the stellar trend is not readily apparent in (a), the criterion to select galaxies was chosen based on the plot in (b) to be $r_{petros} > 1.55$. This line is shown on both plots.

to Vega magnitudes is given by

$$m(F606W)_{Vega} = m(F606W)_{inst} + 2.5 \times log(1300) + zp_{F606W,i} \quad (4.1)$$

$$m(F814W)_{Vega} = m(F814W)_{inst} + 2.5 \times log(1300) + zp_{F814W,i} \quad (4.2)$$

where $zp_{F606W,i}$ and $zp_{F814W,i}$ are the zeropoints for the i^{th} chip for the F606W and F814W filters, given by

$$zp_{F606W,2} = 22.91 \quad (4.3)$$

$$zp_{F606W,3} = 22.91 \quad (4.4)$$

$$zp_{F606W,4} = 22.89 \quad (4.5)$$

$$zp_{F814W,2} = 21.65 \quad (4.6)$$

$$zp_{F814W,3} = 21.66 \quad (4.7)$$

$$zp_{F814W,4} = 21.63 \quad (4.8)$$

The majority of the published literature on dwarf galaxies gives magnitudes in the Johnson-Cousins B or R band. Therefore, we decided to transform the F606W Vega magnitudes to these bands. The transformations from F606W and F814W to B and R were determined from the data of Fukugita et al. (1995). They used filter response functions and spectroscopic energy distributions of various galaxies to synthetically calculate the magnitudes of those galaxies in a number of photometric band systems, and so compute colour transformation laws among the various systems.

Figure 4.10a gives Fukugita et al.'s F606W-R colour plotted against their F606W-F814W colour; Figure 4.10b gives their F606W-B colour plotted against their F606W-F814W colour. Fitting a line to the data in both of

these plots gives the transformations from the Vega F606W magnitude and the Vega F606W-F814W colour to the Vega R or B magnitude:

$$m(R) = m(F606W) - 0.301(m(F606W) - m(F814W)) + 0.003 \quad (4.9)$$

$$m(B) = m(F606W) + 1.60(m(F606W) - m(F814W)) - 0.307 \quad (4.10)$$

This transformation relies on the colour of the object: an object with no F606W-F814W colour cannot be transformed into R or B. Some objects in the Coma and HDF catalogues did not have colours, due to SExtractor assigning the F606W or F814W aperture magnitude a value of -99. If these objects were to appear in the final transformed catalogue, a colour would have to be assigned to them. A histogram was made of the Vega F606W-F814W colour of objects in the final Coma catalogue (Figure 4.11). From this, a typical colour of 0.8 was determined. This value was assigned to all objects in both the Coma and HDF catalogues lacking a Vega F606W-F814W colour.

Using the transformations given in Equations 4.1 to 4.9, the magnitudes of objects in the Coma and HDF catalogues were transformed into the Vega R and B bands.

4.3 Comparison To Published Catalogues

The final Coma catalogue should contain mainly dwarf galaxies, with some contaminating background galaxies also included. To check the identity of the Coma catalogue objects, SExtractor was used to determine the half-light radius of each object in the catalogue. Dwarf galaxies are known to fall in a

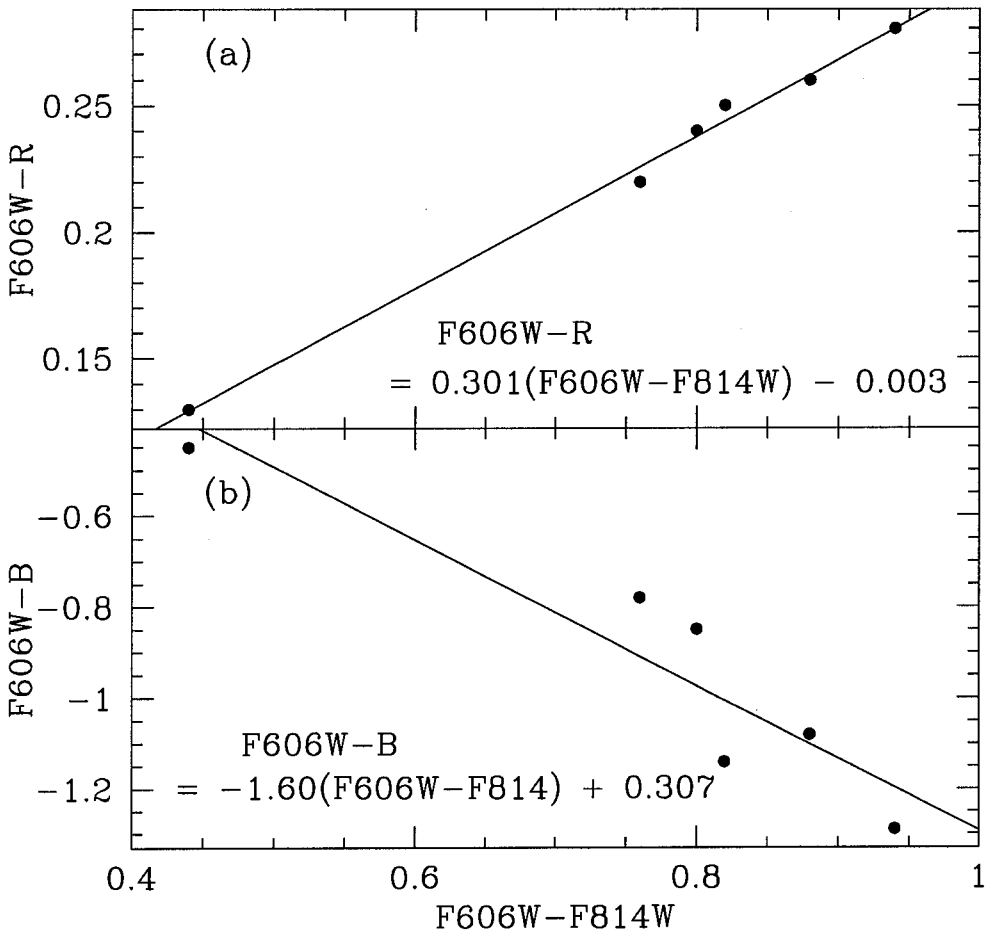

Figure 4.10: Vega colour-colour plots from the data in Fukugita et al. (1995). In (a), the F606W-R colour is plotted against F606W-F814W; in (b), F606W-B is plotted against F606W-F814W. Least square fits to the two data sets are shown.

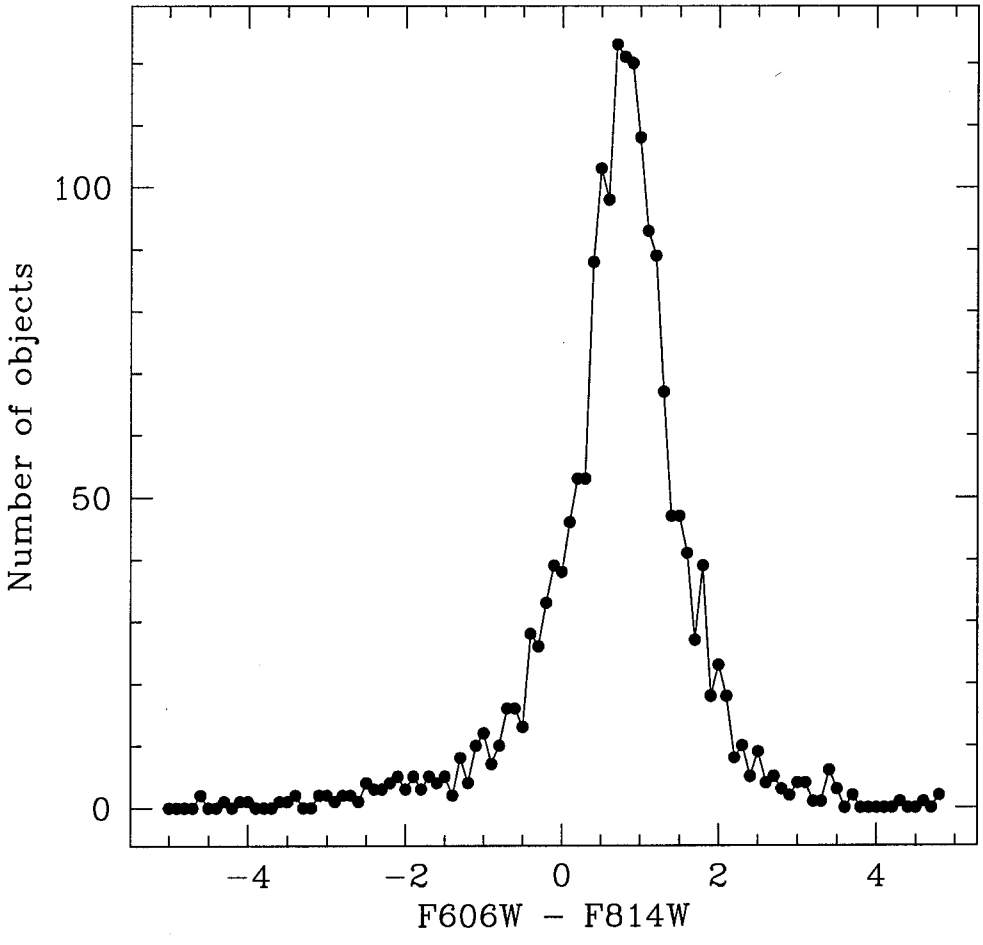

Figure 4.11: Histogram of the Vega system F606W-F814W colour of objects in the final Coma catalogue.

well-defined locus in half-light radius - magnitude space; if the Coma objects are truly dwarfs, they should occupy the same space as other known dwarf galaxies.

Six catalogues of dwarf galaxies in various environments were collected. Table 4.3 summarizes the contents of each catalogue and provides the references. The magnitude and half-light radius for each galaxy was converted into physical units: absolute Vega B magnitudes and kpc for the radii. When a distance to the group or galaxy was given with the catalogue, that value was adopted. For the Virgo cluster, a distance modulus of (m-M) = 30.99 was used (Kavelaars et al., 2000).

Table 4.3: Published catalogues of dwarf galaxies

Catalogue	Number of galaxies	Reference
Local Group	21	Bender et al. (1992)
M101	18	Bremnes et al. (1999)
M81	19	Bremnes et al. (1998)
Northern Field	17	Barazza et al. (2001)
Southern Field	24	Parodi et al. (2002)
Virgo Cluster	365	Binggeli & Cameron (1993)

The magnitudes and half-light radii for objects in our final Coma catalogue were also converted into physical units. Only those Coma objects with instrumental F606W magnitudes brighter than -4.25 were used, in anticipation of the limiting magnitude determined in Section 5.1. The distance modulus of Kavelaars et al. (2000), (m-M) = 35.05, was adopted.

For comparison, the magnitudes and half-light radii of Coma globular

clusters were also determined. These objects are those that were found to be stellar in the process of constructing the globular cluster mask, and so were blocked from future detection runs. This means that no colours were calculated for these objects. As discussed in Section 4.2.6, the transformation to Vega R or B magnitudes depends on an object's instrumental F606W-F814W colour. Therefore, to transform the magnitudes of the Coma globular clusters into Vega B magnitudes, a colour had to be assumed. As discussed in Section 4.2.4, globular clusters have a typical colour. Figure 4.8 shows this colour to be F606W-F814W $\sim$ -0.4. This colour was used to transform the instrumental magnitudes of the Coma globular clusters to Vega B magnitudes using Equation 4.9. Again, only objects with instrumental F606W magnitudes brighter than -4.25 were transformed. Finally, the magnitudes and half-light radii were converted into physical units, using the same distance modulus used for Coma dwarfs.

Half-light radius as a function of magnitude was then compared for all the objects. The results can be seen in Figure 4.12. Although the Coma galaxies are much fainter than most known dwarf galaxies, they fall in the same general locus in half-light radius - magnitude space. The distribution is understandably more scattered: background galaxies are also present in the Coma catalogue, which show a range of half-light radii at each magnitude. The Coma globular clusters, however, clearly occupy a very different region on the plot. From this plot, it is safe to conclude that the objects in the Coma catalogue are not globular clusters, and do have the same general morphology as known dwarf galaxies.

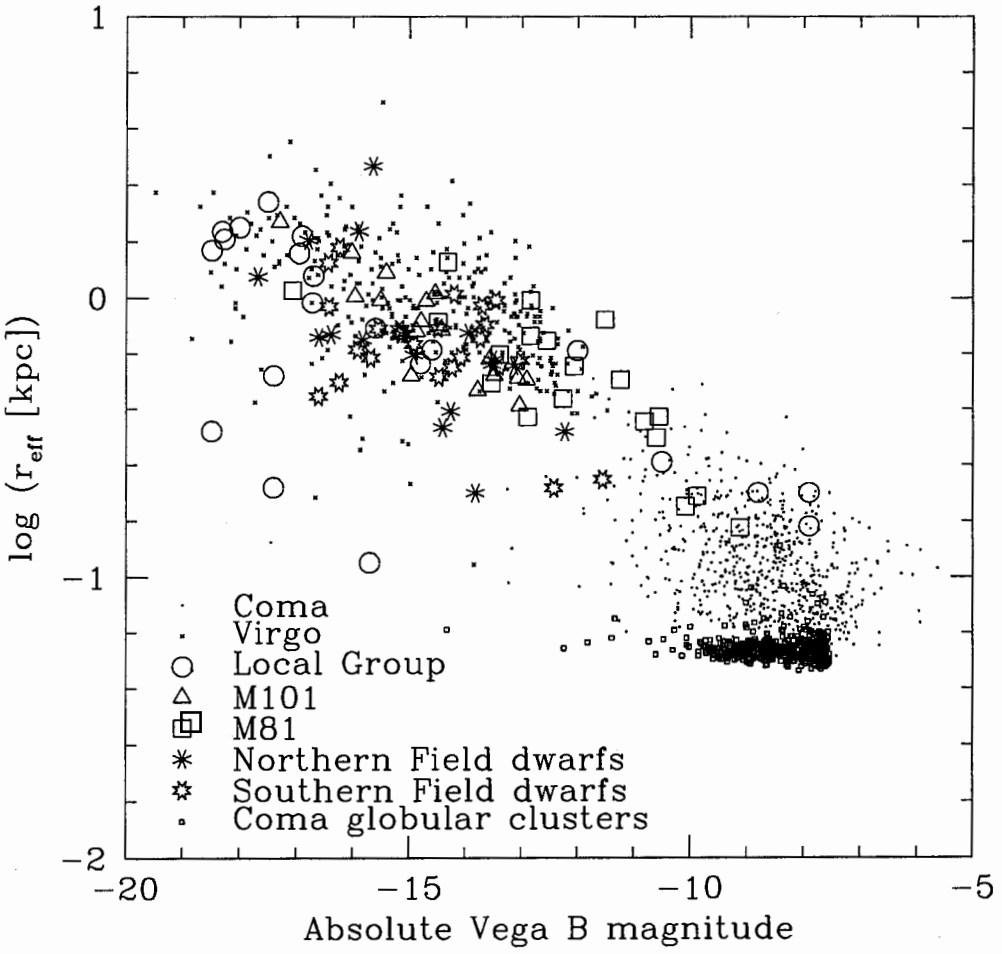

Figure 4.12: Logarithm of effective radius in kpc versus absolute Vega B magnitude for various dwarfs. Plotted are dwarfs from Coma (this work, points); the Virgo cluster (Binggeli and Cameron, 1993, "x"s); the Local Group (Bender et al., 1992, circles); M101 (Bremnes et al., 1999, triangles); M81 (Bremnes et al., 1998, squares); the Northern field (Barazza et al., 2001, stars); and the Southern field (Parodi et al., 2002, suns). Also shown for comparison are the Coma globular clusters identified in this work (small squares).

Chapter 5

Results

5.1 Limiting Magnitude

The limiting magnitude of a catalogue is the magnitude at which the catalogue becomes incomplete; that is, the magnitude at which a satisfactory percentage of objects can no longer be found. In most cases, of course, it is impossible to know how many objects are present but undetectable in an image. Therefore, the common method of determining limiting magnitude is to add fake objects to an image and attempt to retrieve them.

The use of the HDF as the control field in this work presented an interesting alternative method of determining the limiting magnitude. This number count based method is discussed in Section 5.1.1 and the limiting magnitude of the luminosity function is found; in Section 5.1.2, the limiting magnitude is compared to that found using the traditional method.

5.1.1 Determining Limiting Magnitude From HDF Number Counts

As part of this work, the HDF was degraded to match the detection characteristics of the Coma field, so that it could be used as a control field. As a result of this process, some objects in the field (purposely) became too difficult to detect. However, these objects are still detectable on the undegraded HDF. Creating a catalogue of objects from the undegraded HDF and comparing it to the catalogue obtained from the degraded HDF will indicate at what magnitude objects are lost due to the degrading. Since the degraded HDF has the same detection characteristics as the Coma field, this magnitude will then be the limiting magnitude.

The three WF chips of the undegraded HDF were trimmed to include the same physical area as the trimmed degraded HDF chips. SExtractor was run on the trimmed chips. The parameters were set to the same values used to create the Coma and degraded HDF catalogues, with the exception of parameters dependent on the physical size of the pixels. These parameters and with their original and new values are listed in Table 5.1. Note that the `BACK_FILTERSIZE` parameter was not changed; this parameter is in units of the background mesh, not in pixels. In addition to parameter changes, SExtractor was also set to not use a weight image while detecting on the undegraded HDF.

The omit regions used to remove detections falling in suspect regions from the Coma and degraded HDF catalogues were modified such that they would define the same physical areas on the undegraded HDF. They were

Table 5.1: Changes in pixel-dependent SExtractor parameters for detection on the undegraded HDF

Parameter	Value for detecting on degraded HDF	Value for detecting on undegraded HDF
FILTER_NAME	gauss_3.0_7x7.conv	gauss_7.0_11x11.conv
DETECT_THRESH	.86	.37
DETECT_MINAREA	2	12
BACK_SIZE	20	50
BACKPHOTO_THICK	10	25
PHOT_APERTURES	12	30
PIXEL_SCALE	0.1	0.04

then used to remove objects from the catalogue. The F606W and F814W aperture magnitudes were combined to generate instrumental colours for the objects in the catalogue, and magnitudes were transformed to the Vega B and R bands.

The undegraded HDF catalogue was then matched to the degraded HDF catalogue. The objects in the degraded HDF catalogue that had counterparts in the undegraded catalogue were used to construct a plot of number counts versus magnitude. This was compared to a similar plot made using all the objects in the undegraded HDF catalogue.

Figure 5.1 shows the results. It is easy to see that the degraded HDF catalogue begins to miss objects that are present in the undegraded HDF catalogue at a magnitude of R = 26.25. Therefore, the limiting magnitude of this work is R = 25.75.

Also shown on Figure 5.1 are the HDF number counts of Couch. These

were constructed from the data available at the Hubble Deep Field site at the Space Telescope European Coordinating Facility[1], after transforming the magnitudes into the Vega R band. The good agreement between Couch's number counts and the ones obtained here for the undegraded HDF are a reassurance that the undegraded HDF catalogue was constructed correctly.

Of course, a catalogue of the undegraded HDF will become incomplete at some magnitude as well. However, comparing the point at which the degraded and undegraded counts turn over on Figure 5.1, it is obvious that the undegraded HDF does not become incomplete for many magnitudes past the limiting magnitude of the degraded catalogue.

5.1.2 Comparison With Traditional Methods

The limiting magnitude can also be obtained through traditional "add-galaxy" experiments. In Section 3.3, the detection characteristics of Coma and the degraded HDF were compared. Figure 3.10 shows the average percentage of added objects that were recovered as a function of instrumental magnitude for both Coma and the HDF. This is then a plot of completeness versus magnitude for the two fields. As mentioned earlier, Coma and the HDF show somewhat different results. As this plot would indicate the Coma catalogue is less complete at faint magnitudes, the Coma results should be used to determine the limiting magnitude. According to "add-galaxy" experiments, the Coma catalogue is over 80% complete to an instrumental F606W magnitude of -4.25. This can then be considered the limiting instrumental magnitude of the luminosity function.

[1] http://www.stecf.org/hstprogrammes/hdf/catalogs/

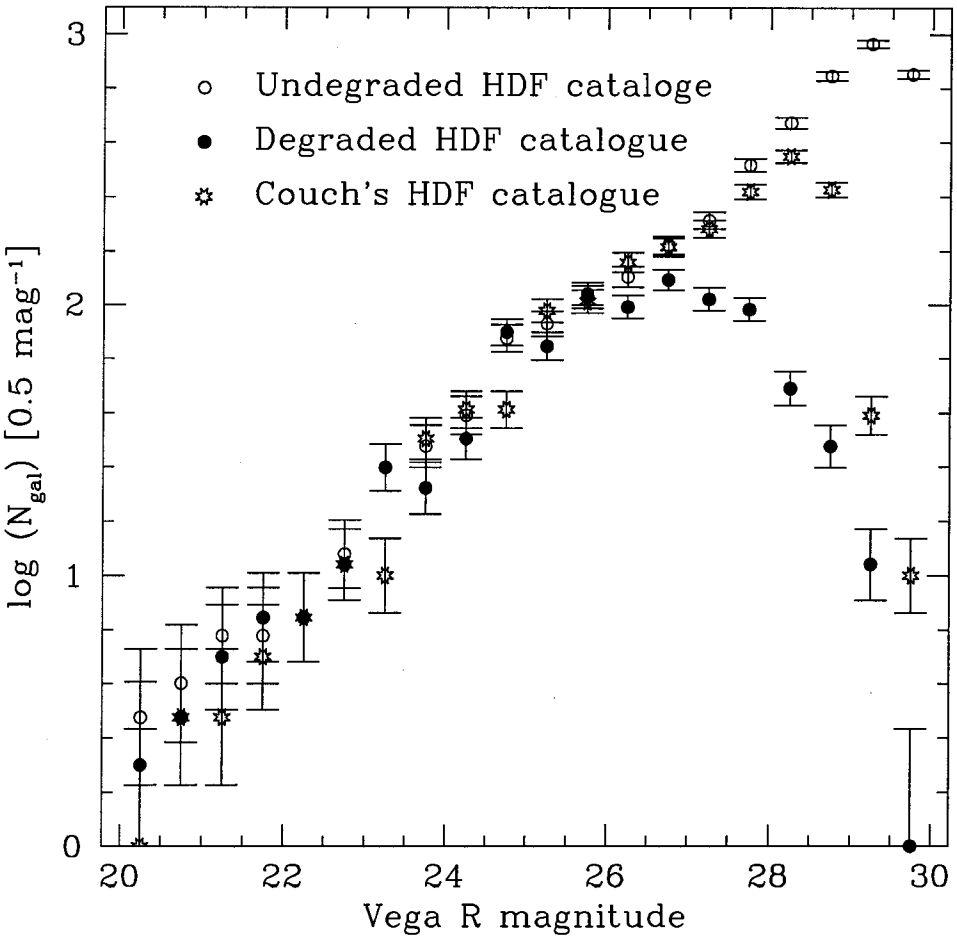

Figure 5.1: Logarithmic number counts in the Vega R band for the HDF. Counts are per 0.5 magnitude bin. Solid circles show the counts obtained from the undegraded HDF images; open circles show the counts from the degraded HDF images. Also shown are the HDF number counts of Couch (suns). All error bars show Poisson error only.

To compare this result to the limiting magnitude as determined by analyzing number counts, plots of number counts versus *instrumental* magnitudes were constructed from the degraded and undegraded HDF catalogues. Figure 5.2 shows the results. The degraded HDF number counts begin to deviate from the undegraded HDF number counts at an instrumental F606W magnitude of -3.75. Therefore, the limiting instrumental magnitude from number counts is also -4.25. The new number counts comparison method for determining limiting magnitude completely agrees with the traditional "add-galaxy" method.

5.2 Cosmic Variance

Background counts across the sky are not uniform. Due to clustering and large scale structure, background counts obtained from any given pointing will differ from those obtained with a different pointing more than can be accounted for by simple Poisson statistics. This "cosmic variance" affects the luminosity function in that a different choice of control field could result in a different statistical background subtraction.

To understand the effects of cosmic variance on the luminosity function, Peebles' (1980) expression for the variance of the count N of objects in a randomly placed cell was evaluated. This expression is given by

$$\langle (N - \eta\Omega)^2 \rangle = \eta\Omega + \eta^2 \int d\Omega_1 d\Omega_2 w(\theta_{12}) \tag{5.1}$$

where η is the mean density of objects on the sky, Ω is the size of the cell, and $w(\theta_{12})$ is the two-point angular correlation function.

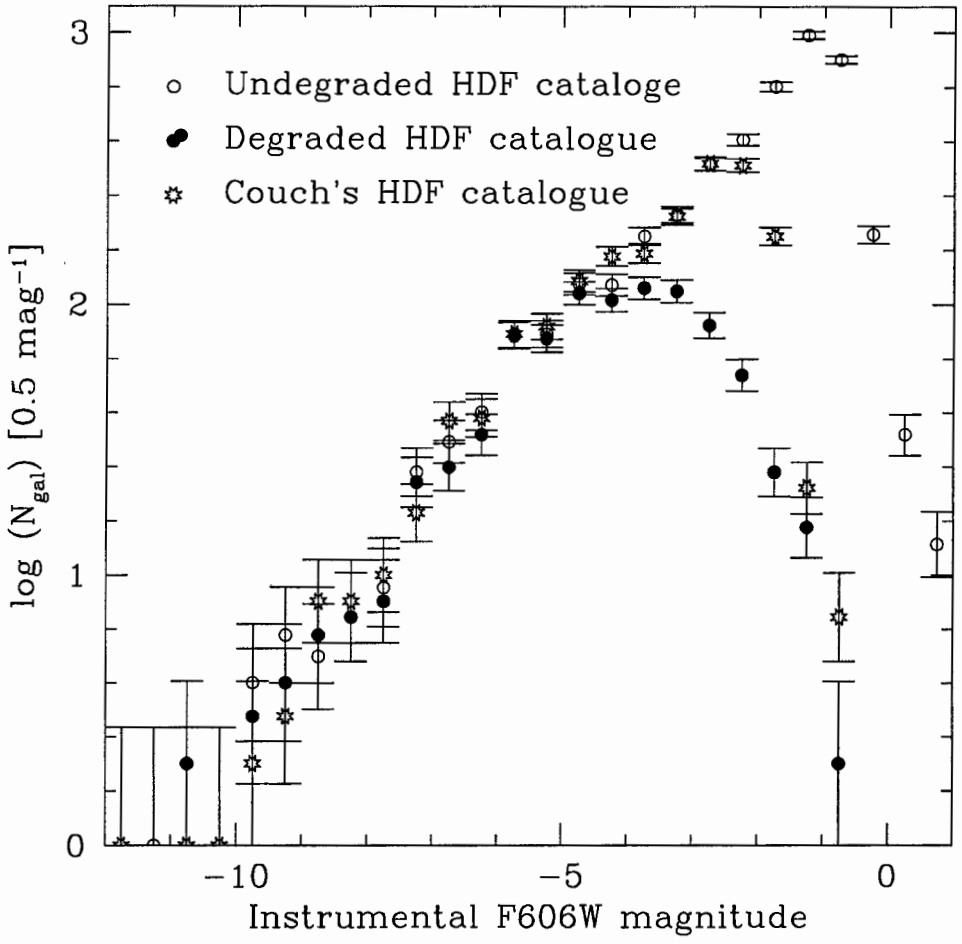

Figure 5.2: As Figure 5.1, but for instrumental F606W magnitudes. The magnitudes in Couch's catalogue were transformed to the instrumental system used in this work.

The number density of background galaxies, the size of the fields used here, and the galaxy angular correlation function can be substituted for η, Ω and $w(\theta_{12})$ respectively. Then, integrating Equation 5.1 over the field used here will give an estimate of the variance in the background counts of this work due to cosmic variance.

In Sections 5.2.1, 5.2.2 and 5.2.3 below, values are determined for $w(\theta_{12})$, η and Ω. In Section 5.2.4, these values are used to calculate the cosmic variance. Finally, in Section 5.2.5 the conversion of these results to error estimates for the luminosity function is discussed.

5.2.1 The Galaxy Angular Correlation Function

The two-point galaxy angular correlation function is usually parameterized by

$$w(\theta_{12}) = A_w \theta^{-\delta} \tag{5.2}$$

Values for A_w and δ were obtained from the work of Wilson (2003). She measured $w(\theta)$ as a function of apparent magnitude, colour and morphology, using data obtained for a weak lensing study of "blank" (i.e., background) fields. She found the data to be well described by a power law with index $\delta = 0.8$. Her best fit values for the logarithm of the amplitude of $w(\theta)$ at $1'$ as a function of Vega V band magnitude are given in Table 5.2.

The luminosity function of this work is in the Vega R band. Ideally, the angular correlation function coefficients would be expressed as a function of Vega R band magnitudes as well. However, a search of the literature yielded no studies of the two-point galaxy angular correlation function in the Vega

Table 5.2: Amplitude of the galaxy angular correlation function as a function of magnitude[a]

Bin	$\log_{10} A_w(1')$
$21.0 < V \leq 22.0$	-1.13 ± 0.06
$22.0 < V \leq 23.0$	-1.49 ± 0.05
$23.0 < V \leq 24.0$	-1.72 ± 0.05
$24.0 < V \leq 25.0$	-2.26 ± 0.09

[a] Taken from Wilson (2003)

R band. Wilson's data can be applied to this work if a constant Vega V-R colour is assumed for the objects in the Coma and HDF catalogues. As determined in Section 4.2.6, the typical Vega F606W-F814W colour of objects in the final Coma catalogue is 0.8. Using the data of Fukugita et al. (1995), the relationship between the Vega V-R colour and the Vega F606W-F814W colour is given by

$$m(V) - m(R) = 0.59(m(F606W) - m(F814W)) + 0.05 \qquad (5.3)$$

Using this, a typical Vega V-R colour of 0.5 is obtained. Therefore, Wilson's coefficients for the bin $m_1 < V \leq m_2$ will be applied to the bin $m_1 - 0.5 < R \leq m_2 - 0.5$ in this work.

5.2.2 Number Density Of Background Galaxies

To determine the mean density of background galaxies, the compilation of published field number counts available from the Durham Cosmology Group[2]

[2]http://star-www.dur.ac.uk/ nm/pubhtml/counts/counts.html

was used. Because counts in the Vega V band were not available, the counts in the Vega R band were used. Following the reasoning used above, the counts in the bin $m_1 - 0.5 < R \leq m_2 - 0.5$ were used as the counts in the bin $m_1 < V \leq m_2$.

A plot of the compiled R band counts is shown in Figure 5.3. The average count in each 0.5 mag bin was determined. This value was divided by 3600 to get the average number of galaxies per arcmin^2 in each bin. As the cosmic variance was to be calculated in 1.0 mag bins, the results from pairs of bins were added together. The final resulting number densities are shown in Table 5.3.

Table 5.3: Number density of background galaxies as a function of magnitude

Bin	Galaxy number density (arcmin^{-2} mag^{-1})
$21.0 < V \leq 22.0$	1.15
$22.0 < V \leq 23.0$	2.69
$23.0 < V \leq 24.0$	6.37
$24.0 < V \leq 25.0$	14.6

5.2.3 The Field Size

The total field size of this work can easily be calculated by adding together the trimmed areas of the three WF chips. Using the chip dimensions given in Table 2.2 and the WF resolution of 0.1″/pix, the total area of the field was calculated to be 4.62 arcmin^2.

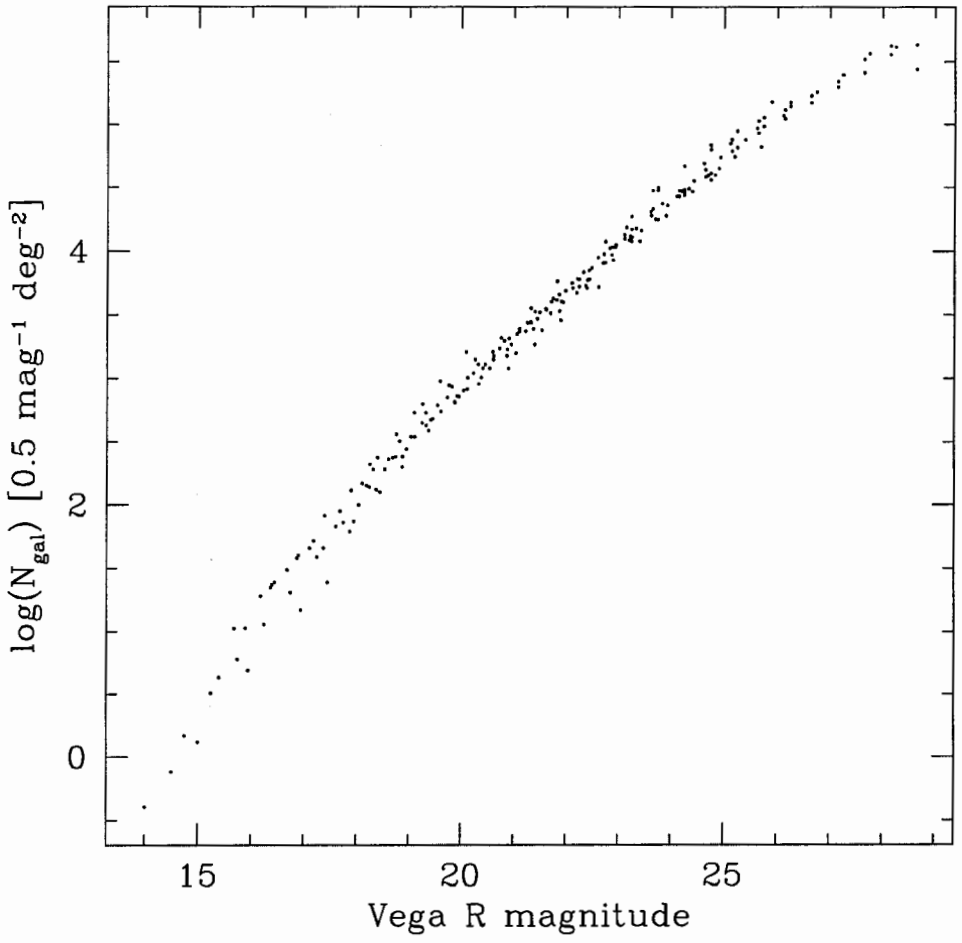

Figure 5.3: Histogram of the logarithmic Vega R band counts of field galaxies. Counts are expressed per 0.5 magnitude bin and per deg^2. The data are taken from the compilation of published field counts available from the Durham Cosmology Group.

5.2.4 Calculating The Cosmic Variance

The cosmic variance was calculated as a function of magnitude in four 1.0 mag bins. A small program was written to evaluate the integral in Equation 5.1 for each bin using the values obtained for $w(\theta_{12})$, η and Ω. The program evaluated the integral as a sum, dividing the field into a number of area elements and evaluating each pair in turn. As an approximation to the true field shape, the field was assumed to be shaped like an "L", formed from three equal sized squares with no spaces between them, such that the total area of the three squares equaled 4.62 arcmin^2. The distance between points was calculated as the linear distance; this approximation is valid for fields as small as the one used here. The program was run a number of times for each bin with an increasing number of area elements until the calculation of the cosmic variance asymptotically approached the true value. The inherent symmetries in the field shape were used to speed the calculation.

Table 5.4 shows the results. For each 1.0 mag bin, the cosmic variance and standard deviation are given.

Table 5.4: Cosmic variance as a function of magnitude

Bin	Cosmic variance	Standard deviation
$20.5 < R \leq 21.5$	7.72	2.78
$21.5 < R \leq 22.5$	21.3	4.62
$22.5 < R \leq 23.5$	48.4	6.96
$23.5 < R \leq 24.5$	96.2	9.81

5.2.5 Converting Cosmic Variance To Bin Error

The luminosity function in this work is expressed in terms of the number of galaxies in 0.5 mag bins. Therefore, the error in each 1.0 mag bin due to cosmic variance, as given in Table 5.4, needs to be divided between two 0.5 mag bins. The logical way to do this is to divide based on the relative population of the two bins: if the first bin has $1/n^{th}$ of the total background counts between the two bins, $1/n^{th}$ of the error due to cosmic variance will go to that bin. If σ_T is the standard deviation due to cosmic variance for a 1.0 mag bin, and n_1 and n_2 are the background counts in two 0.5 mag bins, then the standard deviation due to cosmic variance in the two 0.5 mag bins is given by

$$\sigma_1 = \sigma_T \frac{n_1}{n_1 + n_2} \tag{5.4}$$

$$\sigma_2 = \sigma_T \frac{n_2}{n_1 + n_2} \tag{5.5}$$

The error due to cosmic variance only applies to background counts. For the HDF, all the counts in each bin are due to background galaxies. For Coma, however, only a fraction of the counts in each bin are from background galaxies. The exact number for each bin is unknown; statistically, it is assumed to be the same as the HDF count for the corresponding bin. Therefore, the error due to cosmic variance for each Coma bin equals the error due to cosmic variance for the corresponding HDF bin.

Once calculated, the error due to cosmic variance in each 0.5 mag bin could be added in quadrature to the Poisson error. However, the calculation here was performed only to serve as an estimate of the effects of cosmic

variance on the slope of the luminosity function. A full treatment would have to account for the correlations in error between magnitude bins introduced by cosmic variance. Due to the simplistic approach and many assumptions used here, we choose to present the errors due to Poisson statistics and cosmic variance separately.

5.3 The Luminosity Function

5.3.1 Control Field Number Counts

The final degraded HDF catalogue for each chip was used to determine number counts in 0.5 mag bins, from the magnitude of the brightest catalogue galaxy to the limiting magnitude of R = 25.75, determined in Section 5.1.1. The Poisson error in the counts in each bin was calculated.

Because the control field number counts are subtracted directly from the data field number counts, it is important that the control field counts are scaled to the area of the data field. In this work, the Coma and HDF chips were trimmed to the same area. However, the *effective* areas of the Coma field and the HDF field still differ. This is because of the globular cluster masking process: many more objects were masked in the Coma field than in the HDF field, and so the Coma field has an effectively smaller area in which catalogue objects can be found.

The effective area of a chip can be defined as

$$A_{eff} = xdim \times ydim - n_o - n_m \tag{5.6}$$

where $xdim$ and $ydim$ are the x and y dimension of the chip, n_o is the number of pixels falling in the omit regions (those regions originally occupied by large

elliptical galaxies), and n_m is the number of pixels masked by the globular cluster mask.

Values for $xdim$ and $ydim$ can be found in Table 2.2. To determine n_o for each chip, a small program was written to cycle through the dimensions of each chip and count the number of xy positions that fall inside the elliptical omit regions of that chip. Values for n_m for each of the Coma and HDF chips were obtained from the program which created the globular cluster masks.

Table 5.5 shows the values of $xdim, ydim, n_o$ and n_m for each chip, along with the calculated effective area and scale factor. The scale factor is simply calculated as the effective area of a Coma chip divided by the effective area of the same HDF chip.

Table 5.5: Effective area of Coma and HDF chips

Chip	$xdim$ (pix)	$ydim$ (pix)	n_o (pix)	n_m (pix)	A_{eff} (pix)	Scale factor
Coma chip 2	741	757	99 618	7908	453 411	
HDF chip 2	741	757	99 618	516	460 803	.984
Coma chip 3	756	736	110 734	5776	439 906	
HDF chip 3	756	736	110 734	371	445 311	.988
Coma chip 4	739	741	86 848	8684	452 067	
HDF chip 4	739	741	86 848	511	460 240	.982

For each chip, the HDF number counts and Poisson error in each bin were multiplied by the scale factor. The number counts from the three chips were then combined to determine the total control field number counts. Using Equation 5.4, the error due to cosmic variance was also calculated for each bin where possible.

5.3.2 Data Field Number Counts

The final Coma catalogue for each chip was used to determine number counts in 0.5 mag bins, from the magnitude of the brightest catalogue galaxy to the limiting magnitude. The Poisson error in the counts in each bin was calculated. The number counts from the three chips were then combined.

Equation 5.4 was used to calculate the error due to cosmic variance in each bin where possible. As discussed in Section 5.2.5, the total scaled HDF number counts were used to determine values for n_1 and n_2.

5.3.3 The Luminosity Function

The control field and data field number counts were finally combined to construct the luminosity function. Table 5.6 (at the end of this chapter) shows the data that were used to construct the luminosity function. For the Coma field and the HDF field, the number of galaxies in each 0.5 magnitude bin, the Poisson errors in the count and the errors due to cosmic variance in the count are given. Also shown is the logarithmic surface density of objects in each 0.5 mag bin, defined as the number of galaxies per deg^2, and the Poisson error in that value. The values given for the HDF field are after scaling to match the effective area of the Coma field.

Figure 5.4 plots the number counts of the Coma field and the scaled HDF field. The error bars shown include Poisson error only. As can be seen, cluster counts do not start to dominate over the background until the last two bins. This is due to the small field size surveyed here: bright cluster galaxies are so rare that either zero or $\sim$ 1 are expected in a field this size. The

bright galaxies that are detected in the Coma field are due to background contamination, and so their numbers match those found in the background-only field.

The scaled HDF counts were subtracted from the Coma counts. Table 5.6 shows the counts, surface density and errors obtained in each 0.5 mag bin. The Poisson error and error due to cosmic variance in each bin was calculated as

$$\sigma = \sqrt{\sigma_{Coma}^2 + \sigma_{HDF}^2} \tag{5.7}$$

where σ_{Coma} and σ_{HDF} are the error in the Coma and HDF number counts in that bin. These counts are then the luminosity function; this is shown in Figure 5.5. Again, the error bars show only Poisson error.

Figure 5.6 shows the luminosity function in logarithmic form. The linear errors in each bin were converted to logarithmic form using the properties of the logarithm: if n is the count in a given bin and dn the error in that count, then

$$\begin{aligned} \log(n \pm dn) &= \log\left(n\left(1 \pm \frac{dn}{n}\right)\right) \\ &= \log(n) + \log\left(1 \pm \frac{dn}{n}\right) \end{aligned} \tag{5.8}$$

Again, the error bars show only Poisson error.

5.3.4 The Slope Of The Luminosity Function

The luminosity function is generally parameterized by the Schechter function:

$$n(L)dL \sim \left(\frac{L}{L^*}\right)^{\alpha} \exp\left(\frac{-L}{L^*}\right) \; d\left(\frac{L}{L^*}\right) \tag{5.9}$$

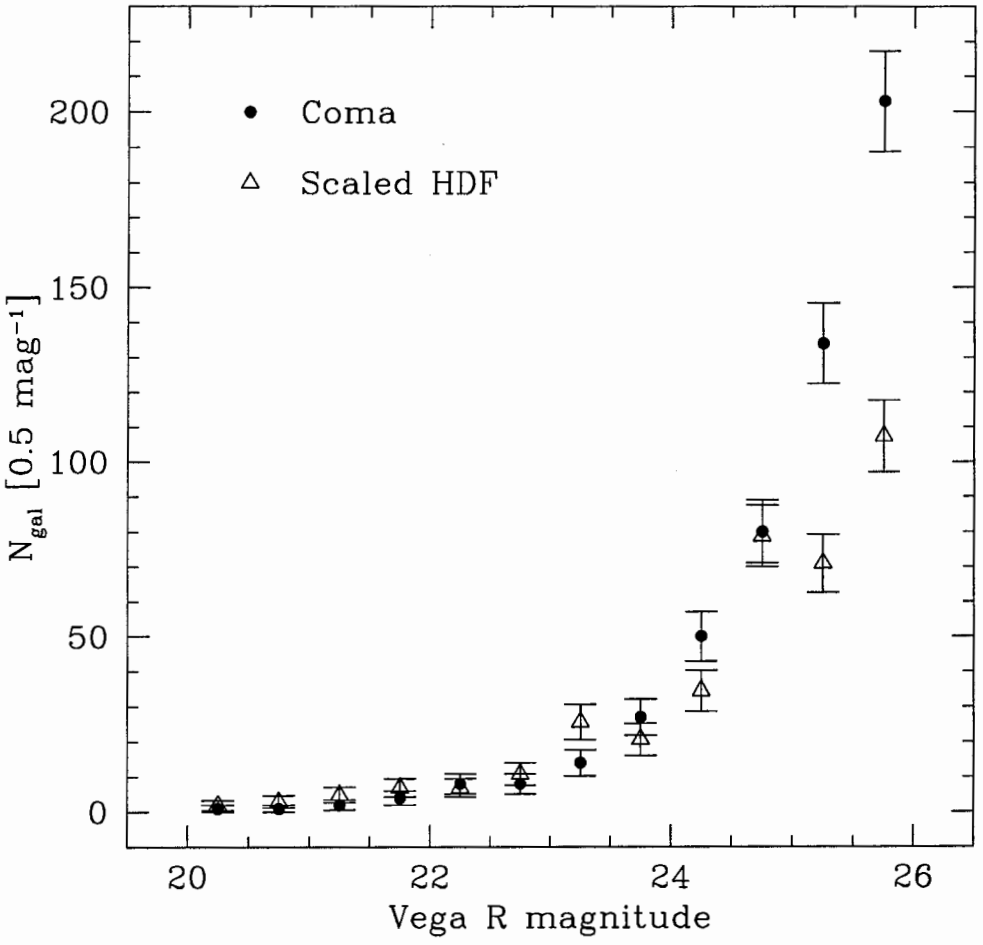

Figure 5.4: Number counts from the Coma cluster field (circles) and the HDF field (triangles). Error bars include Poisson error only.

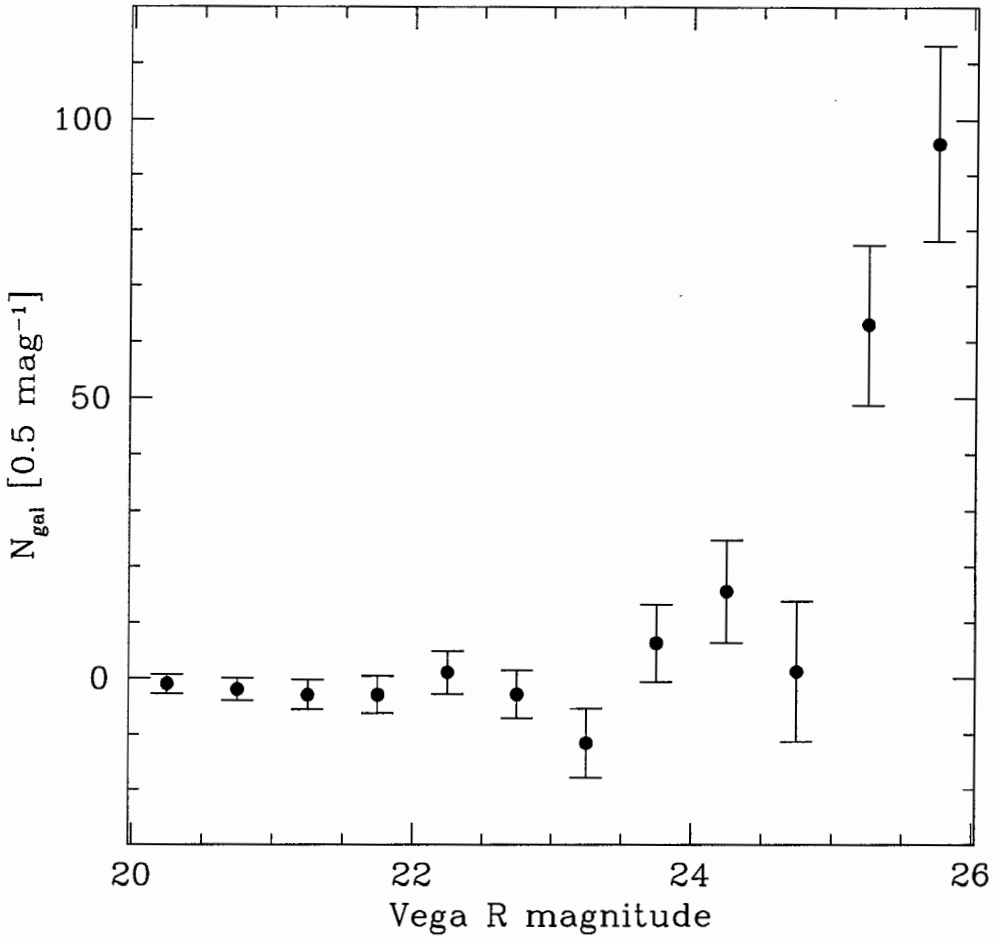

Figure 5.5: Vega R band luminosity function of the Coma cluster, expressed as linear counts per 0.5 magnitude bin. Error bars include Poisson error only.

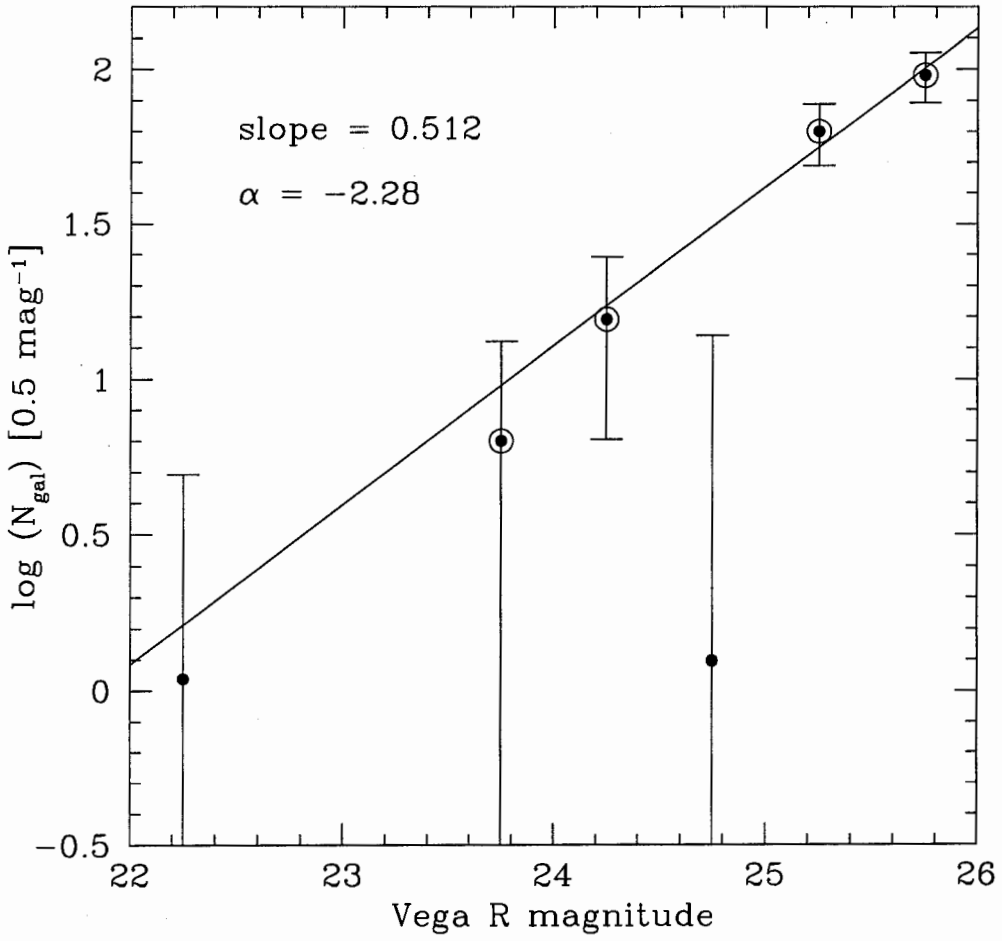

Figure 5.6: Vega R band luminosity function of the Coma cluster, expressed as logarithmic counts per 0.5 magnitude bin. Error bars include only Poisson error. The line is a weighted least squares fit to the four circled points.

This can be written in terms of magnitude as

$$\phi(M)dM \sim 10^{-0.4(\alpha+1)M} e^{-10^{0.4(M^*-M)}} \tag{5.10}$$

At faint magnitudes, this approaches

$$\phi(M)dM \sim 10^{-0.4(\alpha+1)M} \tag{5.11}$$

The slope of the faint end of the luminosity function is generally characterized by the parameter α. To find the relationship between α and the slope, take the logarithm of both sides:

$$\log(\phi(M)) \sim -0.4(\alpha+1)M \tag{5.12}$$

Therefore, the slope m of the luminosity function in logarithmic space is related to the parameter α by

$$\alpha = -1 - \frac{m}{0.4} \tag{5.13}$$

The error in α can be obtained from the error in the slope by differentiating Equation 5.13:

$$\begin{aligned} \frac{dm}{d\alpha} &= -0.4 \\ d\alpha &= \frac{dm}{-0.4} \end{aligned} \tag{5.14}$$

To find a value for α, the logarithmic form of the luminosity function was fit to a straight line using the IRAF task `polyfit`. This task performs a weighted least-squares fit using Bevington's `REGRES` routine (1969). The fit can be weighted in a number of ways; for this work, the weighting was set to

`instrumental`, which set the weight of each point to the squared inverse of the error in that point.

The logarithmic error bars calculated using Equation 5.8 are asymmetric. Because `polyfit` will only accept a single error for each point, symmetric logarithmic errors were calculated from the linear errors based on the idea of a derivative: if n is again the count in a bin, then taking the derivative of the logarithm of the count gives

$$\frac{d}{dn}(\log n) = \frac{\log e}{n} \tag{5.15}$$

Then, the error in the logarithmic count can be expressed as

$$d\log n = \frac{\log e}{n} dn \tag{5.16}$$

Figure 5.6 shows the fit to the luminosity function. Only the four circled points were included in the fit. The fitted value for α was found to be $\alpha = -2.28 \pm 0.34$. Including all six points in the fit gives $\alpha = -2.29 \pm 0.33$. In both cases, the error on the value of α was determined from the calculated error on the slope using Equation 5.14.

Table 5.6: The luminosity function[a]

	Coma				
m_R	N	ΔN_p	ΔN_{CV}	log σ	Δlog σ
20.25	1.00	1.00	–	2.89	2.89
20.75	1.00	1.00	1.04	2.89	2.89
21.25	2.00	1.41	1.74	3.19	3.04
21.75	4.00	2.00	2.31	3.49	3.19
22.25	8.00	2.83	2.31	3.79	3.34
22.75	8.00	2.83	2.07	3.79	3.34
23.25	14.00	3.74	4.89	4.04	3.46
23.75	27.00	5.20	3.68	4.32	3.61
24.25	50.00	7.07	6.13	4.59	3.74
24.75	80.00	8.94	–	4.79	3.84
25.25	134.00	11.58	–	5.02	3.95
25.75	203.00	14.25	–	5.20	4.04
	Scaled HDF				
m_R	N	ΔN_p	ΔN_{CV}	log σ	Δlog σ
20.25	1.97	1.39	–	3.18	3.03
20.75	2.96	1.71	1.04	3.36	3.12
21.25	4.92	2.20	1.74	3.58	3.23
21.75	6.90	2.61	2.31	3.73	3.31
22.25	6.91	2.61	2.31	3.73	3.31
22.75	10.81	3.26	2.07	3.93	3.40
23.25	25.60	5.02	4.89	4.30	3.59
23.75	20.68	4.51	3.68	4.21	3.55
24.25	34.45	5.82	6.13	4.43	3.66
24.75	78.75	8.81	–	4.79	3.84
25.25	70.91	8.36	–	4.74	3.81
25.75	107.32	10.28	–	4.92	3.90
	Coma - HDF				
m_R	N	ΔN_p	ΔN_{CV}	log σ	Δlog σ
20.25	-0.97	1.71	–	–	–
20.75	-1.96	1.98	1.48	–	–
21.25	-2.92	2.62	2.46	–	–
21.75	-2.90	3.29	3.27	–	–
22.25	1.09	3.85	3.27	2.93	3.48
22.75	-2.81	4.32	2.92	–	–
23.25	-11.60	6.26	6.92	–	–
23.75	6.32	6.88	5.20	3.69	3.73
24.25	15.55	9.16	8.67	4.08	3.85
24.75	1.25	12.55	–	2.99	3.99
25.25	63.09	14.28	–	4.69	4.05
25.75	95.68	17.57	–	4.87	4.14

[a] For descriptions of columns, see Section 5.3.3

Chapter 6

Discussion

6.1 Comparison To Other Work

Many studies of the luminosity function of the Coma cluster have appeared in the literature. Table 6.1 (at the end of this chapter) shows a selection of these, indicating their passband, field size, pixel scale, limiting magnitude, and final value for α. As can be seen, this work achieves a much better resolution and a much fainter limiting magnitude than any previous study. This is of course due to the use of Hubble Space Telescope data in this work. For the same reason, this study also covers a much smaller field than any previous work.

Looking at the values for α determined by the various studies, it would appear that there is little agreement on the "true" slope of Coma's luminosity function. Examining the actual data, however, can sometimes give a clearer picture of the situation. To better compare the various determinations of Coma's luminosity function, a composite luminosity function was constructed.

Background-subtracted number counts for the Coma cluster were obtained from the literature. Only those papers which determined the luminosity function in the R band were used, in order to avoid the uncertainties associated with magnitude transformation. In most cases, the number counts were taken from the published tables provided by the authors. For the work of Adami et al. (2000), number counts were determined from their Figure 6.

Mobasher et al. (2003) examined three regions of the Coma cluster; their 1584 arcsec^2 "Coma1" field is the only one that overlaps the field studied in this work, and so only those counts were used. Both their Deep Spectroscopic Survey + Shallow Spectroscopic Survey counts and their Extended Shallow Spectroscopic Survey counts were used.

In cases where completeness corrected counts were provided (Bernstein et al., 1995; Secker et al., 1997; Adami et al., 2000; Mobasher et al., 2003), those counts were used. In the case of Bernstein et al. (1995), their total corrected counts (as opposed to their "resolved" corrected counts) were used.

Counts that were reported in full magnitude bins (Bernstein et al., 1995; Adami et al., 2000; Mobasher et al., 2003) were divided in half and the errors scaled accordingly. The counts and errors were then all scaled to an area of one deg^2, using the field size reported by the authors. Note that no other normalization was performed: this might seem odd, considering the range in field sizes covered by the various studies under consideration. Generally speaking, a survey covering a larger area will include more of the low density cluster outskirts and so find a lower surface density of galaxies; such a survey would have to be scaled differently than a survey examining only the high density cluster core.

The scale radius of the Coma cluster is 320 kpc for $H_o = 70$ km/s/Mpc (Łokas and Mamon, 2003); at the distance of Coma, this corresponds to $\sim 10.8'$, and so a survey covering just the cluster core would cover 369 arcmin^2. Assuming a reasonable surface density profile, moving to a radius $\sim 3r_{scale}$ would roughly halve the average surface density determined from observing only the cluster core. This larger area corresponds to a survey field of ~ 3300 arcmin^2, roughly the size of the largest survey being compared here. A factor of two in surface density corresponds to ± 0.3 in logarithmic surface density; therefore, by normalizing surveys ranging from 1.25% of the cluster core area (this work) to $\sim 300\%$ the cluster core area (Mobasher et al., 2003) in the same way, we introduce a scatter of ± 0.3 to our logarithmic luminosity function. This level of error is acceptable for the comparison done here, as will be seen in Figure 6.1.

Finally, the magnitudes were converted to absolute magnitudes using Trentham's (1998) adopted distance modulus for Coma of $(m - M) = 34.83$.

Figure 6.1 is the composite R band luminosity function for the Coma cluster. The bright end shows good agreement between the various surveys. The counts of Andreon & Cuillandre (2002) are somewhat lower than the trend; in their paper, they scale their counts to those of Trentham (1998) and find the same discrepancy. They attribute this to the fact that they removed giant galaxies (and their surrounding areas) to avoid globular cluster contamination, a step that was not included when quoting their field size.

At the faint end, the counts of this work are a clean continuation of the brighter counts from the literature. They also agree well with Bernstein et al.'s additional counts. These are the counts they felt were accurate but

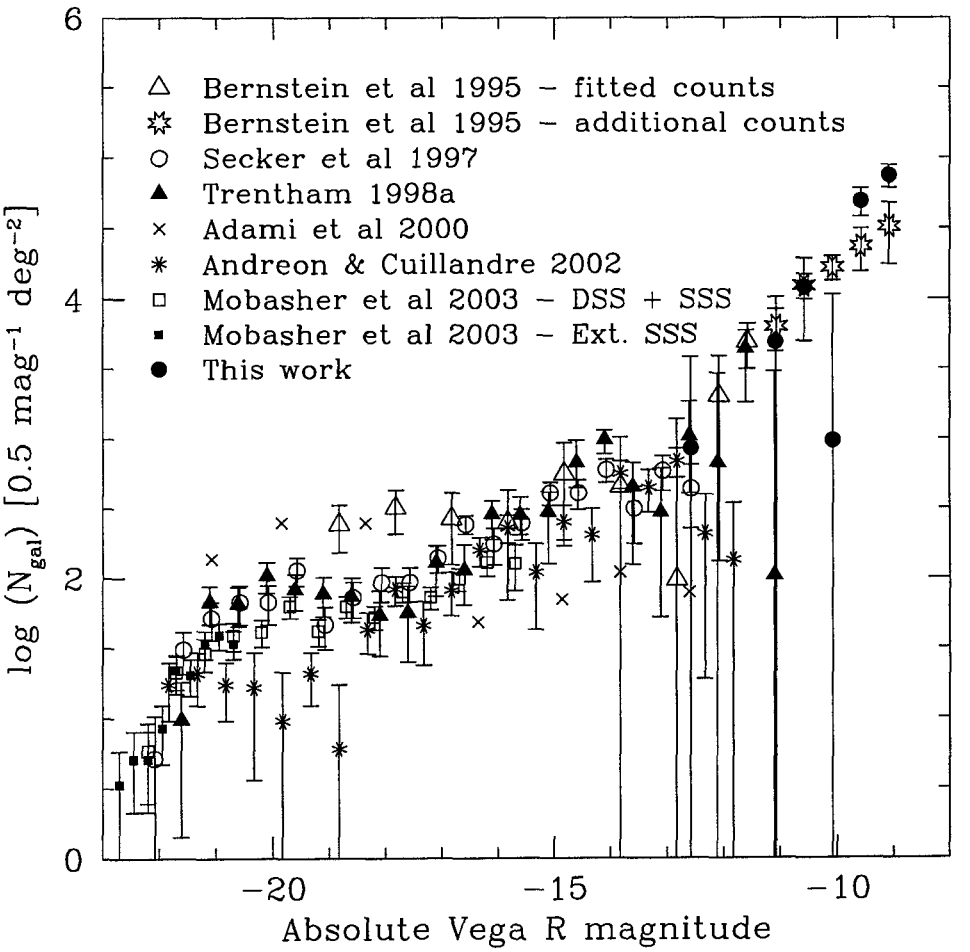

Figure 6.1: The composite Vega R band luminosity function of the Coma cluster, expressed as logarithmic counts per 0.5 magnitude bin and per deg^2. Various counts from the literature are shown: Bernstein et al. (1995, open triangles for points included in their fit and open stars for their additional points); Secker et al. (1997, open circles); Trentham (1998, filled triangles); Adami et al. (2000, "x"s); Andreon & Cuillandre (2002, stars); and Mobasher et al. (2003, open boxes for their DSS+SSS counts and filled boxes for their extended SSS counts). The counts from this work are shown as filled circles. All counts were scaled to an area of one deg^2; no other normalization was performed.

potentially contaminated by globular clusters. They had two arguments for this opinion: the first was based on examining the resolution parameter for objects in the fainter bins; the second was based on scaling the globular cluster system of M87 to NGC 4874.

To determine the shape of their faint objects, Bernstein et al. created histograms of their resolution parameter for the Coma field objects and the background field objects in each magnitude bin. They then subtracted these to obtain the distribution of their shape parameter for Coma objects in the fainter bins. They found that these distributions were consistent with a population of unresolved objects. However, as they go on to discuss, at their resolution fainter galaxies would in fact appear unresolved. They also point out that at low signal-to-noise it is difficult to determine the resolution of objects. Later in their paper, they suggest that HST observations would be ideal for distinguishing globular clusters from dwarf galaxies in the Coma cluster. In this work, we have used HST observations and taken special care to omit all unresolved objects from the catalogue.

The second argument of Bernstein et al. was based on comparing the predicted number of globular clusters to their faint bin counts. Their method of predicting the number of globular clusters was to assume that the globular cluster system of NGC 4874 was a scale model of the M87 globular cluster system. This involved a number of assumptions: the relative distances of the two clusters; the average V-R colour of globular clusters; the similarity of specific frequencies of globular clusters between the two clusters; and the surface density distribution of globular clusters around the two clusters. With this range of assumptions it is easy to imagine that the predicted number of

globular clusters could be in error. In summary, we are confident that our faint bins are not overly contaminated by globular clusters.

Especially interesting to note are the counts of Mobasher et al. and Adami et al. Of the R band studies discussed here, they were the only groups to use spectroscopic redshifts to determine cluster membership. Their counts agree reasonably well with those obtained by using photometric methods to determine cluster membership. This will be discussed further in Section 6.2.2.

Perhaps the most important thing that Figure 6.1 shows is that, although the various studies did not agree on a value for α, the counts themselves agree very well. Caution must always be used when comparing results via parameterized values rather than the underlying data.

6.2 Systematic Explanations For A Steep Faint End Slope

Based on the results of the previous section, it would appear that the faint end of Coma's luminosity function rises steeply. If a real effect, this has interesting implications for galaxy formation in the universe as a whole. There is one school of thought, however, that holds that the steep slope seen in some cluster luminosity functions is due to systematic errors, as opposed to any physical explanation. There are two main arguments used: first, that projection effects cause an artificial steepening of the luminosity function; and second, that statistical background subtraction is a faulty method of constructing luminosity functions. We discuss each of these ideas in turn. We also briefly mention the simulations of Driver et al. (1998), which they

used to asses the conditions under which statistical background subtraction can be used.

6.2.1 Projection Effects

A major study of the role of projection effects in cluster luminosity function determination was made by Valotto et al. (2001). They constructed a mock galaxy catalogue from a large N-body simulation and identified clusters in this catalogue by projecting it on the sky and searching for galaxy overdensities. The distances of the candidate clusters were determined from the velocities of the brightest galaxies in each overdensity, and the cluster richnesses were determined in a similar manner to the Abell and EDCC cluster identification procedures.

Valotto et al. next determined a luminosity function for each of their candidate clusters using statistical background subtraction. The control field counts were taken from an annulus surrounding each cluster candidate. Finally, they compared their computed luminosity functions to the luminosity function input to the simulation. While their input luminosity function had a flat slope with $\alpha = -0.97$, their composite cluster luminosity function determined from the catalogue had a steep slope with $\alpha = -1.41 \pm 0.11$.

Valotto et al. then went back to their catalogue and identified cluster candidates in three-dimensional space using a friends-of-friends algorithm. They took the richest of these, corresponding to Abell richness class 1 and higher, and computed luminosity functions using the same statistical background subtraction techniques as before. The composite luminosity function for these clusters had a flat slope with $\alpha = -1.08 \pm 0.09$, completely consis-

tent with the input luminosity function. From the results of these two tests, Valotto et al. concluded that projection effects severely hamper the determination of the luminosity function for clusters selected in two dimensions, but are not important for clusters selected in three dimensions.

The Coma cluster *is* a cluster selected in three dimensions. It is a true physical structure, identified by many methods in many wavelengths, and is the richest cluster known within 100 Mpc. We are confident that the projection effects found by Valotto et al. for clusters selected in two dimensions cannot apply here.

Another problem involving projection effects was examined by Beijersbergen et al. (2002b). As they discuss, luminosity functions in the literature are generally expressed in terms of the number of galaxies per unit magnitude per unit *area*. Thus, what is actually measured is a luminosity *distribution* (referred to hereafter as a 2-D luminosity function). This becomes important when comparing the 2-D luminosity function from the core region of a cluster to that from a region in the cluster outskirts: the core region will have foreground and background galaxies from the outer parts of the cluster projected onto it, while the cluster outskirts region will not. This could lead to an artificial steepening of the slope in the core region 2-D luminosity function.

To test this, Beijersbergen et al. went back to a previous study of Beijersbergen et al. (2002a). As indicated in Table 6.1, this project examined the 2-D luminosity function of Coma in three passbands over a very large area. They then used that data to investigate the dependence of the 2-D luminosity function on passband and distance from the cluster centre.

Beijersbergen et al. used the data from this earlier study to construct true 3-D luminosity functions, measured as counts per unit magnitude per unit *volume*. They deprojected their observations by assuming spherical symmetry and dividing the cluster area into a series of concentric shells, in each of which the galaxy density per unit magnitude was assumed to be constant. The 3-D luminosity function for each shell could then be found by subtracting the contributions due to projection from the other shells.

Armed with this method, Beijersbergen et al. repeated their study of the dependence of the luminosity function on passband and distance from cluster centre, this time using 3-D luminosity functions. They found that the 3-D luminosity function slopes did slightly flatten due to the removal of projection effects. However, the differences were small and not significant enough to change the conclusions inferred from 2-D slopes. They concluded that while correcting for projection effects is useful, the impact of projection effects is not as severe as Valotto et al. suggested.

In this work, we are studying the very core of the Coma cluster. Undoubtedly, galaxies in the foreground and background regions of the cluster are being projected onto the area we examined. However, as Beijersbergen et al. showed, the effects are small. In addition, we are not attempting any comparison with luminosity functions determined in different cluster regions. All of the studies used in the composite luminosity function (Figure 6.1) are also centred on the cluster core and so should experience the same small projection effects. We do not believe that projection effects are the cause of our steep faint end slope.

6.2.2 Statistical Background Subtraction Errors

The statistical background subtraction method for determining cluster membership is a powerful tool for cluster luminosity function studies. It allows the luminosity function to be determined in reasonable observing times over large areas and to faint magnitudes. Doing a similar study using spectroscopic redshifts to determine cluster membership would require a prohibitive amount of integration time on a large telescope, especially for the extremely faint objects we are observing here.

However, despite their obvious drawbacks, spectroscopic redshifts are indubitably more precise than any statistical method of determining cluster membership. Some authors have attempted spectroscopic studies of Coma's luminosity function. Three of these are listed in Table 6.1: Biviano et al.; Adami et al.; and Mobasher et al..

As can be seen, the values for α reported for these studies are among the lowest published. Some authors have interpreted this as an indication that statistical background subtraction should not be used; that the imprecision of the method causes artificially steep faint end slopes.

It should be noted, however, that spectroscopic studies generally have limiting magnitudes much brighter than photometric studies. Biviano et al.'s limiting magnitude of b = 20 can be roughly transformed to the R band by using $b - r = 1.8$, the average colour from their photometric catalogue (Godwin et al., 1983), and assuming $b - r \sim b - R$. Applying our adopted distance modulus of $(m - M) = 34.83$, their limiting magnitude becomes -16.6 in absolute R magnitudes – well below the turn-up in the slope seen in

Figure 6.1. Mobasher et al.'s counts also stop before the luminosity function moves into truly faint regimes. As can be seen in Figure 6.1, Mobasher et al.'s counts agree well with photometric counts at the bright end, and in no way preclude a steep faint end slope.

The counts of Adami et al. bear a little more discussion, as they do extend into the faint regions of the luminosity function. Adami et al. were working with very small numbers of redshifts: a sample of 88, compared to 265 for Biviano et al. and 1191 for Mobasher et al.. They declined to give error bars for their luminosity function, but assumed that they would be the same as those shown for the counts of Bernstein et al. (1995). With these error bars, their counts are consistent at a 1σ level with those obtained using photometric methods. The general trend of the points may hint at a possible systematic error arising from using photometric methods to construct a luminosity function, but it is dangerous to draw conclusions based on such small numbers. Until a larger sample of redshifts are obtained which show the same results, we remain confident that our steep faint end slope is not due to systematic errors.

6.2.3 Necessary Conditions For Statistical Background Subtraction

No discussion of the potential systematic errors involved in statistical background subtraction would be complete without mentioning the work of Driver et al. (1998). As part of a large project to measure the luminosity function of rich clusters, they performed extensive simulations to explore the validity of statistical background subtraction.

The simulations were based on detailed generated images of deep CCD exposures. Clusters were generated with various properties, their luminosity functions were determined using statistical background subtraction, and the results compared to the input data. The parameters they explored included cluster composition, true faint end slope, cluster richness, cluster redshift, seeing, and zero point errors.

The simulations, of course, depended on the specifics of the observing situation. Testing the observational configuration they would be using for their survey, they found the conditions under which statistical background subtraction could be safely used.

The observational conditions of Driver et al. are not similar to those experienced here in instrument, pixel scale, field of view or exposure time. Therefore, the specific conclusions they reached are unlikely to be applicable to this work. However, as they note, the simulation software they developed is freely available and can generate data for any detection set-up. Those engaged in future studies of luminosity functions may want to consider using this software or equivalent simulations to determine if statistical background subtraction is a good choice for their projects.

6.3 Implications Of A Steep Faint End Slope

Confident that systematic effects have not artificially steepened the slope of our luminosity function, we conclude that the faint end of Coma's luminosity function rises steeply. What are the implications of this result?

Perhaps most importantly, we have shown that there is at least one en-

vironment in the universe where a steep faint end slope is seen. As discussed in the Introduction, cold dark matter cosmologies predict that the faint end of the luminosity function should be steep. This is certainly not seen in all environments; the Local Group, for example, has a nearly flat faint end slope with $\alpha = -1.1 \pm 0.1$ (Pritchet and van den Bergh, 1999, see, however, Willman et al., 2004).

The question therefore now turns to what could cause this difference between faint end slopes in different environments. Tully et al. (2002) put forth one possible mechanism: the reionization of the universe. At reionization, the photoionizing background can heat gas to a temperature comparable to the gravitational energy of small galaxies ($v_{circ} \lesssim 30$ km/s), and so prevent the collapse of this gas into stars and galaxies. The idea is not new (see, for example, Efstathiou, 1992; Quinn et al., 1996; Thoul & Weinberg, 1996), but Tully et al. explored the effects of this mechanism in environments of different densities.

Using semi-analytic galaxy formation simulations which included the effects of reionization, Tully et al. generated a range of galaxy clusters. Tracing the merger trees of these simulations backward, they discovered that dwarf galaxies formed earlier in what would eventually become the richer clusters. This means that whenever reionization occurred, rich environments would have more dwarfs already in place. Put another way, they found that rich environments, which form their dwarfs earlier, would have the chance to form more dwarfs; while poor environments, which form their dwarfs later, would not be able to form as many.

As an observational test of this theory, Tully et al. proposed that clusters

and groups should contain a large population of low mass halos that were unable to form any stars at all. They went looking for these dark halos where their mass would be the most evident: in groups containing only dwarf galaxies. They identified four such groups within ~ 5 Mpc and found mass-to-light ratios for each one. Their results showed $M/L_B \sim 250-1000 M_\odot/L_\odot$, approximately 3 to 10 times higher than the ratio found in groups with larger galaxies. This result is consistent with the idea that dark halos make up a significant fraction of the mass in these groups. More recently, Simon et al. (2003) have proposed that a high velocity cloud seen near the Local Group dwarf galaxy LGS 3 could be a dark halo associated with the Local Group. This result, however, is still highly speculative.

The effects of photoionization in environments of different density was further explored by Benson et al. (2003b). They used more advanced semi-analytic models of galaxy formation to determine the properties of galaxies in a range of environments. These models included detailed treatments of the effects of supernova feedback, photoionization suppression, and dynamical friction and mass loss due to tidal forces. Qualitatively, their results agreed with Tully et al.: when photoionization suppression was switched on, faint end slopes in poor environments became flatter than those in rich environments.

As an interesting side note, Benson et al. also discovered that including the effects of photoionization but switching off tidal losses did not significantly affect the faint end of the luminosity function in any environment. This somewhat counterintuitive result supports the observation that rich clusters are seen to have more dwarfs, in spite of tidal effects which would

seem to be strongest in these environments.

Benson et al. conclude that photoionization suppression indeed affects the luminosity function to different degrees in different environments. However, comparing their results to the data of Trentham & Hodgkin (2002), they find that photoionization alone cannot flatten the slope enough to match the luminosity functions seen in the Local Group and Ursa Major. Some other feedback mechanisms must also be at work.

A detailed study of the mechanisms that shape the luminosity function was performed by Benson et al. (2003a). They ran a series of semi-analytic simulations of galaxy formation, including one additional baryonic process in each model, with the aim of matching both the flat faint end slope and the abrupt bright end drop-off seen in many cluster luminosity functions. Along with the standard processes of supernova feedback, photoionization suppression and galaxy mergers, they also studied three additional mechanisms: supernova and stellar wind heating of the gas, causing it to expand and so lengthen its cooling time; thermal conduction as a method of heating the central regions of the cluster; and so-called "superwinds" that expel gas entirely from the potential.

Benson et al.'s results showed that a good match to both the faint and the bright end of the luminosity function could be obtained with models including the effects of thermal conduction. The efficiency of the conduction required, however, was extremely high, and the temperature gradient in the cluster would have to be somewhat steeper than the gradients generally observed. Adding the effects of superwinds could relax the requirements on the thermal conduction somewhat while still reproducing the shape of the luminosity

function. However, the conduction efficiency would still have to be very high, and the total energy required by the superwinds would significantly exceed that available from supernovae alone.

Obviously, further theoretical efforts are needed to refine the exact processes that shape the luminosity function. Based on the results of this work, these studies must now also explain how these processes can lead to a luminosity function that is different in environments of different densities.

Table 6.1: Previous studies of Coma's luminosity function

Study	Passband	Field size arcmin^2	Pixel scale arcsec/pix	Lim. mag. Vega mags	α^a
Thompson & Gregory 1993	b	14 292	18.56[b]	20	-1.43
Biviano et al. 1995	b	2 496	67.2[b]	20	-1.2 ± 0.2
Bernstein et al. 1995	R	52.2	0.473	23.5[c]	-1.42 ± 0.05
Secker et al. 1997	R	~ 700	0.53	22.5	-1.41 ± 0.05
Lobo et al. 1997	V	1 500	0.3145	21	-1.8 ± 0.05
Trentham 1998a	R	674	0.22	23.83	-1.7
Adami et al. 2000	R	52.2	0.473	22.5	-1
Andreon & Cuillandre 2002	B	720	0.206	22.5	-1.25
	V	1 044	0.206	23.75	-1.4
	R	1 044	0.206	23.25	-1.4
Beijersbergen et al. 2002a	U	4 680	0.333	21.73	$-1.32^{+0.018}_{-0.028}$
	B	18 720	0.333	21.73	$-1.37^{+0.024}_{-0.016}$
	r	18 720	0.333	21.73	$-1.16^{+0.012}_{-0.019}$
Mobasher et al. 2003	R	3 600	0.21	19.5	$-1.18^{+0.04}_{-0.02}$
Iglesais-Páramo et al. 2003	r'	3 600	0.333	20.5	$-1.47^{+0.08}_{-0.09}$
This work	R	4.62	0.1	25.75	-2.28 ± 0.34

[a] When no error is given for α, no value was supplied in the text

[b] These studies were done using photographic plates. The value listed for "Pixel scale" is actually that for plate scale, measured in arcsec/mm

[c] Bernstein et al. reported two limiting magnitudes for their work. The first, shown here, is the faintest magnitude bin included in the fit to their luminosity function. Beyond this, they believed their counts were accurate to R ~ 25.5, but decided not to use the data for fear of globular cluster contamination.

Chapter 7

Conclusions

7.1 Summary

This thesis has measured the faint end of the luminosity function in the core of the Coma cluster. The method of differential counts was used to construct the luminosity function: "blank sky" control field counts were subtracted from the data field counts in order to statistically remove background contamination. HST images were obtained for both the data field and the control field images. For the data field, F606W and F814W band images of NGC 4874 were used. For the control field, the HDF was used. The HDF was degraded to match the detection characteristics of the Coma image.

The SExtractor package was used to create catalogues of objects in the Coma and HDF images. Extensive "add-galaxy" tests were used to determine the best detection and photometry procedures and parameters. A globular cluster mask was created and used to prevent globular clusters from contaminating the catalogues. The size of the faint galaxies in the Coma catalogue were found to be consistent with the sizes of dwarf galaxies studied in other

works.

The final catalogues were found to be complete to $R = 25.75$, allowing the faintest determination of the Coma luminosity function to date. The HDF number counts were scaled to the area of the Coma field and subtracted from the Coma number counts. The slope of the resulting luminosity function was found to be fit by a value of $\alpha = -2.28 \pm 0.34$.

The counts determined in this work were combined with those available in the literature to produce a composite luminosity function. This plot showed good agreement between the various studies, and showed the counts of this work to be a clean continuation of the existing steep faint end slope to fainter magnitudes. Common arguments attributing steep faint end slopes to systematic errors were shown to not apply in this case. The value of α found in this work is consistent with theories that predict photoionization and other feedback effects will affect environments of low density more severely than environments of high density.

7.2 Future Work

All the data used in this work was obtained from the HST archives. Data mining in the archives of HST and other telescopes presents great possibilities for many lines of astronomical research, including the determination of luminosity functions.

For example, the determination of the control field number counts could be greatly improved by examining the background over a larger area of the sky. The HST has been used for a number of large scale surveys of the

distant universe – that is, of the background. The Medium Deep Survey, for example, was completed using the WFPC2 camera. It covers 1 deg^2, 500 times the area of the HDF. More recently, the Great Observatories Origins Deep Survey used the Advanced Camera for Surveys on HST to image 320 arcmin^2 of the sky in two fields; this represents an area 44 times the area of the HDF. Determining background number counts from either of these larger fields would greatly reduce the errors due to cosmic variance and improve the determination of the luminosity function.

More data fields of the Coma cluster are also available from the HST archive. Reasonably deep ($\geq$ 4000 s) images are available for at least four other regions of the Coma cluster: the immediate surroundings of NGC 4889, NGC 4926, NGC 4881 and IC 4051. These four pointings, along with the NGC 4874 pointing used in this work, are shown in Figure 7.1. As can be seen, the images cover a wide range of locations in the Coma cluster. Comparing the luminosity functions obtained from these different fields could provide an interesting study of the distribution of dwarf galaxies as a function of cluster radius.

Such a study would still be plagued with problems related to the small field size at each pointing. However, the number counts from each pointing could be combined appropriately to produce a larger data set and so a more statistically significant luminosity function. With additional data, more in-depth studies of the properties of Coma dwarf galaxies could be performed: luminosity functions in different bands could be compared (Beijersbergen et al., 2002a); a bivariate luminosity function could be constructed (Cross and Driver, 2002; Andreon and Cuillandre, 2002); and the relative clustering

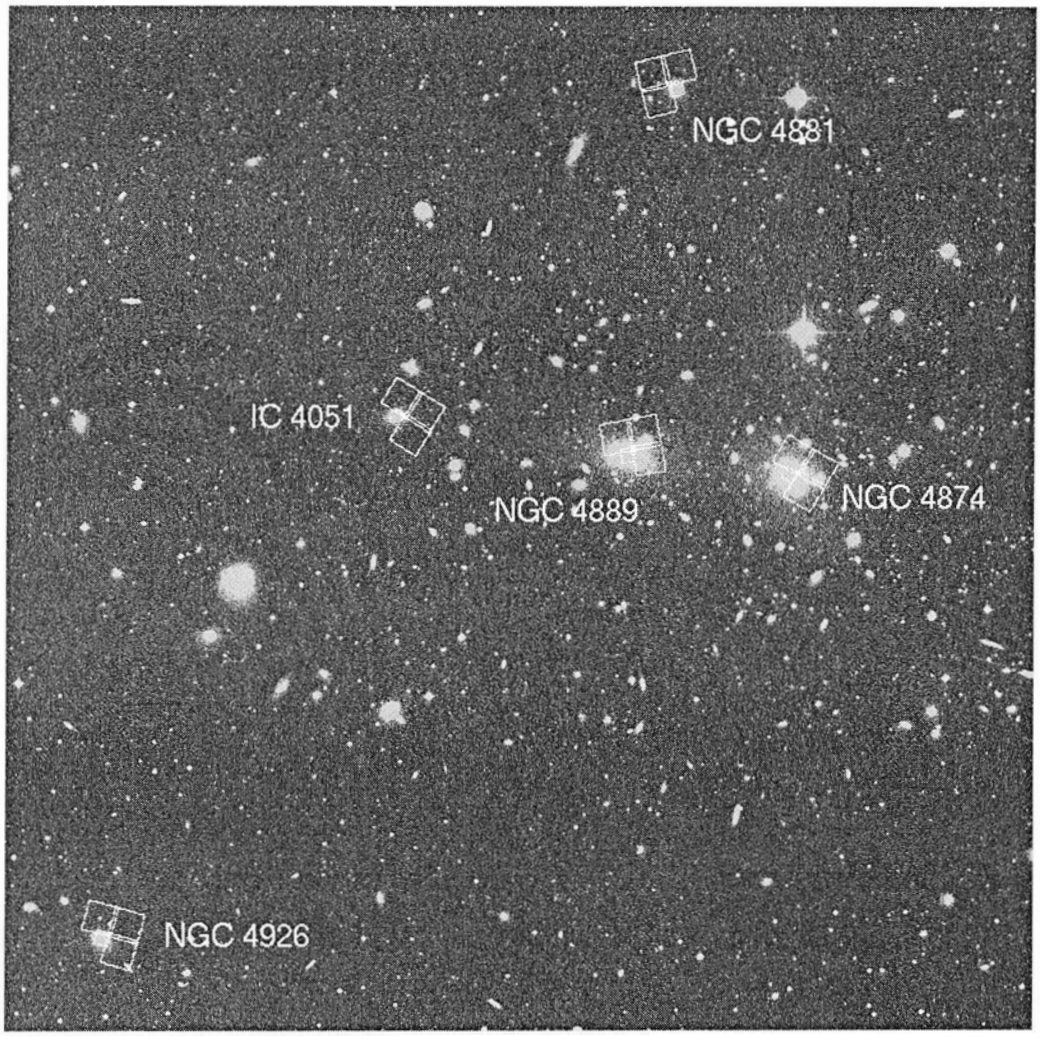

Figure 7.1: An image of the Coma cluster, showing five WFPC2 pointings for which observations are available in the HST archive. The five pointings are in the regions of NGC 4889, NGC 4926, NGC 4881, IC 4051 and NGC 4874. The background image is from the archive of the Digitized Sky Survey, available from the Canadian Astronomical Data Centre.

of dwarf and giant galaxies could be examined (Edwards et al., 2002).

Moving away from Coma, the HST archive could also be mined for deep images of other clusters. Constructing luminosity functions for several clusters, using the same source of data and processing steps for each, could provide a valuable and unique insight into the dependence of the luminosity function on environment.

Bibliography

Adami, C., Ulmer, M. P., Durret, F., Nichol, R. C., Mazure, A., Holden, B. P., Romer, A. K., and Savine, C.: 2000, *Astron. & Astrophys.* **353**, 930

Andreon, S. and Cuillandre, J.-C.: 2002, *Astrophys. J.* **569**, 144

Barazza, F. D., Binggeli, B., and Prugniel, P.: 2001, *Astron. & Astrophys.* **373**, 12

Beijersbergen, M., Hoekstra, H., van Dokkum, P. G., and van der Hulst, T.: 2002a, *M.N.R.A.S.* **329**, 385

Beijersbergen, M., Schaap, W. E., and van der Hulst, J. M.: 2002b, *Astron. & Astrophys.* **390**, 817

Bender, R., Burstein, D., and Faber, S. M.: 1992, *Astrophys. J.* **399**, 462

Benson, A. J., Bower, R. G., Frenk, C. S., Lacey, C. G., Baugh, C. M., and Cole, S.: 2003a, *Astrophys. J.* **599**, 38

Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., and Lacey, C. G.: 2003b, *M.N.R.A.S.* **343**, 679

Bernstein, G. M., Nichol, R. C., Tyson, J. A., Ulmer, M. P., and Wittman, D.: 1995, *Astron. J.* **110**, 1507

Bertin, E. and Arnouts, S.: 1996, *Astron. & Astrophys. Supp.* **117**, 393

Bevington, P. R.: 1969, *Data reduction and error analysis for the physical sciences*, New York: McGraw-Hill, 1969

Binggeli, B. and Cameron, L. M.: 1993, *Astron. & Astrophys. Supp.* **98**, 297

Bremnes, T., Binggeli, B., and Prugniel, P.: 1998, *Astron. & Astrophys. Supp.* **129**, 313

Bremnes, T., Binggeli, B., and Prugniel, P.: 1999, *Astron. & Astrophys. Supp.* **137**, 337

Cross, N. and Driver, S. P.: 2002, *M.N.R.A.S.* **329**, 579

Driver, S. P., Couch, W. J., Phillipps, S., and Smith, R.: 1998, *M.N.R.A.S.* **301**, 357

Edwards, S. A., Colless, M., Bridges, T. J., Carter, D., Mobasher, B., and Poggianti, B. M.: 2002, *Astrophys. J.* **567**, 178

Efstathiou, G.: 1992, *M.N.R.A.S.* **256**, 43P

Fukugita, M., Shimasaku, K., and Ichikawa, T.: 1995, *Pub. Astron. Soc. Pac.* **107**, 945

Godwin, J. G., Metcalfe, N., and Peach, J. V.: 1983, *M.N.R.A.S.* **202**, 113

Groth, E. J.: 1986, *Astron. J.* **91**, 1244

Kavelaars, J. J., Harris, W. E., Hanes, D. A., Hesser, J. E., and Pritchet, C. J.: 2000, *Astrophys. J.* **533**, 125

Kron, R. G.: 1980, *Astrophys. J. Supp.* **43**, 305

Łokas, E. L. and Mamon, G. A.: 2003, *M.N.R.A.S.* **343**, 401

Mobasher, B., Colless, M., Carter, D., Poggianti, B. M., Bridges, T. J., Kranz, K., Komiyama, Y., Kashikawa, N., Yagi, M., and Okamura, S.: 2003, *Astrophys. J.* **587**, 605

Parodi, B. R., Barazza, F. D., and Binggeli, B.: 2002, *Astron. & Astrophys.*

388, 29

Peebles, P. J. E.: 1980, *The large-scale structure of the universe*, Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.

Petrosian, V.: 1976, *Astrophys. J. Let.* **209**, L1

Pritchet, C. J. and van den Bergh, S.: 1999, *Astron. J.* **118**, 883

Quinn, T., Katz, N., and Efstathiou, G.: 1996, *M.N.R.A.S.* **278**, L49

Schlegel, D. J., Finkbeiner, D. P., and Davis, M.: 1998, *Astrophys. J.* **500**, 525

Secker, J., Harris, W. E., and Plummer, J. D.: 1997, *Pub. Astron. Soc. Pac.* **109**, 1377

Simon, J. D., Robishaw, T., and Blitz, L.: 2003, *ArXiv Astrophysics e-prints*

Thoul, A. A. and Weinberg, D. H.: 1996, *Astrophys. J.* **465**, 608

Trentham, N.: 1998, *M.N.R.A.S.* **293**, 71

Trentham, N. and Hodgkin, S.: 2002, *M.N.R.A.S.* **333**, 423

Tully, R. B., Somerville, R. S., Trentham, N., and Verheijen, M. A. W.: 2002, *Astrophys. J.* **569**, 573

Valotto, C. A., Moore, B., and Lambas, D. G.: 2001, *Astrophys. J.* **546**, 157

Willman, B., Governato, F., Dalcanton, J. J., Reed, D., and Quinn, T.: 2004, *ArXiv Astrophysics e-prints*

Wilson, G.: 2003, *Astrophys. J.* **585**, 191

Zwicky, F.: 1957, *Morphological astronomy*, Berlin: Springer, 1957