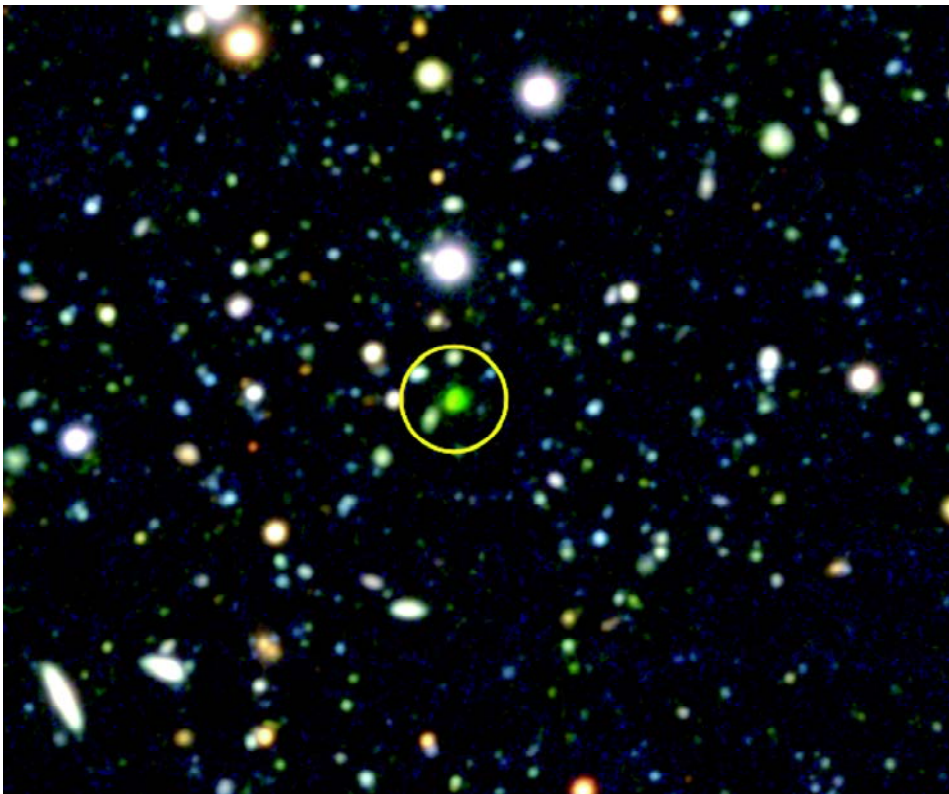# Large Research Project:

Large Scale Structure at z = 4:
Lyman Break Galaxies in a wide field
around Radio Galaxy TN J1338–1942

H.T. Intema, Sterrewacht Leiden

7th December 2004

ii

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This report is part of the final year research project that was performed by the author as a student member of the High Redshift Universe Research Group of Prof. Dr. G.K. Miley at the Sterrewacht Leiden. This is the second of two research projects that are part of the Astrophysics master curriculum at Leiden University. This project was supervised by Dr. H.J.A. Röttgering and Drs. B.P. Venemans.

The starting point for this project was a set of raw (unprocessed) images of a large 25×25 arcmin² FOV (field-of-view) around the high redshift radio galaxy (HzRG) TN J1338–1942 (see cover image) at redshift $z = 4.1$. These images were taken by J.D. Kurk, in collaboration with M. Ouchi and T. Kodama, with the Suprime-Cam on the Subaru telescope on Mauna Kea, Hawaii. These images are part of a larger observing program to search the environment of HzRGs for evidence of galaxy clustering. Luminous HzRGs are found to be the most massive galaxies in the early universe. It is likely that these HzRGs reside in the centres of galaxy clusters or proto-clusters (clusters undergoing formation).

The set of raw images contained 3 wideband filter exposures in B, R_C and i'. The colors $B - R_C$ and $R_C - i'$ of objects in these images were to be used to select candidate Lyman break galaxies (LBGs) from an approximate redshift range $z = 3.5 - 4.5$. The resulting catalog of LBG candidates would be complementary to the catalog of LBGs in a small FOV (3.6×3.6 arcmin²) around TN J1338–1942 by Miley et al. (2004) and Overzier et al. (in preparation) and to the catalog of Lyman-α emitters from 2 intermediate FOVs (both 7×7 arcmin², partially overlapping) around TN J1338–1942.

The aim of this research project was to reduce, calibrate and analyse the raw images, to extract a catalog of LBGs from these images and to analyse the spatial distribution of these LBGs, especially in close proximity to TN J1338–1942. In this report magnitudes are expressed in the AB-magnitude system, uncertainties are gaussian 1σ uncertainties and cosmological distances are comoving distances, unless stated otherwise. The adopted cosmological parameter values are $(\Omega_m, \Omega_\Lambda, h) = (0.3, 0.7, 0.7)$.

1.1 Cosmology

The starting point for 'modern' cosmology was Einstein's general relativity theory, which states that the gravitational force is a geometrical effect of curved space-time. It is the energy content (like mass, kinetic energy, pressure, radiation, etc.) that determines the space-time curvature. Einstein, de Sitter, Lemaître, Friedmann, Robertson and Walker were people who used Einstein's general relativity theory to construct mathematical models of the universe, using only a few basic assumptions:

- On large scales, i.e. larger than a few hundred Megaparsec, the universe is homogeneous: it has the same energy content at all places.

- On large scales the universe is isotropic: it looks the same in all directions.

- On large scales the content of the universe acts like a fluid.

All these cosmological models are characterised by a set of parameters, namely the scale parameter R, the Hubble parameter H, the curvature (constant) k and the density parameter Ω, which are all functions of 'cosmological' time t. They are related by the Friedmann equation:

$$H^2 \left(1 - \Omega\right) = -\frac{kc^2}{R^2} \tag{1.1}$$

The constant c is the speed of light. A subscript '0' denotes the current value (at $t = t_0 = $ now) of a cosmological parameter.

Edwin Hubble was the first to observe that other galaxies are moving away from ours. From the assumption of homogeneousity, this implied that the whole universe was expanding. Theory predicted this, as it was hard to construct a static universe model that was stable against gravitational collapse. Galaxies can be pictured as particles that are drifting on the expanding Hubble flow. When peculiar velocities are neglected, the proper coordinates of galaxies in expanding space can be described by constant coordinates within a comoving reference frame, multiplied by the scale parameter R. The Hubble parameter H is equal to the expansion rate $\dot{R}/R$, where $\dot{R}$ is the time derivative of R. In an expanding universe, H is positive, while R is positive in any real universe.

The density parameter Ω is the ratio of the mass density ρ (or rather energy density ρc^2) over the critical density $\rho_c = 3H^2/8\pi G$, where G is the gravitational constant. The most common contributions to ρ (or Ω) are matter $\rho_m \sim R^{-3}$, radiation $\rho_r \sim R^{-4}$ and vacuum energy $\rho_\Lambda \sim R^0$. There are, in general, three types of solutions for Equation (1.1), namely a closed universe with $k > 0$ ($\Omega > 1$), an open universe with $k < 0$ ($\Omega < 1$) and a flat universe with $k = 0$ ($\Omega = 1$).

From the 1960's on, the precision and sensitivity of astronomical instrumentation have advanced with such pace that observational cosmology has changed our picture of the universe. Two types of observations had a major impact on cosmology. The first is galaxy surveys out to considerable redshift, e.g. 3C/4C/5C (Edge et al. 1959; Pilkington&Scott 1965; Pooley&Kenderdine 1968), CfA (Huchra et al. 1983), IRAS (Lonsdale&Hacking 1989), LCRS (Shectman et al. 1996), ISO-IRAS (Levine et al. 1998), SDSS (Strauss et al. 2002) and 2dFGRS (Colless et al. 2003). The second is measurements on the cosmic microwave background (CMB), e.g COBE (Smoot et al. 1991) and WMAP (Spergel et al. 2003). Measurements of the CMB temperature have shown that the universe at redshift $z \approx 1000$ (with $1 + z = R_0/R$) looks highly isotropic, except for small relative temperature fluctuations of the order $\sim 10^{-5}$ on a ~ 1 degree scale. The galaxy surveys observe the universe out to redshifts of $z \lesssim 1$, which looks homogeneous and isotropic on average on a scale of a few hundred Mpc, but shows strong signs of structure on smaller scales. The time span between $z \approx 1000$ and $z \approx 1$ is a very interesting era, as the small density fluctuations must have evolved into the structure we observe in the local universe.

The current values of the cosmological parameters have been estimated from CMB observations and sky surveys, although some of these estimates varied significantly amongst authors. Recently, it seems that there is consensus on the following values (Spergel et al. 2003): $\Omega_{m,0} \approx 0.3$, $\Omega_{r,0} \approx 0$, $\Omega_{\Lambda,0} \approx 0.7$ and $h \approx 0.7$, where $H_0 = 100\, h$ km s^{-1} Mpc^{-1}, which suggests a Λ-dominated, flat universe. These values were adopted for this report.

1.2   Galaxy Clusters

Although the universe may look homogeneous on the scale of a few hundred Mpc, there is overwhelming evidence that it is not the case on smaller scales. The common paradigm on structure

formation (e.g. Peacock 2001) is that primordial density fluctuations within a certain mass range became gravitationally unstable and collapsed into dense structure surrounded by voids of empty space. The largest inhomogeneous structures observed are galaxies and galaxy clusters. Typical sizes for rich clusters of galaxies in the local universe ($z < 1$) are $10 - 40$ h^{-1} Mpc (e.g. Bahcall 1988, Peacock 2001). The classical picture of galaxy clusters as spherical systems is somewhat oversimplified. Both observations and models show that galaxies often tend to cluster together into elongated structures (e.g. Kauffman 1999; Steidel et al. 2000; Shimasaku et al. 2003; Palunas et al. 2004), like filaments or sheets, which surround large voids of (relatively) empty space. Elongated galaxy distributions are observed in the local universe have lengths up to 100 h^{-1} Mpc, while their widths are only ~ 10 h^{-1} Mpc.

At larger redshift, i.e. at earlier times, there is also evidence of clustering of galaxies. For example, Palunas et al. (2004) found a filament of Lyman-α emitters (LAEs) at $z = 2.36$ of ~ 80 Mpc in length, while Möller&Fynbo (2001) found a filament of LAEs at $z = 3.01$ of ~ 20 Mpc in length. Ouchi et al. (2004b) found clustering signals in their LBG samples at both $z \sim 4$ and $z \sim 5$, which indicates a non-random distribution of LBGs, but there is no evidence presented for the presence of a proto-cluster in their observed fields. Venemans et al. (2002) were the first to discover a proto-cluster at a redshift $z = 4.1$ around the high redshift radio galaxy TN J1338–1942. Shimasaku et al. (2003) claim to have found a proto-cluster at a redshift $z = 4.9$, but this result is questionable due to the low number of proto-cluster galaxies (43) within the large FOV (25×45 arcmin^2). And within the same FOV, Shimasaku et al. (2004) found no trace of clustering at $z = 4.8$.

1.3 High Redshift Radio Galaxy TN J1338–1942

TN J1338–1942 is a radio source in the southern hemisphere of the sky in the constellation Virgo. It was identified as being a high redshift galaxy candidate by its steep radio spectrum (Röttgering et al. 1994; De Breuck et al. 2000). It has an extended radio structure for which an optical counterpart was found in the POSS catalog (Reid et al. 1991). Follow-up spectroscopy (De Breuck et al. 1999) showed that TN J1338–1942 has a very strong Lyman-α emission line, redshifted to $z = 4.1$, therefore making it one of the most luminous LAEs in its class.

Radio-loud galaxies at high redshift are believed to represent the most massive galaxies in the early universe (De Breuck et al. 2002). Observations on nearby galaxy clusters (Matthews et al. 1964) showed that massive cD-type galaxies are usually located in the cores of local clusters. As HzRGs are the best candidate predecessors of cD galaxies, it is assumed that HzRGs will predominantly occupy the cores of (proto-)clusters. This makes steep-spectra radio-loud galaxies a good tracer for finding proto-clusters.

As TN J1338–1942 qualifies for being a very massive galaxy at high redshift, it could very well be centered in a proto-cluster environment. Therefore, the environment of TN J1338–1942 was searched for other galaxies at similar redshift. Venemans et al. (2002 + extension) found 86 LAEs (38 spectroscopically confirmed, including TN J1338–1942) in a narrow redshift range around $z = 4.1$, in an area of ~ 80 arcmin^2 around TN J1338–1942. Within the narrow redshift range, the redshift distribution of LAEs peaks just below the redshift of TN J1338–1942. Miley et al. (2004) and Overzier et al. (in preparation) found 71 LBGs in a wider redshift range around $z = 4$ in an area of ~ 12 arcmin^2 around TN J1338–1942. Both the LAE and LBG observations indicate an excess of galaxies is present in the near proximity (i.e. at small angular distance and at similar redshift) of TN J1338–1942.

The overdense distribution of LAEs around TN J1338–1942 seems to extend beyond the FOV of Venemans et al. Except for the area south-east of TN J1338–1942, the typical observed angular distance from TN J1338–1942 is 3 arcmin, which corresponds to 6 Mpc at $z = 4$. As typical cluster sizes in the local universe are $\gtrsim 10$ Mpc (see Section 1.2), a larger FOV is needed to map the detected overdensity out to its boundaries. The 25×25 arcmin^2 FOV used in this report gives

typical angular distances of 10 arcmin, which corresponds to 21 Mpc at $z = 4$, therefore having a better chance of capturing the whole overdense distribution around TN J1338–1942 (plus a better determination of the mean galaxy density).

1.4 Lyman Break Galaxies

To map the structure of galaxy clusters, it is necessary to measure the angular position and the distance (redshift) of large numbers of galaxies. This normally involves selecting galaxy candidates from images by a selection method (e.g. morphology, color, magnitude) and determining their redshift by follow-up spectroscopy. When mapping galaxy distributions out to higher redshifts, selection becomes more difficult as objects become fainter and confusing galaxies become more numerous. At higher redshifts, galaxies often require special features, e.g. strong emission line in LAEs or beaming in Quasars, to do spectroscopy.

While a less accurate method in terms of redshift values, a more sensitive and efficient way of finding galaxies at higher redshift is by the use of colors, i.e. differences in wideband magnitudes. Photometry using wideband filters is, in general, much faster than spectroscopy because of the higher photon yield per band. The determination of photometric redshift (Puschell et al. 1982) makes use of near-infrared wideband photometry and places several requirements on the galaxy spectrum. For example, Labbé et al. (2003) used this method to find DGRs (distant red galaxies), star-forming galaxies with heavy dust extinction at (rest-frame) UV and optical wavelenghts, out to redshifts of $z \approx 3.5$.

The Lyman break technique (Steidel et al. 1996) measures colors of relatively dust-free star-forming galaxies that emit strongly in rest-frame UV (e.g. Giavalisco 2002). At high redshift ($z \gtrsim 3$), the UV-flux gets redshifted into the optical and near-infrared to within the range of wideband filters on ground-based telescopes. From quasar absorption spectra, it became clear that for high redshift galaxies a significant part of the UV flux with wavelengths beyond the Lyman break at rest-frame wavelength $\lambda \leq 912$ Å is absorbed (see also Madau 1995; Songaila 2004).

The Lyman break is caused by ionisation of neutral hydrogen in the intergalactic medium by UV photons. With clouds of HI along the line-of-sight (LOS) towards a galaxy, a fraction of the flux with $\lambda \leq 912$ Å gets absorbed, depending on the column density of HI. In the optically thick case, practically all of the flux with $\lambda \leq 912$ Å gets absorbed. This causes a sudden drop (Lyman break) in the spectrum of the galaxy at the short wavelength side of 912 Å, assuming that the galaxy emits enough UV photons for this effect to be observed, i.e. dust-free star-forming galaxies.

Steidel et al. (1996) found that there is significant absorption beyond the redshifted Lyman break for UV emitting galaxies at redshift $z \gtrsim 3$. These so-called Lyman break galaxies (LBGs) can be found by comparison between wideband filter images on both the short wavelength side and the long wavelength side of the redshifted Lyman break, as they are heavily dimmed (if not invisible) on the short wavelength side. The normal setup is one wideband filter on the short wavelength side (e.g. U) and two filters on the long wavelength side (e.g. G and R), so that colors can be defined from magnitude differences between adjacent filters (e.g. $U - G$ and $G - R$). The color of the short minus the intermediate wavelength filter magnitude is typically $\gtrsim 1.5$, while the color of the intermediate minus the long wavelength filter magnitude is typically $\lesssim 0$. The latter requires a detection in both the intermediate and the long wavelength filter images and selects galaxies whose continuum spectrum stays flat or increases towards the UV.

Selecting LBG candidates by their colors will select other objects with similar colors (e.g. late-type galactic stars and early-type low redshift galaxies). It is therefore crucial to define the selection criteria in such a way that the completeness of the candidate LBG sample is as high as possible, while the contamination is as low as possible. A trade-off has to be made, which should include the effects of (larger) photometric uncertainties of objects near the detection limit. It

should be noted that LBGs are only a fraction of the true population of high redshift galaxies, as not all galaxies are bright in the UV. For example, galaxies with little star formation and star-forming galaxies with heavy dust extinction (e.g. SCUBA galaxies or DRGs; Barger et al. 1999; Labbé et al. 2003) will escape the detection limit or selection criteria.

For LBGs at $z = 3.5 - 4.5$, there is also significant absorption by HI from the Lyman-α line at rest-frame $\lambda = 1216$ Å (Ouchi et al. 2004a). As UV photons from high redshift galaxies are redshifted on their way to the observer, the Lyα absorption line sweeps across the spectrum. For this report we used selection criteria for LBGs within the redshift range $z \approx 3.5 - 4.5$, based on the colors $B - R_C$ and $R_C - i'$, as defined by Ouchi et al. (2004a), which incorporates both the Lyman break as well as Lyα absorption.

1.5 Overview

In Chapter 2 we describe how the raw (unprocessed) images of the TN J1338–1942 field were obtained and which instruments were used. Chapter 3 contains a detailed description of the data reduction of the raw images and generation of the object catalog. Chapter 4 explains how LGB candidates were selected from the object catalog and what techniques were used to find evidence of structure in the spatial distribution of the candidate LBGs. In Chapter 5, results from Chapters 3 and 4 are evaluated and summarised.

Chapter 2

Observations

2.1 Instrumentation

2.1.1 Subaru Telescope

The Subaru telescope is an optical-infrared telescope located on the 4200 meter top of Mauna Kea (see Figure 2.1) on the Big Island of Hawaii. The telescope is named after the open star cluster 'Pleiades', which is called 'Subaru' in Japanese. The telescope was built and is operated by the National Astronomical Observatory of Japan (NAOJ). It is housed in a cylindrically shaped enclosure to suppress local atmospheric turbulence by preventing the entrance of rising warm, turbulent air from the outside while allowing the warm air produced inside the enclosure to escape rapidly.

The heart of the Subaru telescope is a 20 centimeter thick monolithic primary mirror with a diameter of 8.3 meters (8.2 meters effective aperture) weighing 22.8 tons and supported by an altitude-azimuth mounting (see Figure 2.2). The mean surface deviation of the primary mirror is 14 nanometers. Its focal length is 15 meters, giving it a fast focal ratio of f/2.0 at its primary focus. Other features of the telescope are a highly accurate tracking mechanism using magnetic driving systems, up to seven observational instruments installed at four different foci, namely primary focus, Cassegrain focus, Nasmyth optical focus and Nasmyth infrared focus (see Figure 2.3), and an auto-exchanger system to switch among the observational instruments. The Subaru telescope is equiped with an adaptive optics system of 261 actuators on the back of the primary mirror, that maintains a very high mirror surface accuracy. A PSF FWHM of less than 0.1 `arcsec` is possible, depending on the guide star magnitude and distance from target, but only when using an instrument at the Cassegrain focus.

2.1.2 Suprime-Cam

The Subaru Prime Focus Camera (Suprime-Cam; Miyazaki 2002) is a 5 × 2 mosaic of 10 CCDs (see Figure 2.4), located at the primary focus of the Subaru telescope, although it can operate at the other foci. Because of image curvature due to the fast focal ratio (f/2.0) at primary focus, a 7-element 20-inch diameter corrector lens was placed in front of the Suprime-Cam to ensure sharp images over the Suprime-Cam's large FOV.

Each of the 10 CCDs on the Suprime-Cam is a 2-dimensional pixel-array of 2048 columns × 4096 rows, with each pixel having a scale of 0.20 × 0.20 `arcsec`^{2}. Small gaps exist between adjacent CCDs due to mechanical constraints. The combined area of the 10 CCDs gives the Suprime-Cam a FOV of 34 × 27 `arcmin`^{2} on the sky.

The CCDs are sensitive to optical and near-infrared radiation, covering a wavelength interval of approximately 3500 − 10500 Å. Within this interval, each incident photon that hits the pixel

Figure 2.1: The Subaru telescope facility on top of Mauna Kea, Hawaii.

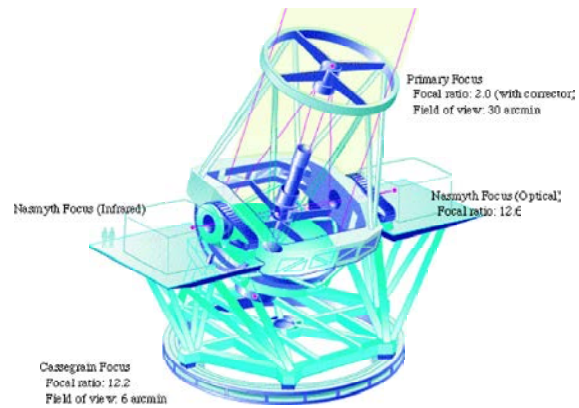Figure 2.2: Inside the Subaru telescope facility.

Figure 2.3: Schematic overview of the Subaru telescope.

surface has a certain chance (quantum efficiency (QE); see Figure 2.5) of being detected, by freeing an electron within the pixel. The number of freed electrons per pixel can be read out of the CCD, which takes 60 seconds for all CCDs in parallel. The total number of freed electrons per pixel is limited to $\sim 80,000$ e⁻, above which the pixel saturates and free electrons leak into adjacent pixels, mainly along the column direction, causing bright lines in the resulting image called 'bleeders' (see Figure 3.1).

The readout mechanism of the CCD converts electron counts into Analog-to-Digital Units (ADUs), with a gain factor of 2.6 e⁻ per ADU. The readout mechanism adds noise to the signal with an RMS value of 10 e⁻ per readout pixel. At the start of each integration, all pixels are preset to an offset value (bias) of $10,000$ ADUs to ensure that the CCDs operate in the linear-response region of their dynamical range. This means that the saturation level is reached at $40,000$ ADUs. Each CCD has an additional overscan region of 32 columns, positioned to the left of the 2048 columns. The signal from this region is used for bias subtraction during data reduction (see Section 3.1.1).

Due to the multiple CCD configuration of the Suprime-Cam, each readout generates 10 output files, one for each CCD. In this report, the term 'image' refers to the output of 10 CCDs during one readout, while the term 'frame' refers to the output of a single CCD during one readout. Various wideband and narrowband filters are available for use with the Suprime-Cam, among which B (Johnson&Morgan 1953), R_C (Cousins 1978) and i' (Fukugita et al. 1996). The transmission curves of these 3 filters are shown in Figure 2.6.

2.2 Observational Data

Following the observing proposal (see Section A), J.D. Kurk (Sterrewacht Leiden), in collaboration with M. Ouchi (University of Tokyo) and T. Kodama (NAOJ), used the Suprime-Cam on the Subaru telescope to observe the field around TN J1338–1942. During the assigned observing time on the nights of January 31st and February 1st 2003, TN J1338–1942 was visible only during the second half of the night. Both nights were within 24 hours of new moon, thereby minimising contamination from sky background radiation.

Limited by an air mass of less than 1.5 and astronomical dawn, the observation window for TN J1338–1942 was 2.5 hours per night. During this time, images of the TN J1338–1942 field were taken using 3 wideband filters: B, R_C and i'. The individual integration times per image in B, R_C and i' were set to 1200, 540 and 300 seconds, respectively, which is a trade-off between maximizing the number of frames for better flat-fielding and cosmic ray removal and minimizing the number of frames for less overhead of CCD readout time. As the sky background noise is stronger at longer wavelengths, the mean sky background noise level per image was roughly the same in all three filters (~ 5000 ADUs per pixel). Also, the B-band images were made closest to astronomical dawn because of their lower relative insensitivity to sky background noise. Table 2.1 contains more details about these observations.

For each filter, the position of the individual images was jittered (offset) by 20 arcsec, to prevent unexposed areas in the combined images due to inter-CCD gaps, dust specks on the CCDs and masked (bad) CCD pixels. Stacking jittered images also removes cosmic rays. During the first night, the last B-band image was too close to astronomical dawn, resulting in a sky background that was 3 times higher than in the 6 other B-band images. Using all 7 B-band images would lower the signal-to-(background-)noise ratio (SNR) by 5% compared to using 6 images. Therefore, the 7th B-band image was discarded.

During the first and second night standard star images were made for photometric calibration of the B-, R_C- and i'-band magnitudes (see Section 3.2). The integration time was 10 seconds per image to avoid saturation. Table 2.2 gives an overview. Air masses ranged from $1.02 - 1.14$, somewhat lower than for the TN J1338–1942 images. The possible implications are discussed in Section 3.2.2.

During evening and morning twilight on both nights, 24 B-band sky-flat images were made
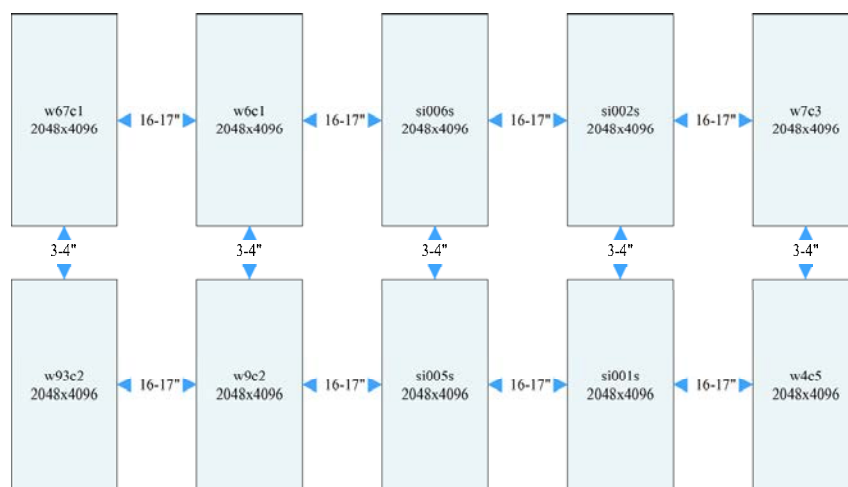
Figure 2.4: Schematic layout of the 10 CCDs on the Suprime-Cam, including the approximate sizes of the inter-CCD gaps.
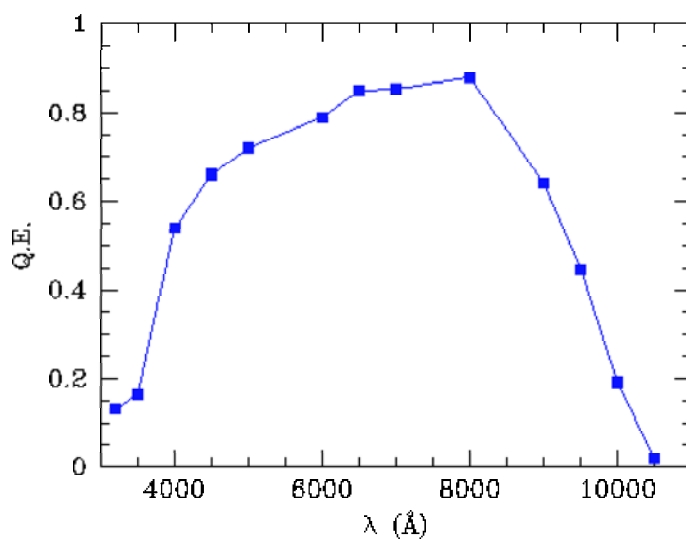
Figure 2.5: Quantum efficiency (QE) of the Suprime-Cam CCDs as a function of wavelength.

with integration times varying between $3-200$ seconds. Of these, 17 images had a sky background level that was within the desired range of $5000-15000$ ADUs. These images turned out to be crucial for flat-fielding the B-band images (see Section 3.1.1).
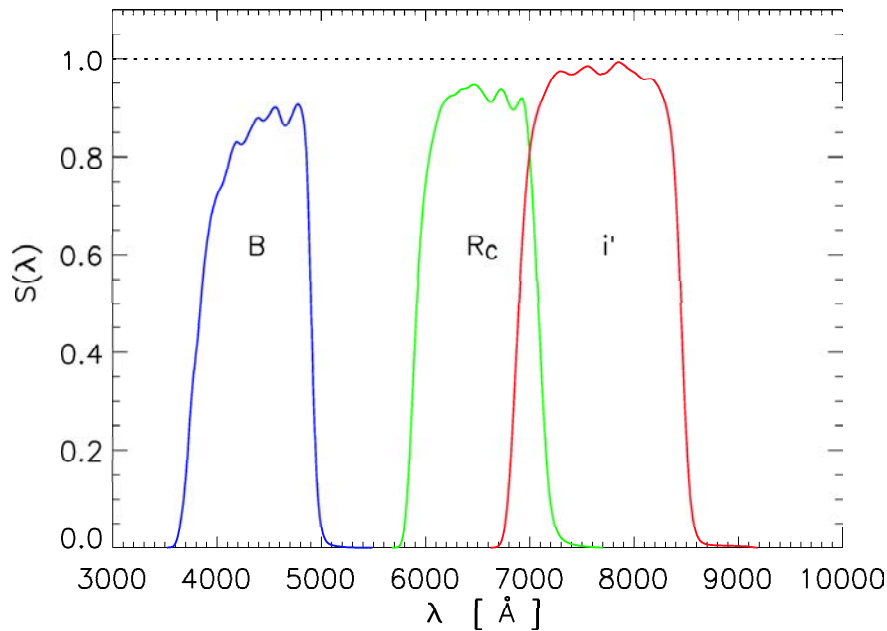
Figure 2.6: Transmission curves of the three selected wideband filters B, R_C and i'.

Filter	Number of Images $[1^{st}+2^{nd}$ night]	Integration Time per Image [s]	Integration Time in Total [m]	Air Mass Range
B	3+4=7 (6)	1200	140 (120)	1.30-1.33
R_C	6+5=11	540	99	1.37-1.66
i'	6+9=15	300	75	1.31-1.42

Table 2.1: Observational data on the imaging of TN J1338–1942. Note that one of the 7 B-band frames was discarded due to high sky background noise.

Standard Star Field	Filter	Number of Images	Number of Stars
Landolt SA95	B	2	10+10=20
Landolt SA107	B	3	9+7+7=23
Landolt SA104	R_C	2	6+7=14
Landolt GD71	i'	2	1+1=2
Hickson 0829+33	i'	1	1

Table 2.2: Observational data on the imaging of the standard star fields (Landolt 1992; Hickson&Mulrooney 1998).

Chapter 3

Data Reduction

3.1 Reducing Images

3.1.1 Reduction Pipeline

A major part of the data reduction was done using the SDFRED software (Subaru Deep Field REDuction v20030220). SDFRED was developed for the data reduction of the Subaru Deep Field (SDF) and Subaru/XMM Deep Field (SXDF) (Ouchi et al. 2001; Ouchi et al. 2004a). SDFRED performs the following actions on Suprime-Cam frames and images:

- Overscan bias subtraction & overscan trimming: Subtraction of the CCD bias level and gradient from frames, calculated from the overscan region. Removal of the overscan region.

- Flat-field generation & flat-fielding: Creation of flat-field frames for each CCD. Dividing out local variations in the exposure and sensitivity of CCDs.

- Distortion correction & differential effect correction: Correction for the atmospheric refraction in the Suprime-Cam's FOV and the field distortion caused by the Suprime-Cam's CCD geometry.

- PSF matching: Convolution with gaussian kernels to match the PSF sizes between images.

- Sky subtraction: Subtracting sky background flux from frames.

- Masking of auto guider probe shades: Invalidating areas where the auto guider probe blocks the FOV.

- Masking bad regions: Invalidating areas in frames that are unusable (e.g. bad pixels).

- Matching & mosaicing: Determining relative positions, rotation angles and sensitivities of images. Stack multiple images into a single image.

SDFRED is a collection of Unix C-shell scripts and C-code executables, which are to be executed in a specific sequence. It makes use of the following software packages:

- Nekosoft v20031223 (Yagi 1996): C-codes for distortion correction, sky subtraction and mosaicing of Suprime-Cam frames.

- SExtractor v2.2.1 (Bertin 1996): Program for extraction of objects from an astronomical image, including astrometry and photometry.

- IRAF v2.12.1-EXPORT (e.g. Jacoby 1998): General purpose software system for the reduction and analysis of astronomical data.

For each filter, the TN J1338–1942 field images were reduced using SDFRED. The flat-field frames were generated from the TN J1338–1942 field images. In the case of the B-band, the flat-fields were of poor quality, due to the relatively low number of input frames. Instead, the flat-fields were made from 17 B-band sky-flat images. The 3 resulting stacked images showed a flat background on both small and large scale, with a well determined noise level (see Section 3.3.2), except at the edges where the sky was only covered by a relatively small number of jittered images. This resulted in relatively high sky background noise and small gaps (incomplete sky coverage) around the edges.

Further reduction and analysis was done using SExtractor, IRAF and other software packages:

- IDL v6.0 (e.g. Fanning 2003): Interactive Data Language for data analysis, visualization, and cross-platform application development.

- IDL Astronomy Library (Landsman 1993): A collection of low-level astronomy software routines in the commercial language IDL.

- IDL ATV v1.5 (Barth 2001): An interactive display and analysis tool for astronomical images.

- SkyCat v2.0-GAIA (Albrecht 1997): A tool that combines image visualization with access to astronomical catalogs and data archives.

- SAOImage DS9 v3.0 (Joye&Mandel 2003): An astronomical imaging and data visualization application.

3.1.2   Image Alignment

For automated source detection (see Section 3.3.1), the B-, R_C- and i'-band images needed to match in pixel scale size and in sub-pixel accuracy in position. Rectangular cut-outs were made from each image (using IRAF→imcopy) in such a way that the noisy (local noise level larger than approximately 3 times the global noise level) and incomplete edges of each image were trimmed off and the (mean) positional offsets between the images were taken out, while the total FOV of the images was kept as large as possible. The mean positional offset between the 3 images was determined by measuring the pixel positions of 5 unsaturated stars across the FOV (using SkyCat→Pick Object). The FOV of the resulting images was 7401×7251 pixels or 24.67×24.17 `arcmin`^{2}.

Although the 3 images had a good pixel-to-pixel match in the center, there was still a mismatch ($\sim$ 1 pixel) at the edges, due to small differences in image scales. To resolve this, the central positions of $\sim$ 1000 unsaturated stars, covering the whole FOV, were extracted (using SExtractor) with sub-pixel accuracy in all 3 images. The stellar positions were used to fit 2-dimensional, 3rd order transformation polynomials, with which the B- and i'-images were transformed (using IRAF→geomap and →geotran) to match the R_C-image. The R_C-image was used as the reference because it has the narrowest PSF FWHM (see Section 3.1.4) and thus the most accurate positions. The rms deviation in relative stellar positions after the transformation was 0.08 pixel or 0.015 `arcsec` for the B-band image and 0.05 pixel or 0.010 `arcsec` for the i'-band image.

3.1.3   Astrometric Calibration

Due to the precise alignment of the B-, R_C- and i'-band images, absolute astrometric calibration was performed on the R_C-band image only. Within the FOV, a set of $\sim$ 350 unsaturated stars ($R_C > 18.0$) was selected from the USNO-A2.0 catalog (Monet et al. 1998). For each star, the central pixel coordinates of the matching star in the R_C-band image were determined (using SkyCat→Tweak an existing calibration). A 2-dimensional, 2nd order polynomial was fitted to

map the pixel coordinates to the (USNO) celestial coordinates (using IRAF→ccmap and cc-setwcs). The resulting astrometric calibration was copied from the R_C-band image to the B- and i'-band images (using IRAF→wcscopy).

The USNO-A2.0 catalog is constructed from digitized photographic plates that are up to 50 years old. This means that nearby stars may have moved a considerable angular distance, up to a few `arcsec`. However, most stars are sufficiently distant to have moved only a little so that the mean stellar reference frame provided by this catalog maintains an accuracy better than 0.3 `arcsec` (Monet et al. 1998). The central position of each detected object is accurate up to a fraction of the PSF FWHM, ~ 0.2 as. Including the alignment accuracy between the images gives a total astrometric accuracy of ~ 0.4 `arcsec`.

3.1.4 PSF Matching

Each point on the sky produces a gaussian-like PSF in the telescope's image plane. This is due to the telescope's beam pattern and to diffraction in the Earth's turbulent atmosphere (seeing). The FWHM of the beam pattern of a circular aperture telescope like the Subaru telescope is given by FWHM = $1.22 \times \lambda/D$ (radians), where λ is the observed wavelength and D is the diameter of the telescope's primary mirror. For the Subaru telescope, $D = 8.2$ meters and $\lambda = 3500 - 10500$ Å (see Section 2.1.2), so FWHM = $0.01 - 0.03$ `arcsec`.

As magnitudes are measured in a fixed size circular aperture (see Section 3.3.1), the PSF FWHM needs to be equal in all 3 images. Measurements on 20 unsaturated stars (using SkyCat→Pick Object) in different parts of the FOV showed that the mean PSF FWHM of stars in the B-, R_C- and i'-band images are 0.98, 0.82 and 0.92 `arcsec`, respectively. From this it is clear that seeing is mainly responsible for the observed PSF FWHM. We convolved the R_C- and i'-band images with gaussian kernels (using IRAF→psfmatch) to create a PSF FWHM of 0.98 `arcsec` in all 3 images. Figure 3.1 shows the resulting images combined into one false-color image.

3.2 Photometric Calibration

3.2.1 Standard Stars

By using the standard star images (see Section 2.2), which contain stars with well determined apparent magnitudes, we are able to calibrate the relation between measured counts (or instrumental flux) F_m and apparent magnitude m. The basic equation for this calibration is given by:

$$m = m_0 + 2.5 \log_{10} (T_m \text{ [sec]}) - 2.5 \log_{10} (F_m \text{ [ADUs]}) \tag{3.1}$$

Here, T_m is the total integration time and m_0 is the magnitude zero-point. The latter depends on the used filter (in our case $m = B$, R_C or i') and the magnitude convention. In this report, magnitudes are expressed in the AB-system (AB$_{95}$; Fukugita et al. 1996).

Table 2.2 in Section 2.2 showed an overview of the standard star fields that were imaged during the assigned observing time. All star fields were imaged under photometric conditions. Standard stars in the SA95, SA107, SA104 and GD71 fields (Landolt 1992) have known magnitudes in the UBV- (Johnson&Morgan 1953) and $R_C I_C$-bands (Cousins 1978), expressed in the Vega-system. The required B-, R_C- and i'-band AB-magnitudes were obtained by use of conversion formulae and tables in Fukugita et al. (1996). The i' AB-magnitude of standard star Hickson 0829+33 (Hickson&Mulrooney 1998) was available through the SDSS DR3 catalog (Abazajian et al. 2004).

The standard star images were reduced with the SDFRED pipeline (see Section 3.1.1) using the same flat-fields that were used on the TN J1338–1942 images, but without the final matching & mosaicing step. For all available standard stars (see Table 2.2, column 4), the counts were measured (using IDL→ATV→Photometry) within a circular aperture of 25 pixels in radius,
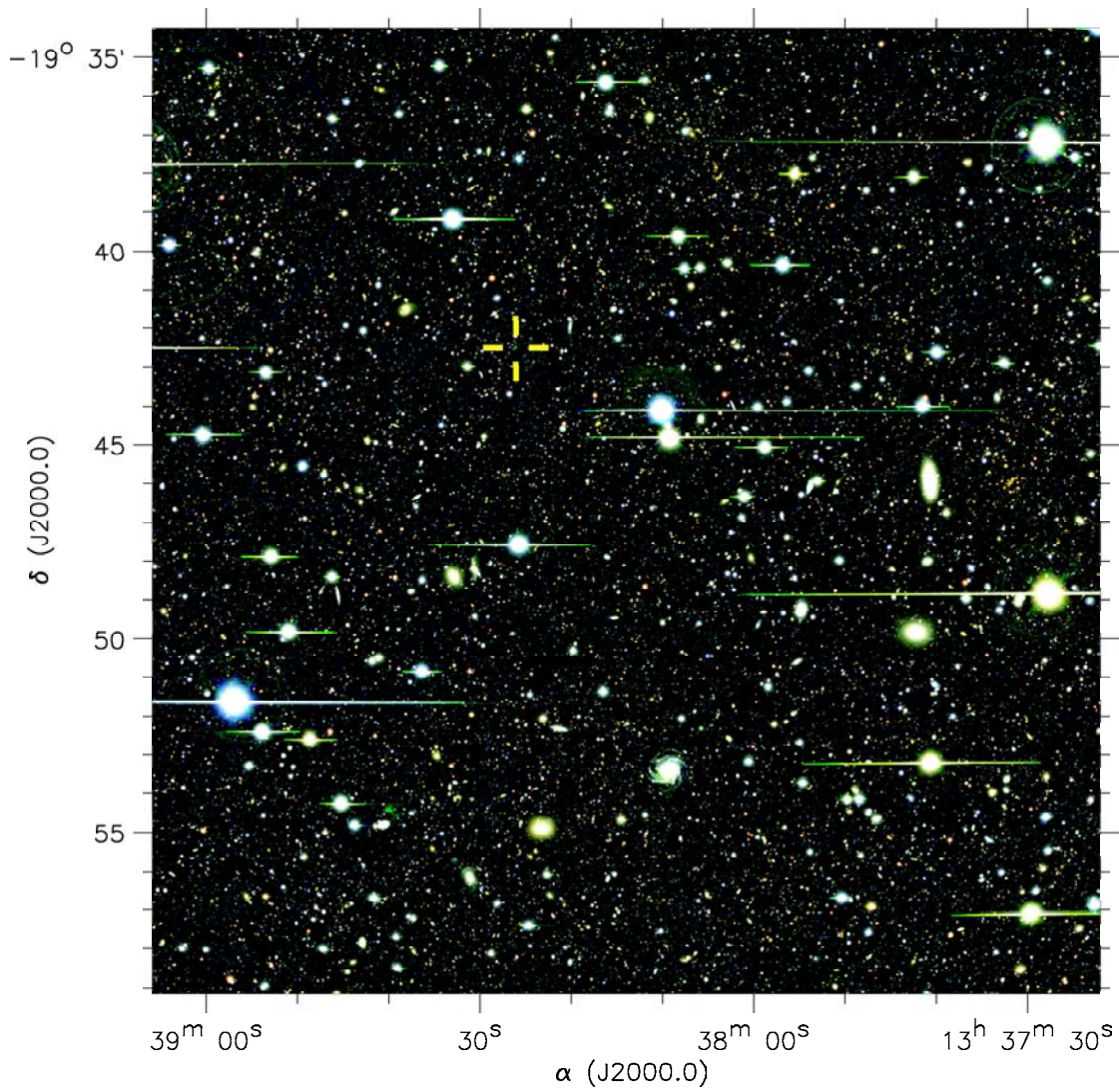
Figure 3.1: False color composite of the reduced images, with B mapped to blue, R_C to green and i' to red. The total area of the image is 7251×7401 pixels or $24.17 \times 24.67 = 596.3$ `arcmin`^{2}. Marked by the yellow crosshair is the location of TN J1338–1942.

while the background was estimated in an annulus between $30 - 50$ pixels in radius. Using Equation (3.1) with $T_m = 10$ seconds, each standard star gives a value for m_0. Averaging these values for each filter, we find $B_0 = 27.39 \pm 0.02$, $R_{C,0} = 27.63 \pm 0.02$ and $i'_0 = 27.82 \pm 0.02$.

3.2.2 Air Mass

Ideally, standard star images are made at the same zenith angle as the object under study and during the same part of the night. This way, the effects of varying air mass are minimised, while calibrating the magnitude zero-points for the different filters. However, this was not quite the case for our standard star observations. Time differences range up to more than 24 hours. Also, as stated in Section 2.2, the air mass (a function of the zenith angle) of the standard star frames is lower than the air mass of the TN J1338–1942 frames.

Air mass (AM) is directly related to atmospheric extinction (A_m). For a simple, plane-parallel atmospheric model, this relation is given by:

$$A_m = -2.5 \ \log_{10}\left(\exp\left(-AM \ \tau_m\right)\right) \approx 1.1 \ AM \ \tau_m \tag{3.2}$$

Here, τ_m is the mean optical depth of the atmosphere at zenith within filter m. As long as τ_m is small ($\ll 1$) and $AM \sim 1$, differences in AM cause only small ($\ll 1$) variations in the magnitude zero-point. As we have no means of evaluating τ_m for $m = B$, R_C or i', we must assume that $\tau_m \ll 1$, so that the magnitude zero-points determined from the standard star images are equal to the magnitude zero-points of the images of the TN J1338–1942 field. For example, suppose that $\tau \lesssim 0.2$, then an air mass difference of $\lesssim 0.5$ causes an error in the magnitude zero point of $\lesssim 0.1$ magnitude.

3.2.3 Galactic Extinction

The Earth's atmosphere is not the only source of extinction. For objects like LBGs, the observed photons have traveled through both intergalactic and interstellar space. As we are relying on the intergalactic matter to cause the Lyman break in an LBG spectrum, we are only correcting for galactic extinction by interstellar matter. Using a galactic extinction calculator (based on Schlegel et al. 1998), the color excess over the FOV is found to be $E(B - V) = 0.094 \pm 0.006$, based on 9 measurements throughout the FOV. The galactic extinction in B, R_C and i' is given by $A_B = 4.325 \ E(B - V) = 0.41 \pm 0.03$, $A_{R_C} = 2.634 \ E(B - V) = 0.25 \pm 0.02$ and $A_{i'} = 2.086 \ E(B - V) = 0.20 \pm 0.01$.

3.2.4 Magnitude Zero-points

During the matching & mosaicing step in the reduction pipeline, multiple images are stacked to create a single resulting image for each filter, which is the *average* of the stacked images. In the case of N_m stacked images, a count of f_m ADUs in the resulting image represents a total measured count of $F_m = N_m \times f_m$ ADUs. The total integration time T_m is N_m times the integration time t_m per image. Combining this with Equation (3.1) and the galactic extinction, the magnitude equation is given by:

$$
\begin{aligned}
m &= m_0 + 2.5 \log_{10}\left(T_m\right) - 2.5 \log_{10}\left(F_m\right) - A_m \\
&\equiv C_{m,0} - 2.5 \log_{10}\left(f_m\right)
\end{aligned}
\tag{3.3}
$$

By defining $C_{m,0} = m_0 + 2.5 \log_{10}\left(t_m \ [s]\right) - A_m$, it is possible to combine all filter-related constants into one. It follows that $C_{B,0} = 34.68 \pm 0.03$, $C_{R_C,0} = 34.21 \pm 0.03$ and $C_{i',0} = 33.82 \pm 0.03$.

3.3 Catalog Generation

3.3.1 Source Extraction

Searching for LBGs at $z \sim 4$ in the B-, R_C- and i'-band images meant searching for very faint objects just above the background noise. As there were many of these (and other) objects within the FOV, the automated object detection program SExtractor was used. To obtain as many faint objects as possible, while keeping false detections to a minimun, SExtractor was configured to detect an object when at least 7 adjacent pixels had a count level above 1σ of the (local) background noise. De-blending parameters were set to maximize separation of partially overlapping sources. These settings for SExtractor were determined empirically by maximizing the number of detected faint objects while minimizing false detections (see Section 3.3.4). Furthermore, SExtractor was configured to carry out photometry on the detected objects.

With SExtractor, it is possible to detect an object in one image and do the object's photometry in another, as long as the images are aligned to sub-pixel accuracy (see Section 3.1.2) over the full FOV. As LBGs at $z \sim 4$ are expected to show up in both the R_C- and i'-band images (but not in the B-band image), the object detection was done on an image that was the sum of the R_C- and i'-band images (to maximize detectability of faint LBGs), while the photometry was done successively on the B-, R_C and i'-band images. This way, a raw catalog of $\sim 120,000$ objects was generated to include for each object an identification number, its center position in Celestial and (sub-)pixel coordinates and the measured counts and magnitude in 3 filters and 2 types of apertures.

SExtractors main photometry of detected objects was done in a fixed size circular aperture of 2 `arcsec` in diameter (using MAG_APER). The most important goal was to have well defined colors, as the selection of LBG candidates is done using the colors $B - R_C$ and $R_C - i'$. By using a relatively small aperture that was centered on each object, colors were determined from the most luminous part of each object. As the fractional count loss outside the aperture is the same in all filters, the apparent magnitudes B, R_C and i' are offset by the same amount. It follows that (ideally) count loss has no effect on the colors. Under the assumption that the spatial structure is the same in all bands, this is also true for objects that are resolved. A magnitude correction for the count loss was applied afterwards (see Section 3.3.3).

The aperture size is a trade-off between maximizing the detectable object counts (increasing the aperture area) and minimizing the background noise (decreasing the aperture area). As most of the LBGS at $z \sim 4$ are unresolved (angular size < 0.98 `arcsec`), they can be treated like point sources with near-gaussian luminosity profiles. A 2 `arcsec` aperture on these objects misses only a small portion of the object counts. When perfectly centered on an unresolved source with a perfect gaussian profile with a FWHM of 0.98 `arcsec` and no background noise, 2 `arcsec` aperture misses only $\sim 6\%$ of the counts. This implies a (negative) magnitude correction of ~ 0.06. In reality, this correction is somewhat larger (see Section 3.3.3). An additional advantage of using this aperture is that measurement of the background noise level is fairly simple (see Section 3.3.2).

To estimate the magnitude correction due to count loss (see Section 3.3.3) SExtractor was configured to perform a second photometry on each object, using an elliptical aperture of flexible size and orientation (using MAG_AUTO; Kron 1980), based on the . This method was configured to select a larger aperture area to collect more object counts, but at the cost of a higher background noise level.

3.3.2 Noise Measurements

There are two major sources that contribute to the uncertainty $\sigma_{f,m}$ in count measurements, namely the background noise $\sigma_{bg,m}$ (including flat-field imperfections and readout noise, see

Section 2.1.2) and the poissonian noise $\sigma_{ob,m}$ in the object counts:

$$\sigma_{f,m}^2 = \sigma_{bg,m}^2 + \sigma_{ob,m}^2 \tag{3.4}$$

The uncertainty σ_m in the apparent magnitude m follows from:

$$\sigma_m^2 = \left(\frac{\partial m}{\partial f_m}\right)^2 \sigma_{f,m}^2 + \sigma_{m,0}^2 = \left(\frac{-2.5}{\ln(10)\ f_m}\right)^2 \sigma_{f,m}^2 + \sigma_{m,0}^2 \tag{3.5}$$

Here, $\sigma_{m,0}$ is the uncertainty in the magnitude zero-points (Section 3.2.4)

For a poissonian process, the noise is equal to the square root of the counted entities, in our case electrons. For a gain factor $g = 2.6\ \mathsf{e}^-$ per $\mathtt{ADU}$ (see Section 2.1.2), the object counts f_m represents $g \times N_m \times f_m$ electron counts, so the poissonian noise is $\sqrt{g\ N_m\ f_m}$ electrons. Expressed in the same units as f_m, this becomes:

$$\sigma_{ob,m} = \frac{\sqrt{g\ N_m\ f_m}}{g\ N_m} = \sqrt{\frac{f_m}{g\ N_m}} \tag{3.6}$$

A direct method for background noise estimation inside a 2 $\mathtt{arcsec}$ aperture is to measure the total counts within the aperture at a large number of random, but empty, positions (no objects within aperture) in the FOV, create a histogram and fit a gaussian (Labbé et al. 2003). For each filter, there were 50,000 count measurements (using IDL→idlastro→aper) on the images. For the B-band, the resulting histogram and gaussian fit are plotted in Figure 3.2. The plots for R_C and i' are very similar.

From the histogram it is clear that the high end tail of the gaussian profile is contaminated with object counts. Therefore the gaussian fits were limited to $\lesssim +1\sigma$. These fits give the 1σ noise level $\sigma_{bg,m}$ within the 2 $\mathtt{arcsec}$ aperture. For the B-, R_C- and i'-bands, $\sigma_{bg,m} = 265.5$, 249.9 and 219.0 $\mathtt{ADUs}$, respectively.

3.3.3 Aperture Correction

For unresolved sources, it is relatively straightforward to apply a correction for the count loss due to the 2 $\mathtt{arcsec}$ aperture. We apply this correction to all detected sources, under the assumption that most of the LBGs at $z \sim 4$ are unresolved. Figure 3.3 shows the difference between two B-band magnitudes, one measured in an elliptical aperture of variable size and orientation (MAG_AUTO(B); Kron 1980) and the other measured in a 2 $\mathtt{arcsec}$ fixed size circular aperture (MAG_APER(B)). For MAG_APER(B)$\lesssim$ 24, a distinct horizontal trend is present, which represent relatively bright but unresolved objects. Within the high signal-to-noise range $20 \lesssim$ MAG_APER(B) $\lesssim$ 24, it is easy to identify the extended sources below the horizontal line. For these objects the fractional count loss is larger than for unresolved objects. The horizontal line is offset from zero due to count loss outside the 2 $\mathtt{arcsec}$ aperture. A best fit gives a (negative) magnitude offset $\Delta B = 0.20 \pm 0.005$. Plots for R_C- and i'-band magnitudes give similar results, with $\Delta R_C = 0.16 \pm 0.005$ and $\Delta i' = 0.22 \pm 0.005$. This is very similar to the value of 0.2, found by Ouchi et al. (2004a). The differences between the offsets in 3 bands are most likely the result of small imperfections in the PSF matching (Section 3.1.4). Incorporating these offsets in the magnitude zero-points gives $C_{B,0} = 34.47 \pm 0.03$, $C_{R_C,0} = 34.05 \pm 0.03$ and $C_{i',0} = 33.60 \pm 0.03$.

3.3.4 Completeness and Reliability

The completeness and the reliability are important among the parameters that define the quality of the object catalog. Completeness is the number of objects in the catalog compared to the number of available objects in the FOV given a certain magnitude limit, while reliability is the
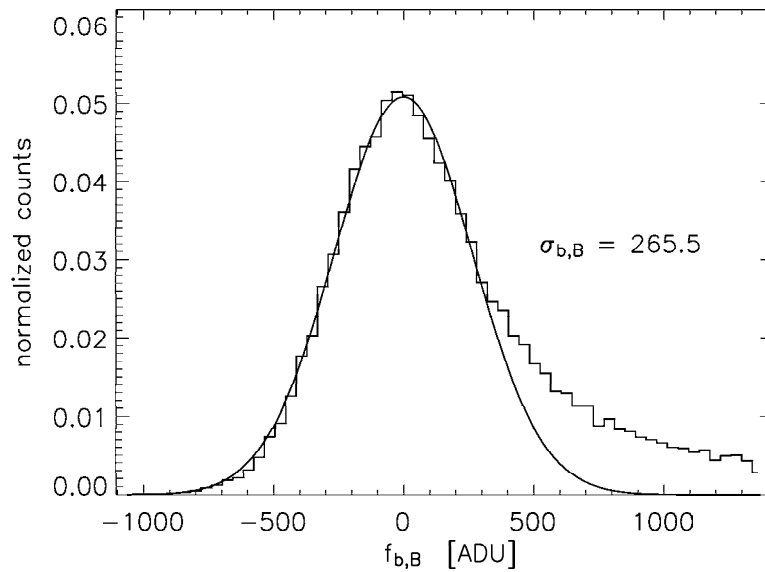
Figure 3.2: Histogram of 50,000 background count measurements in the B-band image. Discarding the high end tail (due to object count contamination), the B-band background noise $\sigma_{bg,B}$ can be estimated by fitting a gaussian (fat solid line).
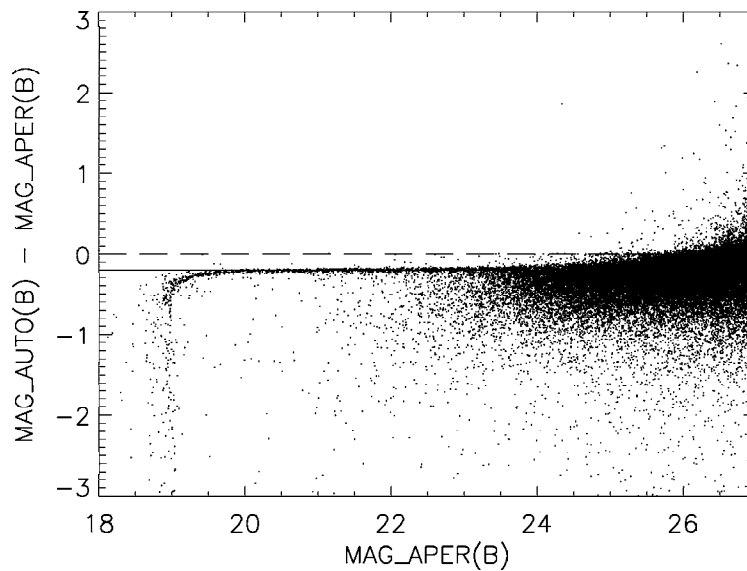
Figure 3.3: Differences between B-band magnitudes in two different apertures (MAG_AUTO(B) and MAG_APER(B)) due to count loss in the second aperture. The dashed line denotes a zero difference between the two apertures, while the solid line is the best fit to the constant (negative) offset in the range $20 \lesssim$ MAG_APER(B) $\lesssim 24$. On the left, the effects of saturation of the brightest objects enlarges the negative offset, while on the right the offset gets swamped by faint objects with large photometric uncertainties.

number of true objects in the catalog compared to the total number of objects in the catalog. Completeness and reliability are dependent on magnitude, as both drop when approaching the detection limit of an image.

The detection limit is the minimum SNR times the background noise $\sigma_{bg,m}$. Using Equation 3.3, Table 3.1 gives the magnitude detection levels corresponding with SNR $= 1 - 5$. The bottom row shows the reliability when we assume that the only source of false detections is gaussian background fluctuations. There are, however, other possible sources of false detections (e.g. bleeders, fringes around bright stars, bad flat-fielding, etc.). Practically all of these can be taken out by visual inspection, but this can be time-consuming in the case of $\sim 120,000$ objects. Therefore, visual inspection is done later on the (smaller) sample of LBG candidates (see Section 4.1.1).

For completeness estimates assumptions have to be made about objects that are undetected. A fairly simple and straightforward way is plotting the (logarithmic) number density versus magnitude bin (see Figure 3.4). Assuming that the straight line, fitted on $i' < 25$, can be extrapolated (e.g. Steidel&Hamilton 1993, Palunas et al. 2004), the catalog starts to become incomplete for $i' > 25$. Completeness is $\sim 93\%$ at $i' = 25.25$, $\sim 82\%$ at 25.75, $\sim 72\%$ at 26.25 and $\sim 61\%$ at 26.75.
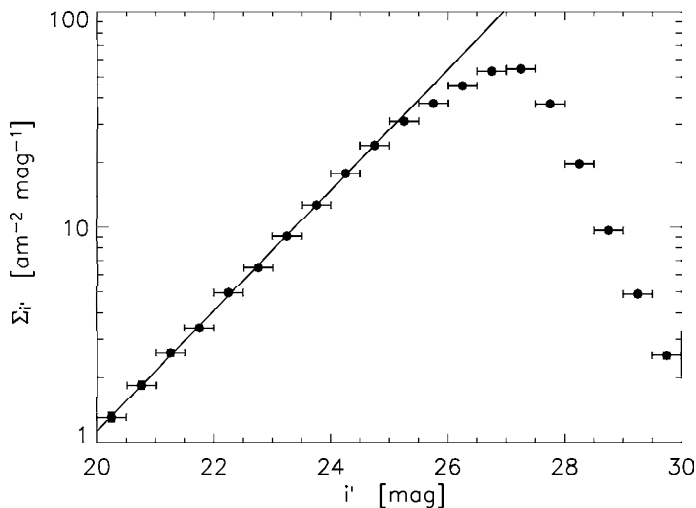
Figure 3.4: Number counts of $\sim 120,000$ detected objects, sorted into $\Delta i' = 0.5$ magnitude bins (black dots, including uncertainty bars). Assuming the fitted straight line can be extrapolated to higher magnitudes, the completeness of the object catalog drops for $i' \gtrsim 25$.

Filter	$1\sigma_{bg,m}$	$2\sigma_{bg,m}$	$3\sigma_{bg,m}$	$4\sigma_{bg,m}$	$5\sigma_{bg,m}$
B	28.42	27.67	27.23	26.92	26.67
R_C	28.06	27.31	26.87	26.55	26.31
i'	27.75	26.99	26.55	26.24	26.00
Reliability	84.1%	97.7%	99.9%	$\sim 100\%$	$\sim 100\%$

Table 3.1: Limiting magnitudes for different filters and different SNRs and the reliability in the case that false detections are caused only by gaussian background fluctuations.

Chapter 4

Analysis

4.1 LBG Candidates

4.1.1 Selection Criteria

As discussed in Section 1.4, LBGs are detected by measuring and comparing their magnitude in wideband filters on the short and long wavelength side of the redshifted Lyman break. Ouchi et al. (2004a) searched for LBGs at $z = 3.5 - 4.5$ in the blank SDF and SXDF fields using the Suprime-Cam images with wideband filters B, R_C, and i'. Their selection criteria for selecting LBG candidates are based on detecting absorption by neutral hydrogen from both the Lyman break ($\lambda < 912$ Å) and the Lyman forest ($\lambda < 1216$ Å) in the flat UV-continuum emission of star-forming galaxies.

For galaxies at $z = 3.5 - 4.5$, the Lyman break is redshifted to $\sim 4100 - 5000$ Å, while the start of the Lyman forest is redshifted to $\sim 5500 - 6700$ Å. From Figure 2.6 it follows that the redshifted Lyman break is positioned within the B-band, while the Lyman forest starts at the blue edge of the R_C-band. Therefore, luminous LBGs at $z \sim 4$ are detectable in both the R_C- and i'-band images, while heavily dimmed or undetectable in the B-band image.

We assumed that the selection criteria from Ouchi et al. (2004a) were also applicable to our data, as we tried to adapt to their reduction and photometry methods as closely as possible. The selection criteria were constructed using the colors of model objects and real objects, as shown in Figures 4.1 and 4.2, taking into account larger photometric uncertainties for fainter objects. The selection criteria are a trade-off between selecting as many high redshift ($z > 3$) galaxies as possible, while keeping the number of contaminants (lower redshift galaxies ($z \leq 3$) and galactic stars) as low as possible (see Section 4.1.2). They are given by:

$$B - R_C > 1.2, \quad R_C - i' < 0.7, \quad B - R_C > 1.6\,(R_C - i') + 1.9 \tag{4.1}$$

To ensure some level of photometric completeness, only objects with magnitude $i' < 3\sigma_{bg,i'}$ were included, while the B magnitude was clipped to $1\sigma_{bg,B}$ in case of a non-detection in B (meaning $B > 1\sigma_{bg,B}$).

From the $\sim 120,000$ objects in our raw catalog (see Section 3.3), $\sim 62,000$ were selected with $i' < 3\sigma_{bg,i'} = 26.55$. Figure 4.3 shows the colors of these objects. While clipping the B-magnitude to $1\sigma_{bg,B} = 28.42$, 917 objects were selected using the selection criteria in Equation (4.1). We were cautious to use the term 'LBG candidates' already because false detections were not taken out yet (see Section 4.1.2). The selection rule $R_C - i' < 0.7$ makes sure that the R_C-magnitude is at least lower than $2\sigma_{bg,R_C} = 27.31$. In fact, it was found that $R_C < 3\sigma_{bg,R_C} = 26.87$ for all 917 selected objects.

Figure 4.3 shows that the bright i'-objects (red dots) trace out some structure in the color-color diagram, while the structure in faint objects (black dots) is much more diffuse. This is

23
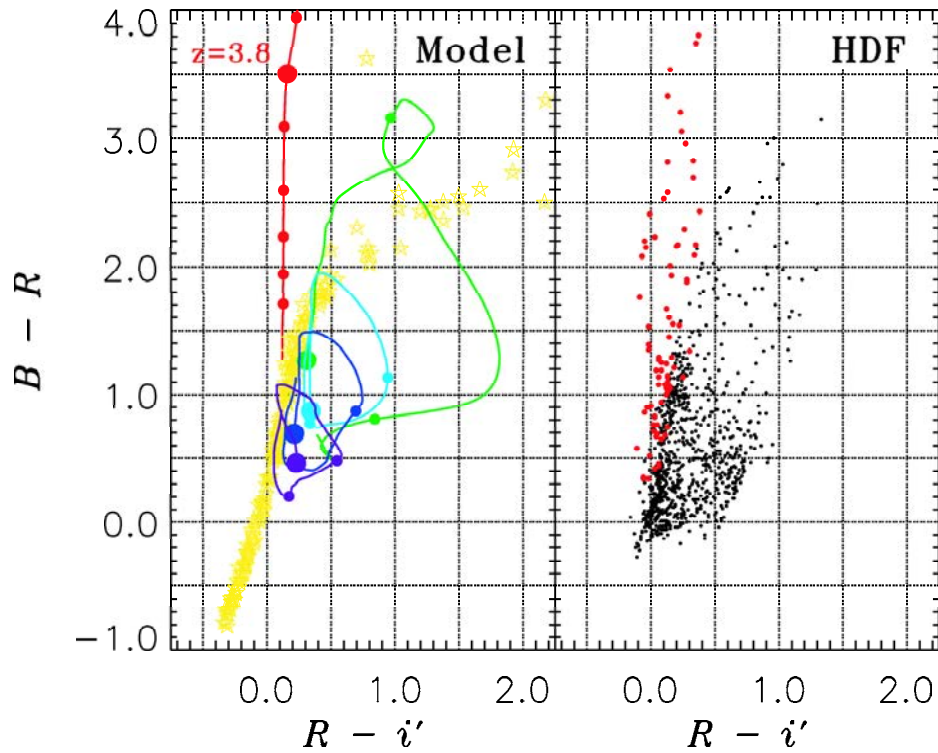
Figure 4.1: *Left*: Colors of model galaxies and 175 galactic stars (yellow). The red line shows the track of a typical $z = 3$ star-forming galaxy, artificially redshifted from $z = 3$ to $z = 4$ (dots at $z = 3.3$ to 3.9). The green, cyan, blue and violet lines are (typical) elliptical, Sbc, Scd and irregular galaxies, artificially redshifted from $z = 0$ to $z = 3$ (dots at $z = 0$, 1 and 2). *Right*: Colors of 1048 HDF-N galaxies, whose photometric redshift is indicated by a black dot ($z \leq 3$) or a red dot ($z > 3$). Adopted from Ouchi et al. (2004a).
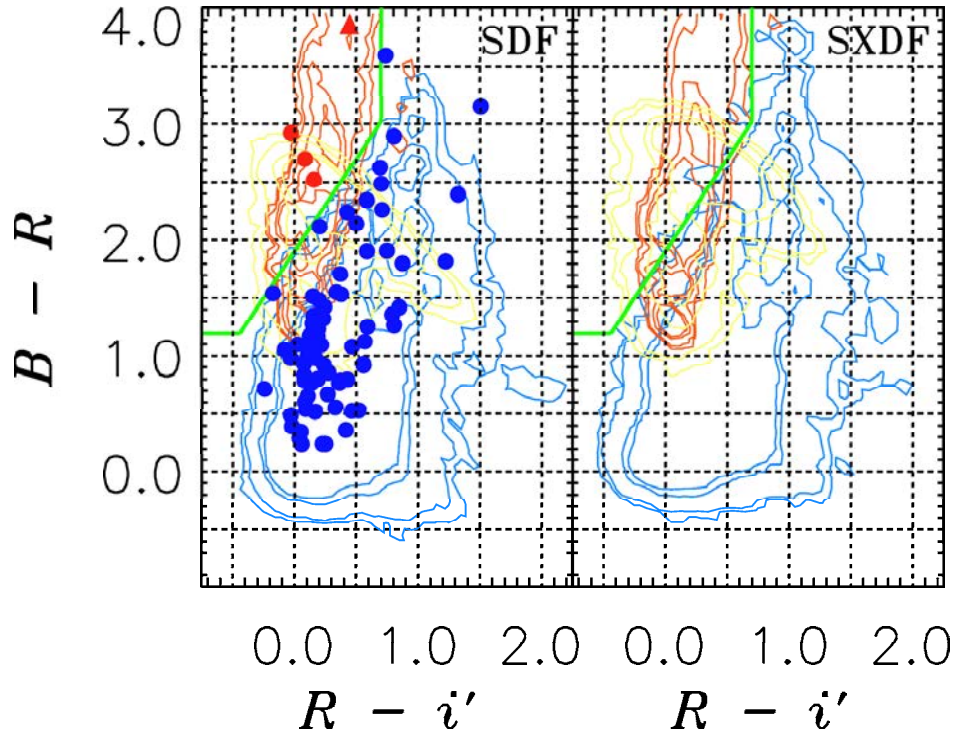
Figure 4.2: Probability maps of high redshift ($z > 3$) and low redshift ($z \leq 3$) galaxies in both the SDF and SXDF fields, as estimated by Monte Carlo simulations using the HDF-N photometric redshift catalog. Blue lines show contours that include 99.5, 99.0, 95.0 and 90.0% of the detected low redshift galaxies (from edge to center), while red and yellow lines show contours that include 95, 90, 70 and 50% of the detected high redshift galaxies for bright ($i' = 23.5 - 24.0$) and faint ($i = 25.5 - 26.5$) galaxies, respectively. For the SDF, red and blue dots represent 85 galaxies with spectroscopically determined redshifts in the range $z > 3$ and $z < 3$, respectively. Adopted from Ouchi et al. (2004a).
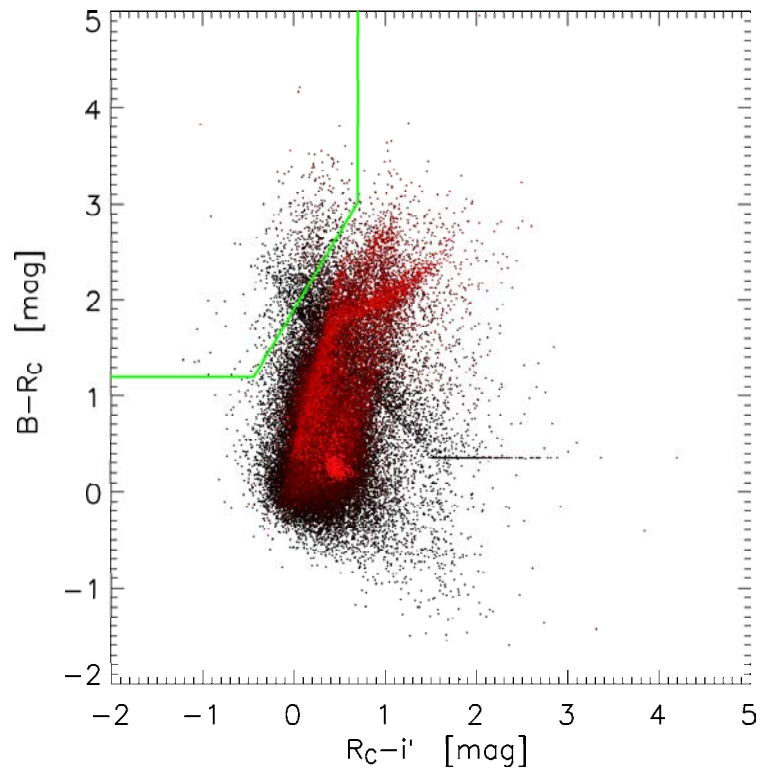
Figure 4.3: Color-color diagram of $\sim 62,000$ detected objects with $i' < 26.55$. High to low i'-magnitude is color-coded from black to red. Marked in green is the area of the LBG selection criteria from Ouchi et al. (2004a). The horizontal feature on the right is due to objects that are undetected in B and R_C, while the diagonal trend from upper-left to lower-right is caused by objects that are undetected in B and very close to the detection limit in i'.

caused by the larger photometric uncertainties for faint objects. In Figure 4.3, the objects selected by the criteria in Equation (4.1) do not seem to trace out a redshift dependent track like the high redshift galaxies in Figure 4.1. However, as most LBG candidates are faint objects, the photometric uncertainties can cause scatter that conceals the presence of such a track. Intrinsic differences between LBGs also might increase the scatter.

4.1.2 Completeness and Contamination

As mentioned in Section 3.3.4, automated source detection introduces the possibility of false object detections other than background noise fluctuations, such as bleeders and fringes around bright stars. To take these out, the 917 selected objects were visually inspected in all 3 images, using a custom-made inspection tool. Examples of these inspections are presented in Figure 4.4. This way, 43 of the 917 objects ($\sim 5\%$) were rejected for being false detections, resulting in a final catalog of 874 LBG candidates (which includes TN J1338–1942). Table 4.1 shows the distribution of the 874 LBG candidates over i'-magnitude intervals.

The average number density of LBG candidates in the total FOV is 1.47 ± 0.05 `arcmin`$^{-2}$ (poissonian uncertainty) to a depth of $i' = 26.55$. This is lower than the number density of 2.65 `arcmin`$^{-2}$ found in the SDF, but higher than the 1.12 `arcmin`$^{-2}$ found in the SXDF (Ouchi et al. 2004a). Possible causes for these differences are discussed in Section 5.1.2. For the completeness and reliability estimates of our catalog, we used the simulation results of Ouchi et al. (2004a) concerning the SDF and SXDF. To compensate for differences between the data sets, we approximated by averaging between the SDF and SXDF results (where needed). The completeness $C(i', z)$ as a function of i'-band magnitude and redshift was determined by inserting and extracting artificially redshifted model galaxies (see Section 4.1.1) of varying magnitudes into and from the real data. Results were sorted into redshift bins of $\Delta z = 0.2$ and magnitude bins of $\Delta i' = 0.5$. The contamination $f_c(i')$ as a function of magnitude was determined by investigating the number of stars, low redshift ($z < 3$) galaxies and high redshift ($z > 3$) galaxies in the simulations and in the HDF-N that were selected by the criteria in Equation (4.1). Note that the magnitudes in the Ouchi et al. (2004a) paper are uncorrected 2 `arcsec` aperture magnitudes, which are higher than our magnitudes by 0.2 magnitude.

Figures 4.5 and 4.6 show the completeness and contamination estimates, as well as their averages, weighted over all i'-band magnitudes. The solid line in Figure 4.5 shows how the selection criteria in Equation (4.1) pick out LBG candidates in the approximate redshift range $z = 3.5 - 4.5$. The average contamination in our sample is estimated at 3.2%.

4.2 Spatial Distribution

4.2.1 Density Contours

The candidate LBGs were selected from a redshift range of $z \approx 3.5 - 4.5$, which corresponds to a depth of ~ 700 `Mpc`. In this redshift range, the FOV covers an area of $\sim 50 \times 50$ `Mpc`^{2}. So we were probing a volume which has a depth that is ~ 14 times as large as its width. The observed LBG density in this volume is $\sim 6 \times 10^{-5}$ `Mpc`$^{-3}$.

By looking through this volume, any galaxy structure that is present is projected along the line of sight, as there is no precise information available about the exact distance (or redshift)

i' range	$21-22$	$22-23$	$23-24$	$24-25$	$25-26$	$26-26.55$
LBG count	1	0	15	109	390	359

Table 4.1: Magnitude distribution of LBG candidates.
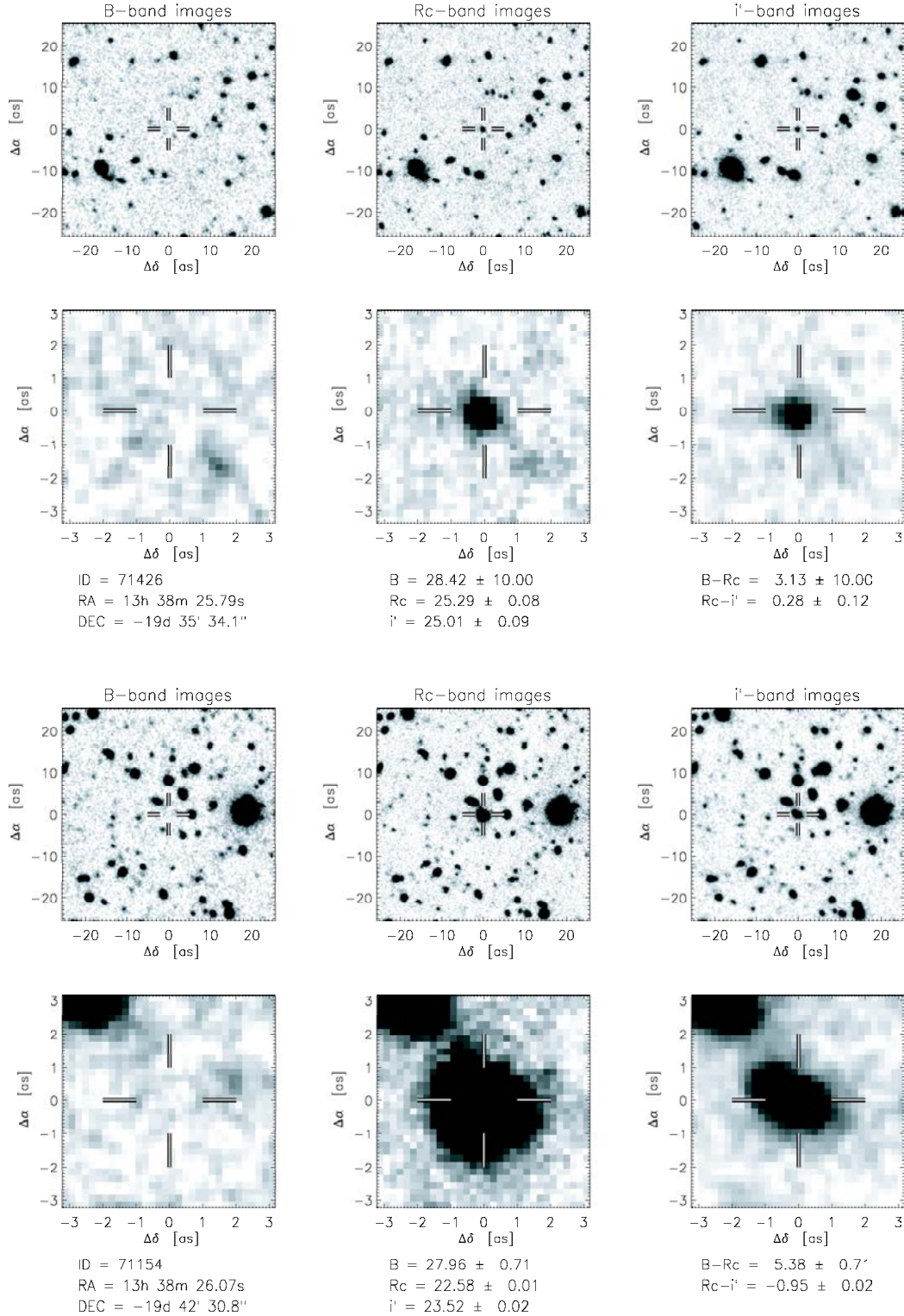
Figure 4.4: Example of 2 objects under inspection using a custom-made inspection tool. For each object, the B-, R_C- and i'-band images were plotted on 2 magnification scales ($1\times$ and $8\times$). The object displayed in the upper 6 panels is a typical example of an unresolved LBG candidate. The object in the lower 6 panels is TN J1338–1942. (An uncertainty in magnitude of 10.0 means that the object's magnitude was above the detection limit for that band.)
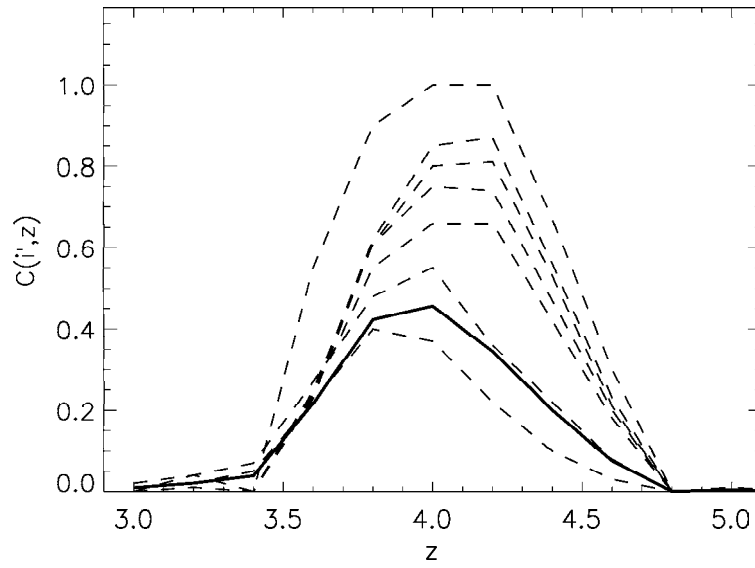
Figure 4.5: Completeness $C(i', z)$ as a function of i'-band magnitude and redshift. From top to bottom, dashed lines represent i'-band magnitudes 23.05 and below, 23.55, 24.05, 24.55, 25.05, 25.55, 26.05. The solid line is $C(z)$, the weighted average of $C(i', z)$ over all values of i'. For $i' = 26.55$, $C(i', z)$ was estimated to be half the values of $C(i', z)$ at $i' = 26.05$.
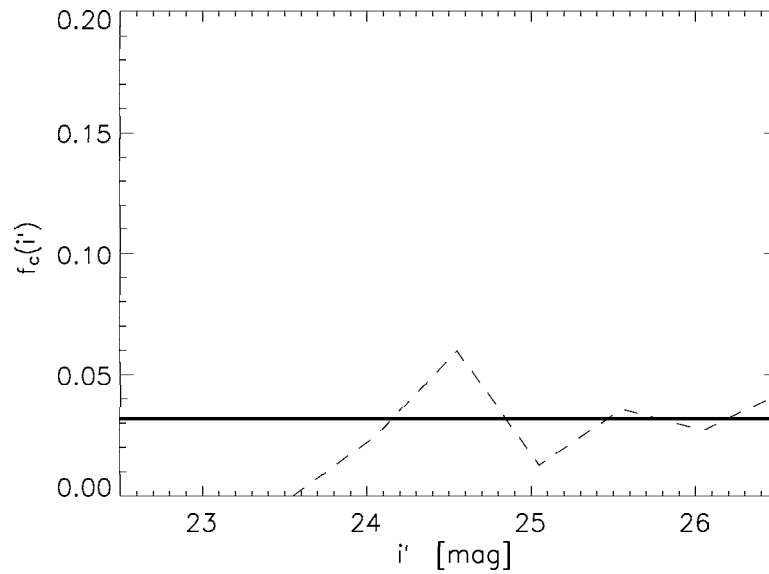
Figure 4.6: Contamination $f_c(i')$ as a function of i'-band magnitude (dashed line). The solid line is the weighted average of $f_c(i')$ ($= 3.2\%$) over all values of i'.

of the observed galaxies. When the galaxy contribution in the volume contains structure, a projected version of this structure is observed, polluted by foreground and background galaxies along the same line of sight. Figure 4.7 shows the angular distribution of the 874 LBG candidates within the FOV, including TN J1338–1942. It is difficult to recognise any structure, although some clustering seems to be present.

Our technique of making structure visible was to smooth out the mass contained within each LBG (which is basically a point mass from our point of view) over an area around the LBG (e.g. Steidel et al. 2000; Buttery 2003). As there is no detailed distance information available for the LBGs, we did not attempt to convert the luminostity of the LBGs into mass. Instead, the same constant (mean) mass is assumed for all LBGs. The surface density function in the FOV can be represented by a sum of (2-dimensional) Dirac delta functions at the positions of the LBGs:

$$\Sigma(\alpha, \delta) = \sum_{i=1}^{N} \delta(\alpha - \alpha_i, \delta - \delta_i) \tag{4.2}$$

Integrated over the whole FOV, this yields the total number (N) of LBGs in the FOV. To smooth the surface density, a (2-dimensional) gaussian convolution is applied on the objects within the FOV. It follows from convolution theory that a normalised gaussian is centered on the position of each delta function:

$$\widetilde{\Sigma}(\alpha, \delta) = \Sigma(\alpha, \delta) \otimes G(\alpha, \delta) = \sum_{i=1}^{N} G(\alpha - \alpha_i, \delta - \delta_i) \tag{4.3}$$

The normalised gaussian has the following functional form:

$$G(\alpha, \delta) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\alpha^2 + \delta^2}{2\sigma^2}\right) \tag{4.4}$$

It is assumed that α and δ are expressed in the same units (e.g. `arcsec`).

We choose the width (σ) of the gaussian in such a way that the gaussian is at half maximum when the distance from its center is equal to the average distance between objects in the FOV. This width has no relation to any physical scale. For N objects in a FOV with an area Ω this means:

$$\sigma = \sqrt{\frac{2\Omega}{\pi \ln(2) N}} \tag{4.5}$$

This way, objects blend when their angular separation is smaller than the average distance, thereby creating a sense of clustering. In our case, $N = 874$ and $\Omega = 596$ `arcmin`^{2}, so $\sigma = 47.5$ `arcsec`. After convolution, the map is normalised so that the surface density integrated over the total FOV gives 874, the total number of candidate LBGs in the FOV.

¿From a mathematical point of view, convolving a function with a gaussian of width σ in the spatial domain means multiplication with a gaussian of width $1/\sigma$ in the spatial frequency domain. A delta function in the spatial domain gives a constant amplitude for all spatial frequencies. The gaussian convolution acts as a low pass spatial frequency filter, allowing only for the spatial frequency components with wave-numbers $k \lesssim 1/\sigma$ to pass. This filters out the sharp discontinuity of the delta functions, resulting in more gradual transitions in the surface density function.

4.2.2   Correction Map

The convolution method in Section 4.2.1 succeeds in visualising overdense areas in the FOV, but becomes less accurate in areas where a relatively large part of the FOV is obscured, e.g. by

Figure 4.7: Spatial distribution of 874 LBG candidates in the FOV. The position of each candidate is marked by a circle whose si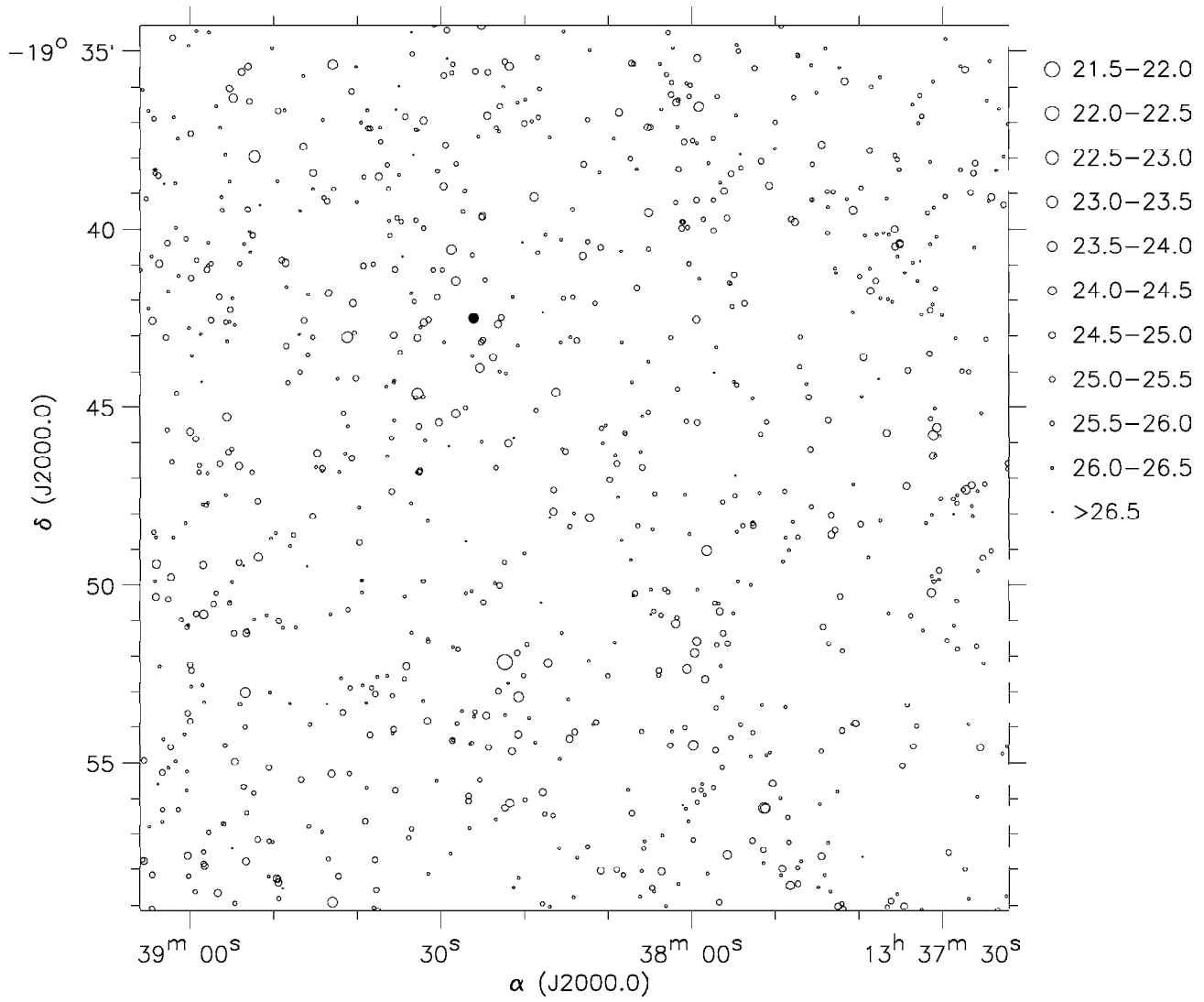ze indicates the i'-band magnitude as noted in the legend. The filled circle (at $13^h38^m26.1^s$, $-19°42'31''$) is TN J1338–1942.

foreground objects or by the limiting border of the image. Just as gaussian tails of unobscured objects add to the surface density in obscured areas (e.g. across the border), obscured objects should have contributed to the surface density in unobscured areas. The effect of this is most prominent along the border of the FOV, where the surface density drops outwards.

The drop in surface density along the border can be compensated, by using a reflecting boundary condition (the gaussian tail that runs over a border is reflecteded back into the FOV) or a periodic boundary condition (the gaussian tail that runs over the border enters the FOV at the opposite border). But these boundary conditions do not compensate for obscuration by foreground objects.

Instead, we used a large number of random objects to generate a correction map. Random objects were created by picking a random position in the FOV and verifying that there was no bright object present at that position. This verification was done by measuring the counts within a 2 `arcsec` circular aperture in all 3 bands and checking that the counts in each band lie within the range of -3 to $+5$ times the background noise $\sigma_{bg,m}$ (see Section 3.3.2). This way, random objects are allowed to be positioned in empty space, on top of faint objects or near bright objects, but never on top of bright objects.

In total, 100,000 random objects were generated, organised in 20 catalogs of 5,000 objects each. The reason for using 20 catalogs instead of 1 is that the same random objects were also used for the estimation of the angular correlation function (see Section 4.2.3). The correction map was created by applying the identical gaussian convolution of Section 4.2.1 on the 100,000 random objects and dividing the convolved map by its median value. Dividing the smoothed surface density map of Section 4.2.1 by the correction map gives a corrected surface density map, which is shown in Figure 4.8. There is no apparent correlation between the underdense areas in Figure 4.8 and the positions of bright, extended (obscuring) objects in Figure 3.1. It seems that overdense areas near the borders of the FOV are a bit overemphasized by the correction map.

To test the significance of the apparent clustering in Figure 4.8, similar surface density maps were created from 20 catalogs of 874 random objects each. The mean, median, minimum and maximum values of the surface density of the random maps were very close to the values of the object map. However, for the object map we found the RMS value of the fluctuations around the mean surface density to be 0.661 `arcmin`$^{-2}$, a factor ~ 1.4 times the combined result of 0.467 ± 0.028 `arcmin`$^{-2}$ (values ranged from 0.432 to 0.512 `arcmin`$^{-2}$) for the random maps ($\sim 7\sigma$ deviation). We compared the total area within the maps with a surface density higher than twice the mean. The object map gave 20.8 `arcmin`^{2}, a factor ~ 6.3 times the 3.3 ± 2.1 `arcmin`^{2} (values ranged from 1.10 to 7.24 `arcmin`$^{-2}$) for the random maps ($\sim 8\sigma$ deviation). We also compared the total area within the maps with a surface density lower than half the mean. The object map gave 46.9 `arcmin`^{2}, a factor ~ 2.5 times the 19.1 ± 3.7 `arcmin`^{2} (values ranged from 12.9 to 26.6 `arcmin`$^{-2}$) for the random maps ($\sim 7\sigma$ deviation).

4.2.3   Angular Correlation Function

A popular method of quantifying the projected clustering of objects is the 2-point angular correlation function (Peebles 1973, 1980), as it is relatively easy to calculate and can be related to the angular power spectrum by use of a Fourier transform. When the mean surface density of the projection of a large sample of galaxies is found to be $\bar{\Sigma}$, then the chance δP of finding a galaxy within a solid angle $\delta\Omega$ at a random angular position is given by:

$$\delta P = \bar{\Sigma} \, \delta\Omega \tag{4.6}$$

When galaxies are randomly distributed over space, the projected angular positions of galaxies are also randomly distributed (a poissonian distribution). Then the chance of finding a second galaxy within a solid angle $\delta\Omega$ at an angular distance θ from a randomly chosen first galaxy is given by Equation (4.6), as the chance of finding the second galaxy is independent on the choice of the first galaxy.

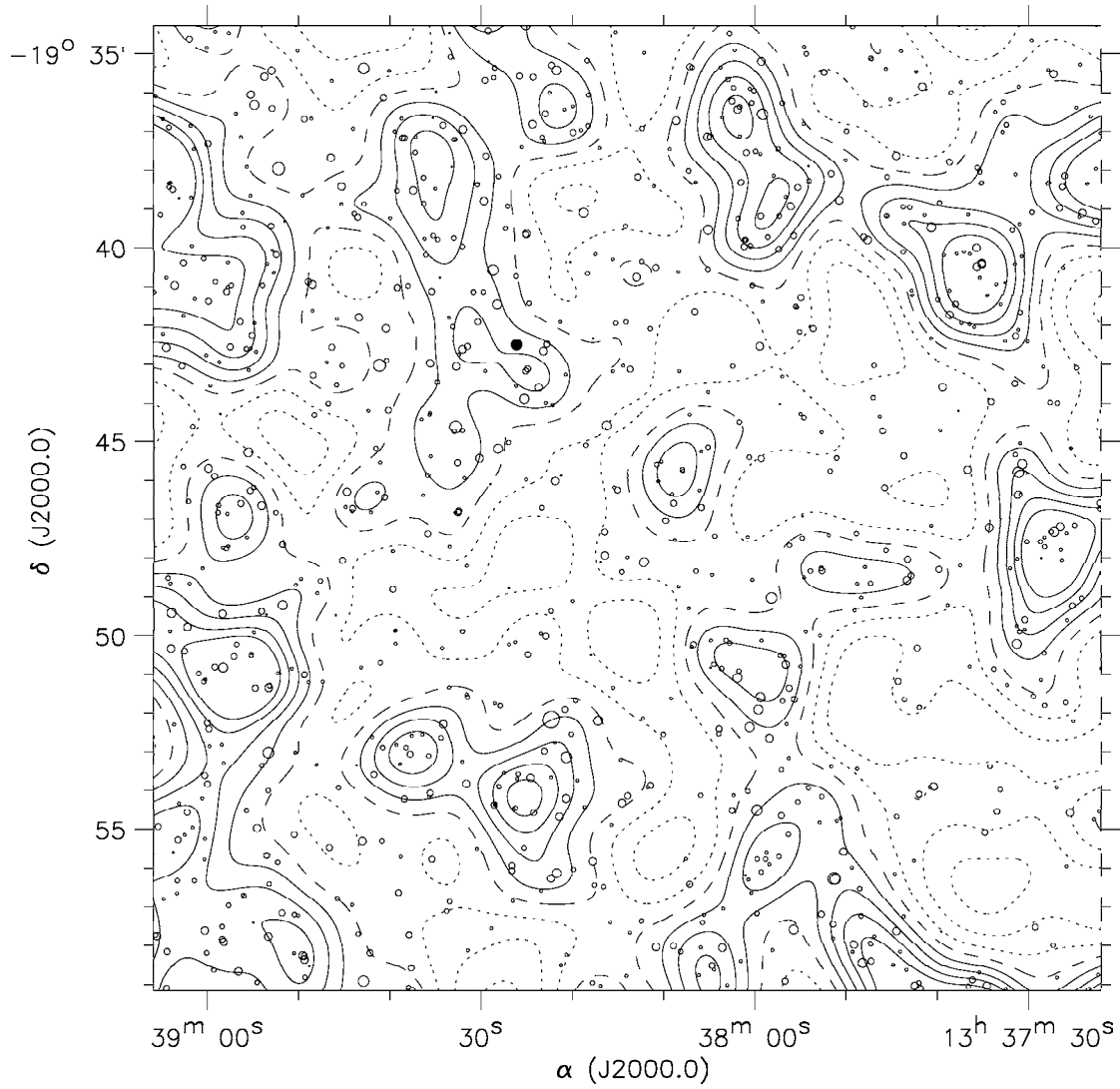Figure 4.8: Corrected surface density contour map of the candidate LBGs in the FOV. The dashed line represents the average surface density of 1.47 $\mathtt{arcmin}^{-2}$, the solid lines represent overdensities of +25, +50, +75 and +100% (edge to center) and the dotted lines represent underdensities of −25, −50 and −75% (edge to center). The LBG candidates are plotted in the same way as in Figure 4.7.

However, when galaxies tend to cluster, the distribution is no longer poissonian and the angular distances between galaxies will be correlated on smaller scales. In this case, the chance of finding a second galaxy within a solid angle $\delta\Omega$ at an angular distance θ from a randomly chosen first galaxy is:

$$\delta P = \bar{\Sigma} \left(1 + \omega(\theta)\right) \delta\Omega \tag{4.7}$$

This defines the angular correlation function (ACF) $\omega(\theta)$. The ACF is the Fourier transform of the angular power spectrum (Peebles 1980, Peacock 2001).

When galaxies are clustered, the ACF is expected to rise for smaller angular distances, while for large angular distances it will converge to zero. As the ACF does not depend on direction, it is most sensitive to gaussian-like density fluctuations, but less sensitive to elongated structures (e.g. a galaxy wall).

The ACF was calculated by using 2 different estimators by Landy&Szalay (1993) and Hamilton (1993) respectively:

$$\omega_{LS}(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}, \quad \omega_H(\theta) = \frac{DD(\theta) \times RR(\theta)}{DR^2(\theta)} - 1 \tag{4.8}$$

Here, $DD(\theta)$ is the binned number of observed galaxy-galaxy (DD) pairs with an angular distance between $\theta - \Delta\theta$ and $\theta + \Delta\theta$, divided by the total number of DD pairs ($= N_D(N_D - 1)/2$).

To take out any systematic selection effects (e.g. limited FOV, obscuration by foreground objects) in the galaxy catalog, the galaxy catalog is compared to large catalogs of randomly placed objects (see Section 4.2.2) that have the same positional constraints as the object catalog and produce a high density coverage of the FOV. In Equations (4.8), $RR(\theta)$ is the binned number of random pairs with an angular distance between $\theta - \Delta\theta$ and $\theta + \Delta\theta$, divided by the total number of random pairs ($= N_R(N_R - 1)/2$), while $DR(\theta)$ is the (binned) number of galaxy-random pairs with an angular distance between $\theta - \Delta\theta$ and $\theta + \Delta\theta$, divided by the total number of galaxy-random pairs ($= N_D \times N_R$).

To keep the number of pairs and calculations within reasonable limits, while maintaining a large total number of random objects, 20 catalogs were created, each containing 5,000 random objects. For each of the random catalogs, the ACF estimators $\omega_{LS}(\theta)$ and $\omega_H(\theta)$ were calculated using our LBG catalog, for 10 logarithmic bins ranging from $2 - 1,000$ `arcsec`. The 20 results per estimator, one for each random catalog, were combined into a mean and standard deviation, giving both the best estimate from 100,000 objects and the uncertainty due to variations in the random catalogs.

The formal uncertainty in the ACF estimators (assuming an infinite number of random objects) due to the limited sample of real galaxies is given by Hewitt (1982):

$$\sigma_\omega = \sqrt{\frac{|1 + \omega_{obs}(\theta)|}{DD(\theta) \times N_D(N_D - 1)/2}} \tag{4.9}$$

Here, $\omega_{obs}(\theta)$ is either $\omega_{LS}(\theta)$ or $\omega_H(\theta)$. This uncertainty was quadratically added to the random catalog uncertainty to obtain an estimate of the total uncertainty in the ACF estimators' data points.

Groth&Peebles (1977) found that ACF estimators are (negatively) offset by a integration constant, due to the assumption that the observed galaxy surface density is equal to the true surface density:

$$\omega(\theta) = \omega_{obs}(\theta) + IC, \quad IC = \frac{1}{\Omega^2} \int_{FOV} \int_{FOV} \omega(\theta_{12}) \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 \tag{4.10}$$

This integration constant cannot be evaluated without knowing the true ACF.

For low redshift galaxies, it is found that the ACF closely resembles a power-law over a fairly large range of θ (e.g. 10 `arcsec` $\lesssim \theta \lesssim$ 3 `deg` in Peebles 1980):

$$\omega(\theta) = A_\omega \theta^{\ \beta} \tag{4.11}$$

The value of β is found to be close to 0.8 (Groth&Peebles 1977; Peebles 1980). We assumed a power-law form for our ACF estimations and attempted a fit (using IDL→curvefit, a weighted, non-linear least squares fit using a chi square test), including the IC:

$$\omega_{obs}(\theta) = A_\omega \theta^{-\beta} - IC \tag{4.12}$$

The results of the fitting parameters (A_ω, β, IC), including uncertainties, for both estimators are presented in Table 4.2. When allowing all 3 parameters to vary, the uncertainties in all three parameters are relatively large. The fitting was repeated for a fixed value of $\beta = 0.8$. Both estimators give similar results, but the fitted value A_ω for ω_H is in both cases a bit higher than for ω_{LS}. For varying β, the clustering amplitude A_ω is more than 3 times larger than its uncertainty ($> 3\sigma$ significance), while for fixed β, the clustering amplitude A_ω is more than 7 times larger than its uncertainty ($> 7\sigma$ significance). The fitted power-laws for the Landy&Szalay estimator are shown in Figure 4.9. The plots for the Hamilton estimator are very similar.

Due to contamination (see Section 4.1.2), the apparent clustering of LBG candidates is diluted (Ouchi et al. 2004b). While the contaminant objects might possess some clustering as well, it is likely that this clustering signal is much weaker than the clustering signal of the true LBG population. When making this assumption, the clustering amplitude of the LBG ACF can be corrected for contamination:

$$A_{\omega,c} = \frac{A_\omega}{(1 - f_c)^2} \tag{4.13}$$

Here, f_c is the contamination fraction. Using the weighted average of $f_c(i')$ of 0.031, the corrected ACF amplitudes are given in column 5 of Table 4.2.

4.2.4 Spatial Correlation Function

Similar to the definition of the angular correlation function (ACF; see Section 4.2.3), a spatial correlation function (SCF) can be defined, which is the chance above average of finding a second galaxy within a volume δV at a distance r from a randomly chosen first galaxy:

$$\delta P = \bar{n} \left(1 + \xi(r)\right) \delta V \tag{4.14}$$

For low redshift galaxies, the SCF function is well approximated by a power-law (e.g. Peebles 1980):

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma} \tag{4.15}$$

Estimator	A_ω	β	IC	$A_{\omega,c}$
ω_{LS}	4.96 ± 1.39	1.11 ± 0.11	0.0069 ± 0.0031	5.29 ± 1.48
ω_H	5.37 ± 1.44	1.13 ± 0.10	0.0067 ± 0.0030	5.73 ± 1.54
ω_{LS}	2.00 ± 0.27	0.8 (fixed)	0.0159 ± 0.0029	2.13 ± 0.28
ω_H	2.04 ± 0.26	0.8 (fixed)	0.0163 ± 0.0029	2.17 ± 0.28

Table 4.2: Overview of the fitting parameters obtained by fitting a power-law (see Equation (4.12)) to the ACF estimators by Landy&Szalay (1993) and Hamilton (1993).
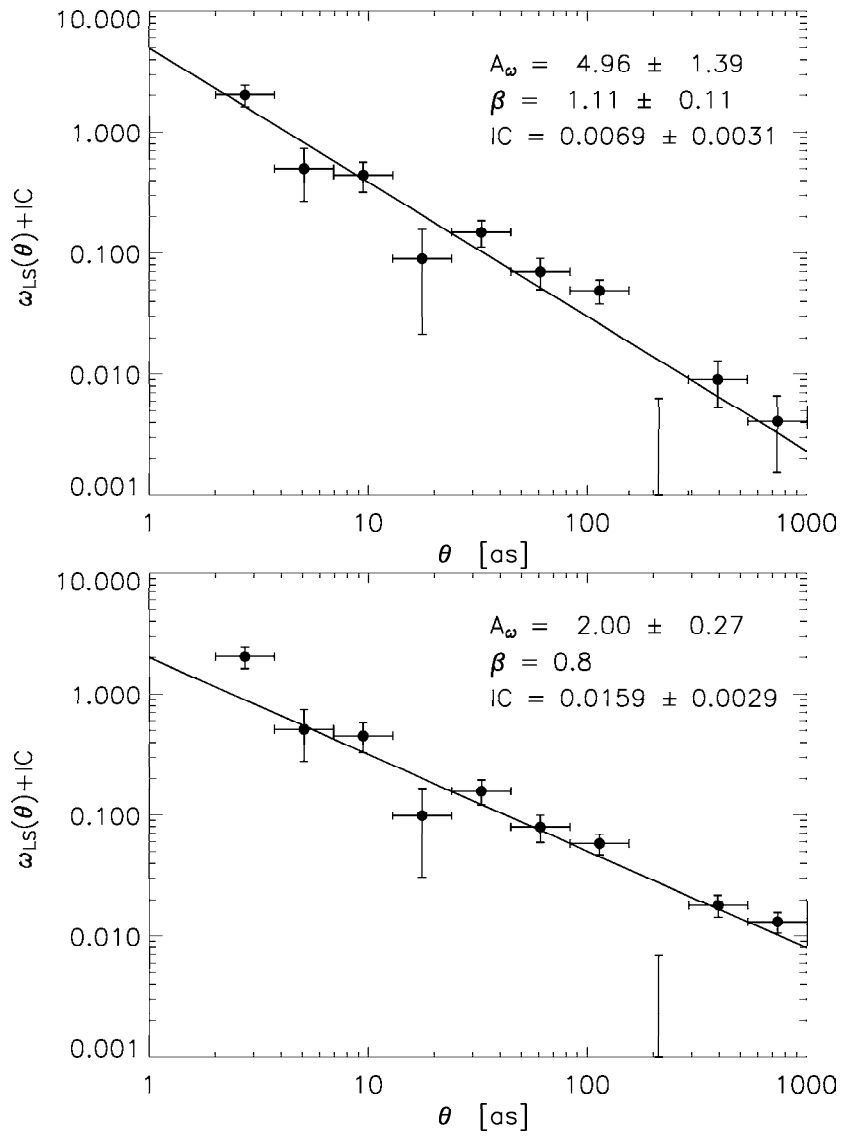
Figure 4.9: Overview of the Landy&Szalay (1993) ACF estimator datapoints, with both a variable slope power-law fit (upper) and a fixed slope power-law fit (lower).

The correlation length r_0 is the radius within which δP is more than twice the average chance.

The ACF is related to the SCF by the projection of the galaxy sample on the plane of the sky. The transformation from $\xi(r)$ to $\omega(\theta)$ was first described by Limber (1953), while Groth&Peebles (1977) formulated the relativistic version of the so-called Limber equation, including spatial curvature. For small angles ($\theta \ll 1$ `rad`) and a power-law SCF (Equation (4.15)), the relativistic Limber equation is given by Efstathiou et al. (1991):

$$\omega(\theta) = \frac{\sqrt{\pi}\,\Gamma\left(\frac{\gamma-1}{2}\right)}{\Gamma\left(\frac{\gamma}{2}\right)}\,A\,r_0{}^\gamma\theta^{1-\gamma} \tag{4.16}$$

The constant A is evaluated by the following integral:

$$A = \frac{\int_0^\infty D_A{}^{1-\gamma}(z)(\mathrm{d}D_P(z)/\mathrm{d}z)^{-1}N^2(z)F(z)\mathrm{d}z}{\left[\int_0^\infty N(z)\,\mathrm{d}z\right]^2} \tag{4.17}$$

Here, $D_A(z)$ is the proper angular diameter distance and $D_P(z)$ the proper line-of-sight (LOS) distance (e.g. Hogg 2000). $N(z)$ is the redshift distribution of objects and $F(z)$ describes the change of the SCF as a function of redshift:

$$\xi(r,z) = \xi(r)F(z) \tag{4.18}$$

These formulae can be used to invert the Limber equation, i.e. to derive the SCF from the ACF. From Equations (4.12) and (4.16), it follows that:

$$\gamma = \beta + 1, \quad r_0 = \left(\frac{\Gamma\left(\frac{\gamma}{2}\right)A_\omega}{\sqrt{\pi}\,\Gamma\left(\frac{\gamma-1}{2}\right)A}\right)^{\frac{1}{\gamma}} \tag{4.19}$$

We calculated the parameters γ and r_0, using the equations above and the values for β and $A_{\omega,c}$ found in Section 4.2.3. For $N(z)$, we used the weighted completeness function $C(z)$ (see Figure 4.5). Following Efstathiou et al. (1991) and Ouchi et al. (2004b), we used the following functional form for $F(z)$:

$$F(z) = \left(\frac{1+z}{1+z_c}\right)^{-(3+\epsilon)} \tag{4.20}$$

Here z_c is the central redshift of the observed sample of galaxies, which is $z = 4$ in our case. For $\epsilon = \gamma - 3$, the clustering pattern is fixed in comoving coordinates, while for $\epsilon = 0$ the clustering pattern is fixed in proper coordinates. Choosing $\gamma - 3 \le \epsilon \le 0$ allows for evaluating different clustering evolution scenarios. It turned out that varying γ had very little effect on the results ($\Delta r_0 \sim 0.01$ `Mpc`). The results are presented in Table 4.3, with r_0 expressed in comoving distance.

Estimator	$A_{\omega,c}$	β	$r_0\ [h^{-1}\ \mathtt{Mpc}]$	γ
ω_{LS}	5.29 ± 1.48	1.11 ± 0.11	$3.63^{+1.55}_{-1.10}$	2.11 ± 0.11
ω_H	5.73 ± 1.54	1.13 ± 0.10	$3.62^{+1.46}_{-1.06}$	2.13 ± 0.10
ω_{LS}	2.13 ± 0.28	0.8 (fixed)	$4.56^{+0.33}_{-0.35}$	1.8 (fixed)
ω_H	2.17 ± 0.28	0.8 (fixed)	$4.61^{+0.32}_{-0.35}$	1.8 (fixed)

Table 4.3: Overview of the SCF power-law parameters (4.15), derived from the power-law fits (4.12) on the ACF estimators by Landy&Szalay (1993) and Hamilton (1993).

Chapter 5

Discussion

5.1 Results

5.1.1 Photometry

To test the photometry in i', we used the catalog of candidate LBG galaxies by Miley et al. (2004) and Overzier et al. (in preparation). This catalog contains 71 LBG candidates in an area of 11.6 $\mathtt{arcmin^2}$ around TN J1338–1942 (red rectangle in Figure 5.2) which were extracted from $g'r'i'z'$-images (Fukugita et al. 1996) by the HST/ACS (Pavlowski et al. 2004). The ACS F775W-filter ($= i'$) is a close match to the Suprime-Cam i'. We found 62 of the 71 objects (within 1 $\mathtt{arcsec}$ search radius) in our raw catalog of 120,000 objects, of which 56 have an (Suprime-Cam) i'-band magnitude less than 26.55. ACS F775W magnitudes were measured using Sextractor's MAG_AUTO function (see Section 3.3.3) and have a photometric uncertainty less than 0.18 magnitude.

Figure 5.1 shows a trumpet diagram of F775W-magnitudes compared to i'-magnitudes for objects with $i' < 26.55$. If we assume that F775W-magnitudes are without photometric uncertainty, we find that F775W$-i'$ differences average to ≈ 0 and most F775W$-i'$ differences lie within 3 times the photometric uncertainty in i', except for a few outliers. The upper outliers are the result of incorrect identifications due to objects whose angular separation is very small ($\lesssim 1$ $\mathtt{arcsec}$). The single outlier below is due to count loss outside the 2 $\mathtt{arcsec}$ aperture of i' (see Section 3.3.3).

We conclude that the i'-band photometric calibration and uncertainty estimation are fairly accurate. We believe this also holds for the B- and R_C-band magnitudes, because the process towards photometric calibration of these bands was very similar to the i'-band. Figure 5.1 shows that the 3σ photometric uncertainties in i' ranges up to ~ 1 magnitude for $i' \rightarrow 26.55$. This is similar for $R_C \rightarrow 26.8$ and $B \rightarrow 27.0$. The photometric uncertainty causes a larger scatter in the color of faint LBG candidates, with typical magnitudes $i' \sim 26.3$, $R_C \sim 26.3$ and $B \sim 28.3$. About 40% of the number of LBG candidates in our catalog have $i' > 26.0$ and qualify as faint.

5.1.2 Lyman Break Galaxies

A comparison is made between our LBG candidates and LBG candidates from a much smaller, but deeper, field around TN J1338–1942 by Miley et al. (2004) and Overzier et al. (in preparation). As stated in Section 5.1.1, they found 71 candidate LBG galaxies in an area of 11.6 $\mathtt{arcmin^2}$ around TN J1338–1942 using colors $g' - r'$ and $r' - i'$ (composed of HST/ACS wideband filter magnitudes), thereby selecting galaxies from a redshift range similar to ours. The FOV (red rectangle), including the positions of the LBG candidates (red 'plus' signs), are plotted in the Suprime-Cam's FOV in Figure 5.2. Of these 71 candidates, 62 were detected in our raw catalog
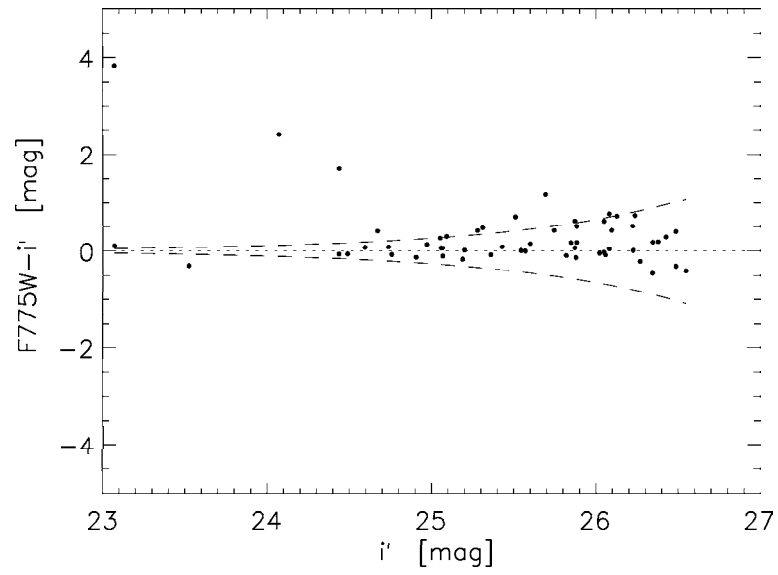
Figure 5.1: Magnitude difference between ACS F775W ($= i'$) and Suprime-Cam i'. The dashed lines represent plus and minus 3 times the total photometric uncertainty in Suprime-Cam i'.

of $\sim 120,000$ objects, of which 56 objects have an i'-band magnitude below $3\sigma_{bg,i'} = 26.55$. Of these 56 objects, 13 were detected as LBG candidates in our catalog, which makes a strong case for these candidates. Within the same FOV, our catalog contains 24 LBG candidates in total, $\sim 34\%$ of the number of ACS LBG candidates.

As Miley et al. (2004) and Overzier et al. (in preparation) use a different filter set with different selection criteria, it is hard to directly compare the two sets of LBG candidates. The selection criteria define which redshift interval is probed. The completeness and contamination are also dependant on the limiting magnitude and the PSF FWHM. As the HST is a space based telescope, diffraction and extinction in Earth's atmosphere is eliminated. The PSF FWHM of the ACS images is nearly telescope diffraction limited ($\lesssim 0.1$ `arcsec`), ~ 10 times better than our Suprime-Cam images, so it can separate sources and do photometry with much better accuracy.

When making the (crude) assumption that the ACS catalog of LBGs is 100% complete and 0% contaminated, it follows that our catalog of LBG candidates is $\sim 23\%$ complete within the $i' < 3\sigma_{bg,i'} = 26.55$ detection limit, while the contamination is $\sim 46\%$. While the completeness agrees within uncertainty limits with the completeness found in Section 4.1.2, the contamination seriously disagrees. It is more likely that both LBG samples are subsets from a larger collection of high redshift galaxies, and that both subsets only partially overlap. This makes a strong case for the 13 LBG candidates that are present in both catalogs, so the $\sim 23\%$ completeness might be interpreted as a lower limit estimate.

Candidate LBG counts in the field of TN J1338–1942 were compared with the candidate LBG counts of both the SDF and SXDF fields in Ouchi et al. (2004a). Ouchi et al. found 1438 LBG candidates with $i' < 26.3$ in the SDF field and 732 with $i' < 25.8$ in the SXDF field. These fields are blank fields, i.e. there was no a-priori knowledge on the presence of clusters. We found 874 LBG candidates with $i' < 26.55$ in the TN J1338–1942 field. As most of the objects are detected at the faint end of i' and the FOV of the fields are comparable, we detect far less LBG candidates per magnitude range than Ouchi et al. ($\sim 50\%$ for $i' > 25$). Comparing 3σ noise levels (Table 5.1), we see that the B-band images of both the SDF and SXDF fields are deeper than the TN J1338–1942 field. As most of the non-LBG candidates are rejected on the basis of their low $B - R_C$ color (break), a B-band magnitude difference of $0.2 - 0.4$ can make a big difference. Furthermore, Ouchi et al. do not mention rejection of false detections from their catalogs, which reduced our LBG candidate count by $\sim 5\%$. There is also the possibility of a low density due to cosmic variance, but this is found to be only $\sim 10 - 20\%$ for LBGs (Somerville et al. 2004), while the LBG catalog of Miley et al. (2004) and Overzier et al. (in preparation) suggests that the LBG density is higher than we observed.

5.1.3 Lyman-α Emitters

Using VLT wideband (R_C) and narrowband (custom) imaging and follow-up spectroscopy in the field of TN J1338–1942, Venemans et al. (2002 + extension) found 38 spectroscopically confirmed (including TN J1338–1942) and 48 narrowband selected candidate LAEs at $z \approx 4.1$. The total area of 79.7 `arcmin`^{2} was covered with 2 partially overlapping images (blue rectangles in Figure 5.2). There are 104 LBG candidates from our catalog that lie within this area. A comparison

	SDF	SXDF	TN J1338–1942
B	27.6	27.4	27.23
R_C	26.9	27.0	26.87
i'	26.7	26.0	26.55

Table 5.1: Comparison of 3σ limiting magnitudes between the SDF and SXDF fields and the TN J1338–1942 field.

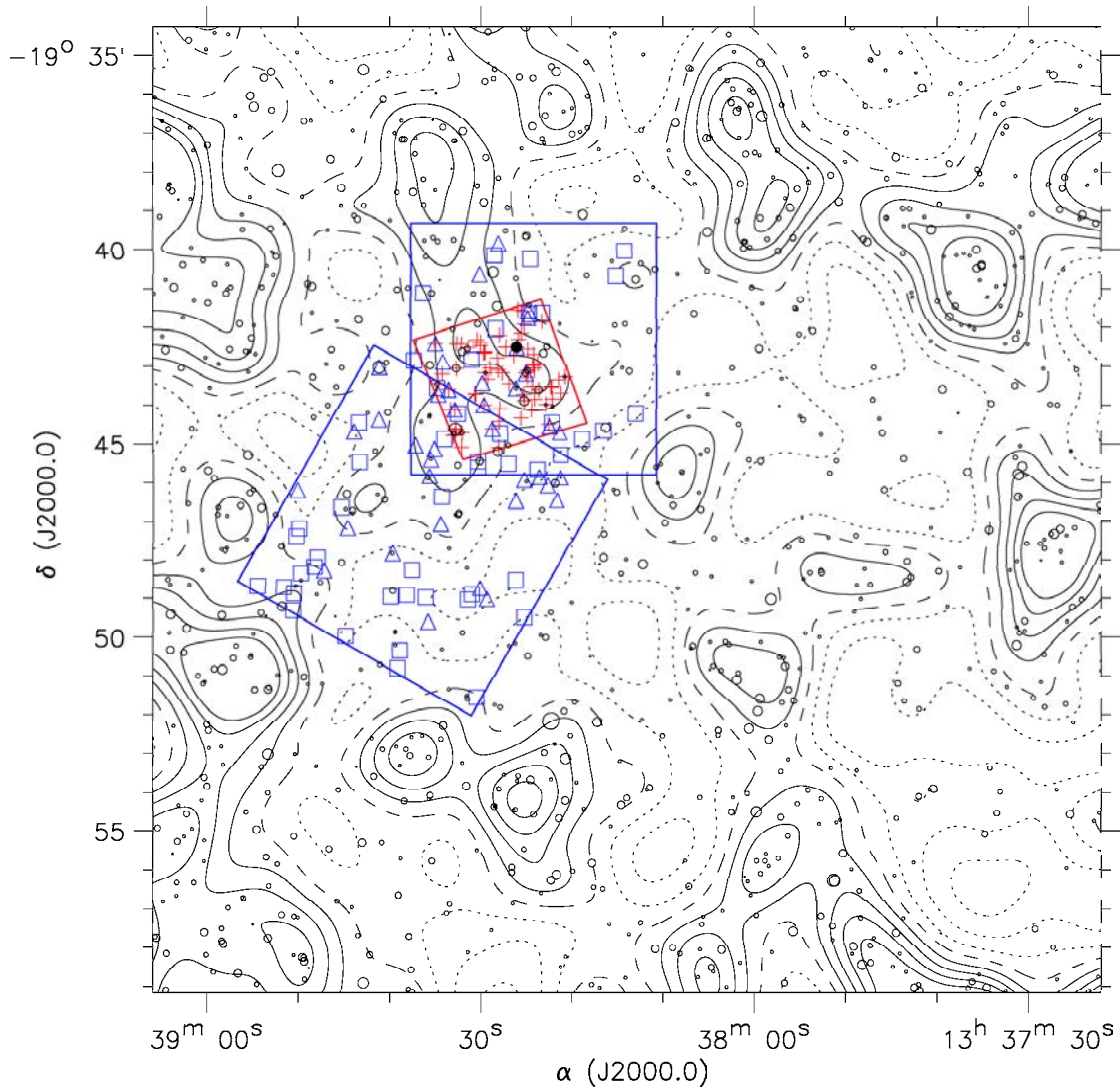Figure 5.2: Surface density contour map of the candidate LBGs in the FOV, as in Figure 4.8. The blue rectangles represent the FOV for searching LAEs at $z \approx 4.1$ (blue triangles for confirmed, blue squares for candidate) by Venemans et al. (2002 + extension), while the red rectangle represents the FOV for searching LBGs (red 'plus' signs) with the ACS by Miley et al. (2004) and Overzier et al. (in preparation) .

was made between our LBG catalog and the LAE catalog. Of the 38/48 (confirmed/candidate) LAEs, 30/25 were detected in our raw catalog of $\sim 120,000$ objects, of which 20/10 objects have an i'-band magnitude below 26.55. Of these 20/10 objects, 7/0 were detected as LBG candidates.

LAEs are selected by their strong Lyman-α emission line, so this does not mean the galaxy has a spectrum that is suited for LBG detection. However, because of the narrow redshift range in which they are detected, they are interesting reference material. When considering only the confirmed LAEs, we find that 7 out of 20 (35%) detectable galaxies at $z \approx 4.1$ are selected by our selection criteria. This agrees with the expected completeness (see Figure 4.5), when assuming that LAEs are a subset of LBGs. When including the candidate LAEs, this drops to 7 out of 30 ($\sim 23\%$), but within uncertainty margins of poissonian statistics this still agrees. There is surely some selection effect present, as the confirmed LAEs are, in general, brighter than the candidate LAEs (relatively easier to obtain spectra from brighter objects).

5.1.4 Clustering

In Section 4.2.2, we found that the smoothed surface density map (Figure 4.8) deviates from random fluctuations. We assumed that the process behind this deviation involves the gravitational attraction between galaxies that leads to the formation of galaxy clusters. Smoothing with a gaussian of $\sigma = 47.5$ `arcsec` means that all structural detail $\lesssim 2 \times 47.5$ `arcsec` ≈ 1.5 `arcmin` is smoothed out. At redshift $z = 4$ this equals ~ 3 `Mpc`. Knowing that low redshift galaxies and clusters have sizes of $\lesssim 100$ `kpc` and $\gtrsim 10$ `Mpc`, respectively, the smoothing operates right in between these size scales.

Already at 25% above the average surface density, the density contours start forming separate isles at considerable distance from each other while the density contours at 25% below average trace out at larger, connected areas. This agrees with the idea of clusters of galaxies surrounded by voids of relatively empty space ('the field'). Similar maps of random catalogs with the same number of entries (874) showed that the overdensity peaks are larger in area and lower in amplitude while the underdensity areas look more like the complement of the overdensity areas, i.e. disconnected. When assuming that the 25% overdensity contours trace out (proto-)cluster boundaries, it follows that the typical angular diameter of a cluster is $\sim 5 - 6$ `arcmin`. At $z \approx 4$ this means a typical transverse size of ~ 10 `Mpc`.

The question is: what is the likelihood that the overdense areas in Figure 4.8 are actual protoclusters instead of random fluctuations? Clustering 'signals' must be coming from a relatively narrow redshift interval similar to the transverse size of the cluster, while candidate LBGs outside this narrow interval and contaminants are providing the 'noise'.

When assuming that the line-of-sight (LOS) depth of a cluster is also ~ 10 `Mpc`, then at $z \approx 4$ this means a redshift interval $\Delta z \approx 0.015$. The redshift range $z = 3.5 - 4.5$ probes a LOS depth of ~ 700 `Mpc`, that is ~ 70 times the cluster depth. Although the completeness is highest around $z = 4$ and decreases outwards (see Figure 4.5), there is still an 'equivalent depth' that is $\sim 30 - 40$ times the cluster depth. Unless a clustering signal is really strong, it could easily be swamped by galaxies outside the narrow redshift range along the same line of sight.

Evidence was presented by Venemans et al. (2002) for a proto-cluster around TN J1338–1942. They selected LAEs from a narrow redshift range $z = 4.07 - 4.12$, which probes a LOS depth of ~ 33 `Mpc`. Within their FOV of 79.7 `arcmin`^{2} there are 104 LBG candidates present in our catalog. The mean surface density of LBG candidates within this FOV is 1.30 `arcmin`$^{-2}$, which is lower than the mean surface density for the total catalog of LBG candidates. Of the 104 LBG candidates within their FOV, 7 match with confirmed LAEs, which sets a lower limit on the actual number of LBG candidates within the narrow redshift range. Averaging 104 LBGs over an 'equivalent depth' of ~ 350 `Mpc` gives an estimate of ~ 10 LBGs within the narrow redshift range $z = 4.07 - 4.12$. This means that the identification of the 7 LBGs by itself does not signal an overdensity at $z \approx 4.1$. However, Steidel et al. (2000) found that only $20 - 25\%$ of the LBGs at $z = 3$ have a strong Lyman-α emission line to satisfy typical LAE selection criteria. This implies

that the expected number of LBGs at $z = 4.07 - 4.12$ is $\sim 30 - 35$. Assuming that all of these LBGs are in our selection of 104 LBG candidates, this would imply an overdensity of LBGs of a factor $3 - 5$ within the redshift range $z = 4.07 - 4.12$.

Figure 5.2 shows the angular distribution of the confirmed and candidate LAEs (in blue), overplotted on the surface density map of LBG candidates. The spatial distribution of the LAEs partially coincides with the overdense areas. There is a reasonable agreement in the overdense area directly surrounding TN J1338–1942, while within the southern half of their FOV the agreement is worse. The same Figure also contains the angular distribution of 71 LBG candidates (in red) found by Miley et al. (2004) and Overzier et al. (in preparation). Within this limited FOV, the LBG distribution agrees well with the surface density map, especially in the area marked by the 50% overdensity contour, just next to TN J1338–1942.

5.1.5 Correlation

One way of testing whether the observed surface density fluctuations are random or not is by determination of the two-point angular correlation function (ACF). In the case of random fluctuations, the ACF is approximately zero. We found that the ACF is well fitted with a power-law (Figure 4.9) for 2 estimators (Landy&Szalay 1993; Hamilton 1993). The clustering amplitude is measured with $> 7\sigma$ significance, which basically rules out the possibility of random fluctuations.

Although the power-law trend in the data points is easily recognised, there is one data point on the interval $\theta \approx 150 - 300$ `arcsec` $= 2.5 - 5$ `arcmin` that lies significantly below the fitted power-law. When looking at the surface density map (Figure 4.8), the average angular distance between the cores of overdense areas and the surrounding underdense areas lies in the approximate range of $2.5 - 5$ `arcmin`. So there is a significantly smaller chance of finding a second galaxy at a distance $\theta = 2.5 - 5$ `arcmin` when the first galaxy lies within an overdense core. As most of the LBG candidates do lie in overdense areas (by definition), it is likely that this left a signal in the ACF. Reversing this argument, it might very well be possible to extract the typical angular size of overdense areas by searching for a drop in the ACF. In our case, the typical overdense area size estimate would be $\sim 2 \times 2.5 = 5$ `arcmin` ($= 10$ `Mpc` at $z = 4$).

The two-point spatial correlation (SCF) of the candidate LBGs, calculated from the power-law fitting parameters of the ACF, shows a spatial correlation length of $r_0 = 3.6 \pm 1.5 \, h^{-1}$ `Mpc` for an SCF power-law slope of -2.1 ± 0.1 and $r_0 = 4.6 \pm 0.3 \, h^{-1}$ `Mpc` for a fixed slope of -1.8. These values correspond to an angular separation of $150 - 190$ `arcsec`. This is well within the domain of the fitted ACF so that no extrapolation is required. Both values of r_0 do agree well (within uncertainty margins) with the value of $r_0 = 4.1 \pm 0.2 \, h^{-1}$ `Mpc` found by Ouchi et al. (2004b) for LBGs at $z \approx 4$. Ouchi et al. conclude that r_0 stays constant at $\sim 5 \, h^{-1}$ `Mpc` or increases slightly with increasing redshift over the redshift range $z = 3 - 5$, which suggests that during the era of $z = 5$ to 3, the clustering of LBGs in the evolving universe does not become stronger. Overzier et al. (2003) made a overview diagram of correlation lengths versus redshift for both models and various galaxy types, including LBGs. This diagram suggests a moderate increment of r_0 towards lower redshift, but this does not conflict with Ouchi et al. (within uncertainty margins).

5.1.6 Catalog Divisions

The set of 874 LBG candidates was divided into a bright and faint subset of 437 LBG candidates each, based on their i'-band magnitude. The division boundary was at $i' = 25.86$, which is $5.8 \, \sigma_{bg,i'}$. Due to smaller photometric uncertainties, the bright subset has smaller uncertainties in its colors than the faint subset. Therefore, the contamination of the bright subset is expected to be lower compared to the faint subset, while the completeness is expected to be higher. For each of the subsets a smoothed surface density map was created using the same method as decribed in Sections 4.2.1 and 4.2.2. Figure 5.3 shows the result for the bright subset, while Figure 5.4 shows the result for the faint subset.

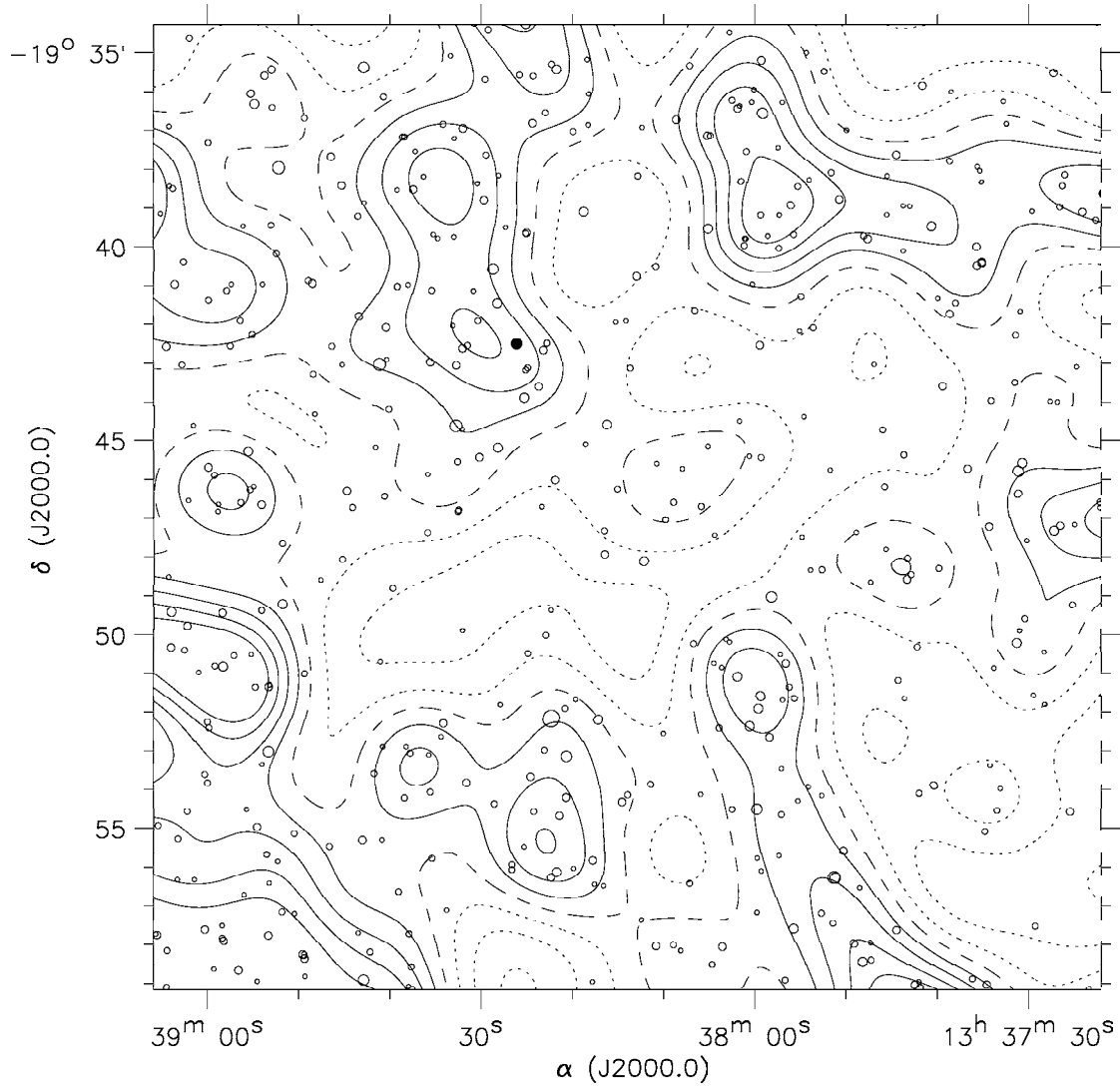Figure 5.3: Surface density contour map of 437 bright candidate LBGs ($i' < 25.86$) in the FOV, as in Figure 4.8.
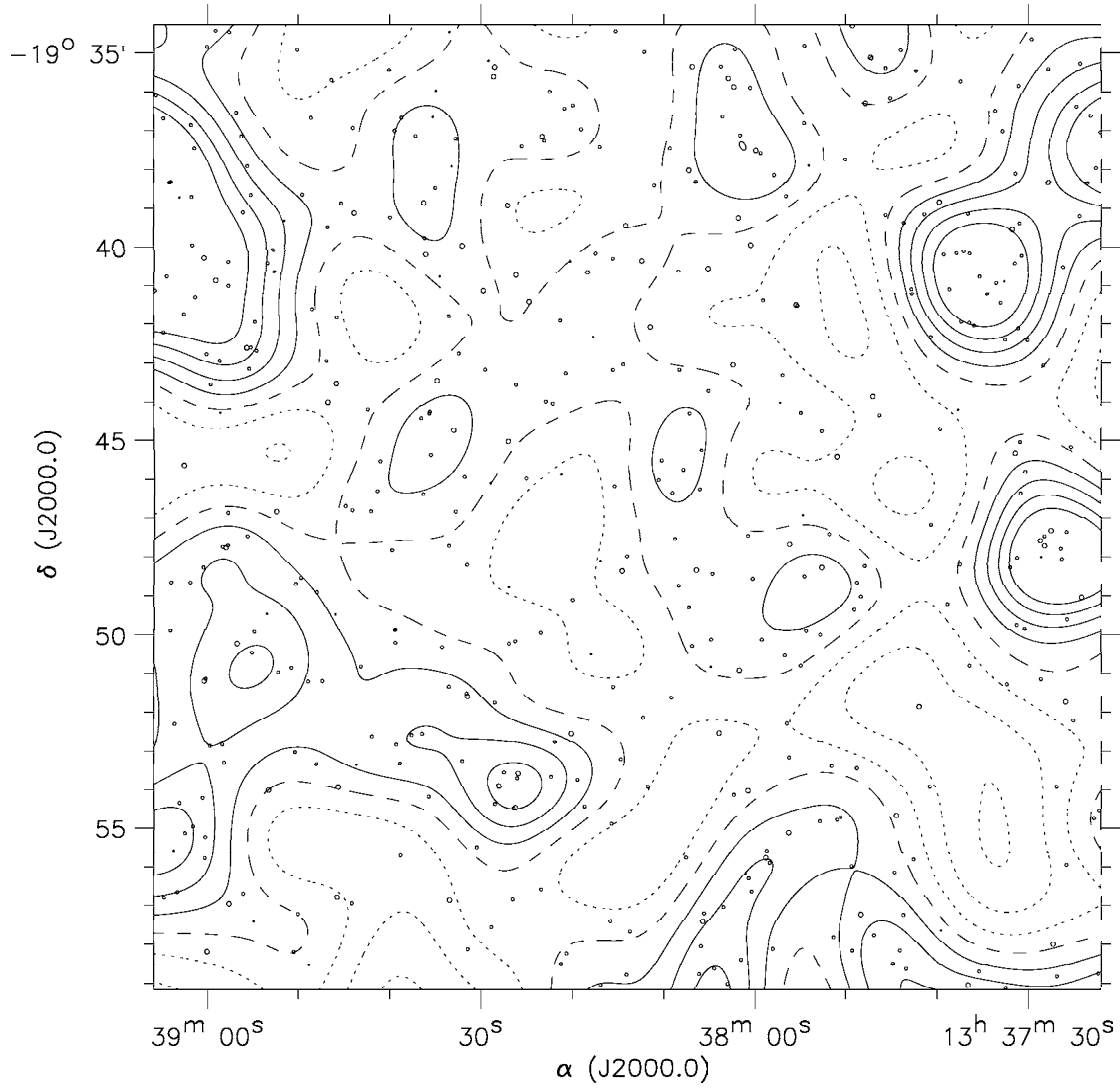
Figure 5.4: Surface density contour map of 437 faint candidate LBGs ($i' > 25.86$) in the FOV, as in Figure 4.8.

It is striking how different these maps are compared to each other. The overdense regions in the bright map are more numerous and stronger peaked, therefore creating a larger contrast between the overdense and underdense regions. Several overdense areas in the bright map are not or barely visible in the faint map. These maps suggest that the clustering of bright objects is stronger than of faint objects, when assuming that the higher contamination in the faint subset does not dominate the clustering pattern in that map. Comparing both maps against the total map of all 874 LBG candidates shows that most of the overdense areas in the total map originate from the bright map, including the overdensity surrounding TN J1338–1942.

Again, the set of 874 LBG candidates was divided into 2 equally sized subsets of 437 LBG candidates each, based on the strength of the Lyman break or the height of the $B - R_C$ color. From Figure 4.1 it follows that a higher $B - R_C$ color means a higher redshift of the LBG candidate. For fainter LBG candidates, the larger photometric uncertainty and the upper limit on the B-magnitude of $1\sigma_{bg,B}$ tends to lower $B - R_C$, thereby influencing the relation between $B - R_C$ and redshift. The subset of 437 high-break LBG candidates with $B - R_C > 2.35$ is expected to select both brighter LBG candidates and LBG candidates within the higher part of the redshift interval $z = 3.5 - 4.5$, while the complement is a subset of 437 low-break LBG candidates with $B - R_C < 2.35$. The smoothed surface density map of the high-break subset is shown in Figure 5.5, while Figure 5.6 shows the low-break map.

Comparing the high-break and low-break maps reveals that there are very few bright LBG candidates in the low-break map. Although there are some differences, these maps look very similar to the bright and faint maps. TN J1338–1942 occupies an overdense area in the high-break map, while this area is not present in the low-break map. This is also true for several other overdense areas. The overdense areas that are present in both the bright map and the high-break map might be labelled as high confidence areas, because their members are bright LBG candidates (with small photometric uncertainties) and a large Lyman break.

5.1.7 Cluster Galaxy Density and Proto-Cluster Density

In this section, we make some bold assumptions and present some tentative consequences that may arise from our data. First, let's assume that the overdense area around TN J1338-1942 in Figure 4.8 is a proto-cluster with a depth similar to its width, not contaminated by foreground or background proto-clusters. In Section 5.1.4 we found that the typical width at $z \sim 4$ was ~ 10 Mpc. When assuming this is also the cluster's depth, we found that the 'equivalent depth' of the probed volume was $\sim 30 - 40$ times the cluster depth, which is $300 - 400$ Mpc. Therefore, the observed overdensity is enhanced by a factor of $30 - 40$ when only considering the volume of the cluster depth. An observed $25 - 50\%$ overdensity implies a true overdensity of $750 - 1500\%$, meaning a cluster LBG density of $8.5 - 16$ times the average LBG density. In Section 4.2.1 we found an average LBG density of $\sim 6 \times 10^{-5}$ Mpc^{-3} for the full depth of our probed volume, which should be doubled when considering the 'equivalent depth'. Furthermore, in Section 4.1.2 we found that the completeness at $z \sim 4$ is $\sim 40\%$. Together, this implies a cluster galaxy density of $\sim 4 \times 10^{-3}$ Mpc^{-3}. This estimate does not include galaxies which are undetectable by the Lyman break technique, e.g. galaxies without star formation or dust obscured star forming galaxies. The value found is considerably lower than the average galaxy density of ~ 1 Mpc^{-3} in the local universe (e.g. Giuricin et al. 1993), while the contrast with the galaxy density inside rich clusters will be even larger. Assuming that galaxy mergers decreased the number density of galaxies towards lower redshift increases this contrast even more. This implies that the detected LBGs are only a very small fraction ($< 4 \times 10^{-3}$) of the total number of galaxies at redshift $z \sim 4$, unless a large number of galaxies formed between $z \sim 4$ and $z \sim 0$.

Second, let's assume that all the overdensities in the bright map of Figure 5.3 are actual proto-clusters. As stated in Section 4.2.1, the search for LBGs probed a comoving volume of $\sim 50 \times 50 \times 700 = 1.8 \times 10^6$ Mpc^3, equal to $6.0 \times 10^5 \, h^{-3}$ Mpc^3 for $h = 0.7$. There are ~ 5 overdensities in the bright map, which means an average proto-cluster density of $\sim 8 \times 10^{-6} \, h^3$ Mpc^{-3}. Including
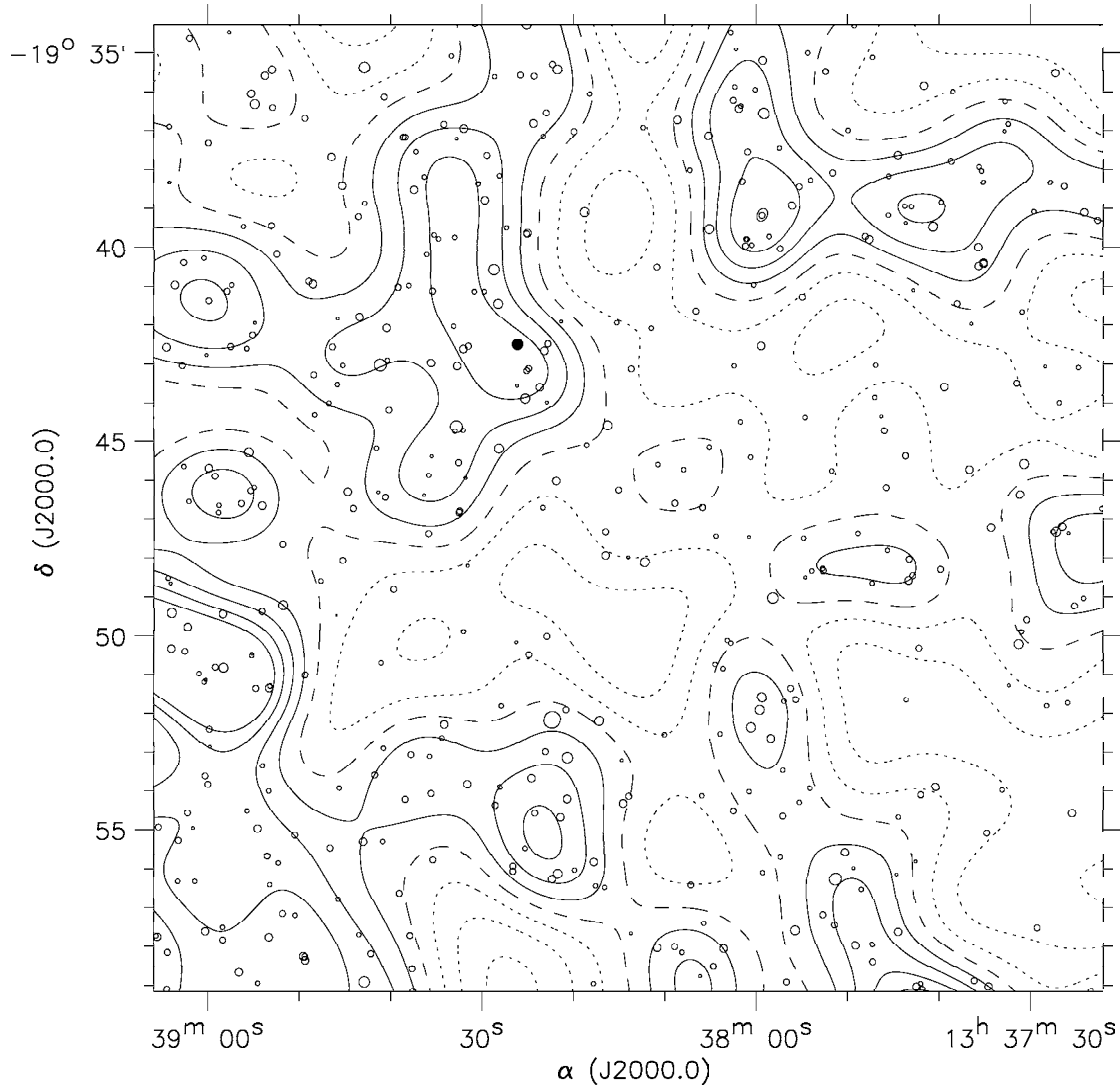
Figure 5.5: Surface density contour map of 437 high-break candidate LBGs ($B - R_C > 2.35$) in the FOV, as in Figure 4.8.
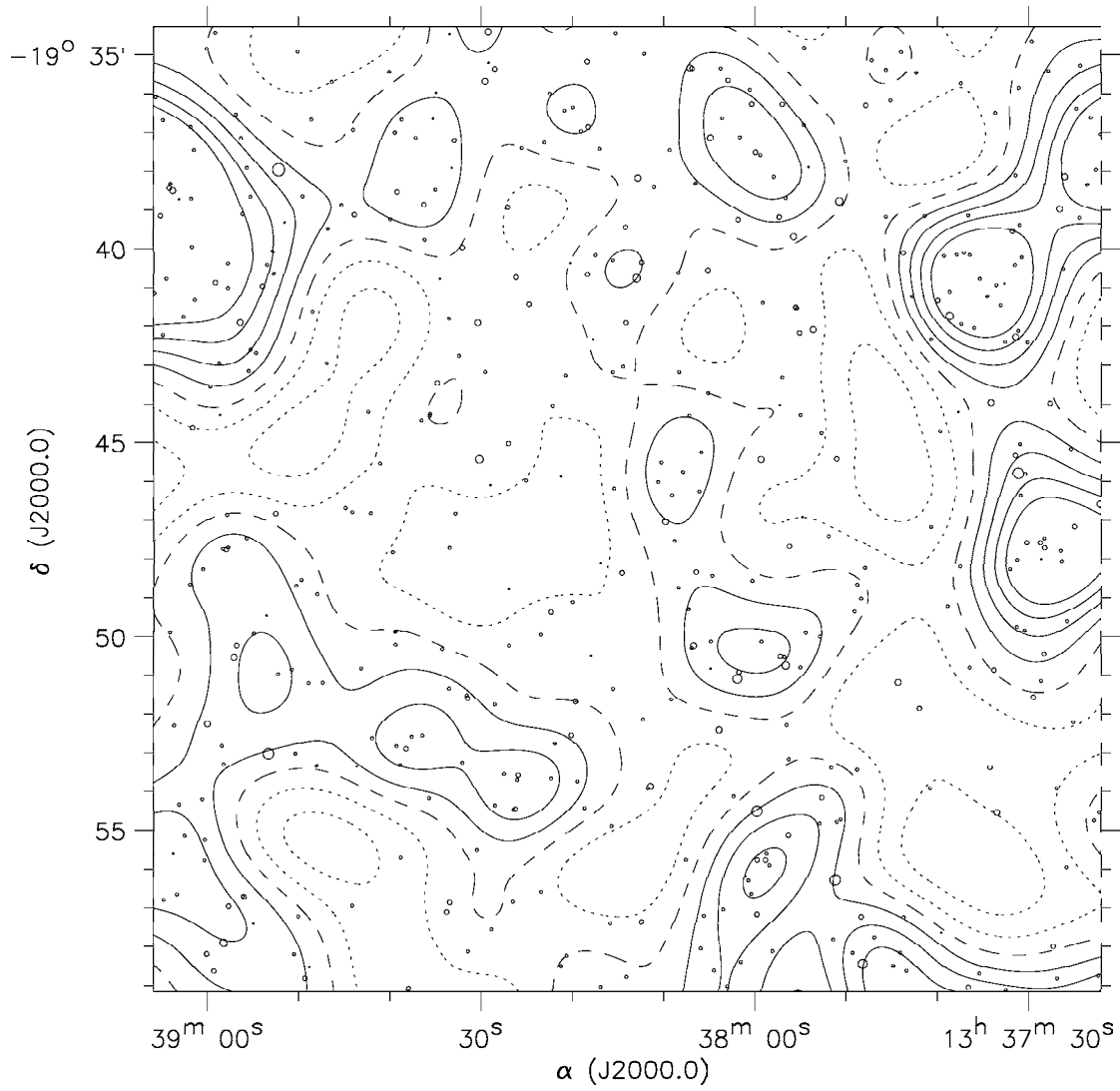
Figure 5.6: Surface density contour map of 437 low-break candidate LBGs ($B - R_C < 2.35$) in the FOV, as in Figure 4.8.

the overdensities in the faint map increases the number density to $\sim 14 \times 10^{-6}$ h^3 Mpc^{-3}. Values found for rich cluster density at low redshift ($z < 0.1$) lie in the range $6 - 12 \times 10^{-6}$ h^3 Mpc^{-3} (Bahcall&Soneira 1983; Postman et al. 1992; Peacock&West 1992; Zabludoff et al. 1993; Mazure et al. 1996). Both our values are similar to this range, which suggests that the comoving cluster number density remained approximately constant or decreased slightly from $z \sim 4$ to $z \sim 0$. This implies that there have been little or no mergers between galaxy clusters during this era.

5.2 Conclusions

Selection criteria based on $B - R_C$ and $R_C - i'$ colors by Ouchi et al. (2004a) select Lyman break galaxies (LBGs) within a redshift interval $z \approx 3.5 - 4.5$. Using these selection criteria, we found 874 candidate Lyman Break Galaxies in a 25×25 arcmin^2 field around the high redshift radio galaxy TN J1338–1942 with a detection limit of $i' < 26.55$. This gives an average surface density of 1.47 ± 0.05 am^{-2}, or a volume density of $\sim 6 \times 10^{-5}$ Mpc^{-3} within a $(\Omega_m, \Omega_\Lambda, h) = (0.3, 0.7, 0.7)$ cosmology. Using simulation results from Ouchi et al. (2004a), we estimated that the completeness of the LBG sample peaks at $z \approx 4$ to $\sim 40\%$, while the contamination is $\sim 3\%$.

Visual inspection of the LBG candidates resulted in high confidence in the resulting catalog (by taking out bad or doubtful detections), although the number counts per magnitude bin are lower than Ouchi et al. (2004a) found in the SDF and SXDF blank fields. This is probably caused by the differences in limiting magnitudes, especially in B. We found that $\sim 23\%$ of the LBG candidates by Miley et al. (2004) and Overzier et al. (in preparation) are selected by our selection criteria, which agrees with our completeness estimate. The same holds for $\sim 23 - 35\%$ of the LAEs (candidate and confirmed members) by Venemans et al. (2002 + extension). The LAEs are selected from a narrow redshift interval $z = 4.07 - 4.12$, centered around the redshift $z = 4.11$ of TN J1338–1942. When adopting a LBG-to-LAE ratio of $4 - 5$ (Steidel et al. 2000), an LBG overdensity of $3 - 5$ in redshift space is found for the range $z = 4.07 - 4.12$. The comparison with Miley et al. and Overzier et al. does not rule out a contamination that is higher than 3%. However, a higher contamination would result in a higher estimate of r_0 (Equations (4.13) and (4.19)), while our current estimate is in good agreement with estimates of Ouchi et al. (2004b).

We attempted to visualize the clustering structure of the candidate LBGs by smoothing and correcting the surface density distribution. The resulting pattern differs from random maps in that (surface) overdense areas are stronger peaked and more distant from each other (more isolated) while underdense areas are larger and more connected. These observed differences are supported by some simple area statistics that show $\sim 7\sigma$ deviations from random. The observed density contrast is moderate: the maximum value in the surface density distribution is less than 4 times the mean value and ~ 25 times the minimum value.

Strong evidence for non-random fluctuations is provided by the two-point angular correlation function (ACF) with a clustering amplitude that is more than 7σ significant over its uncertainty. The two-point spatial correlation function (SCF) is calculated by de-projection of the ACF by using the Limber equation. This results in 2 correlation length estimates $r_0 = 3.6 \pm 1.5$ h^{-1} Mpc and $r_0 = 4.6 \pm 0.3$ h^{-1} Mpc, that both agree well with the correlation lengths at $z \approx 4$ found in the SDF and SXDF by Ouchi et al. (2004b).

When assuming these observed surface overdensities are due to the clustering of galaxies and the observed LBGs are a fair sample of the true (local) galaxy population, then the observed surface density structure supports the idea that galaxies are organized in clusters and filaments surrounded by voids of low galaxy volume density. The observed typical angular diameter of overdense areas is $\sim 5 - 6$ arcmin, which corresponds to ~ 11.5 Mpc at $z = 4$. This is similar to the proto-cluster size of 12 Mpc at $z = 4.9$ found by Shimasaku et al. (2003). The ACF shows a drop at $\theta = 2.5 - 5$ arcmin, which could be related to the typical angular radius of the surface overdense areas. The observed proto-cluster density of $\sim 8 \times 10^{-6}$ h^3 Mpc^{-3} agrees well with the local rich cluster density found by other authors.

TN J1338–1942 is found to inhibit one of the overdense areas in the surface density distribution, although the overdensity peak within this area is not as high compared to other areas within the FOV. However, this overdense area was found to be one of the areas with a high confidence label, because most of the LBGs within this area are relatively bright and have a large $B - R_C$ color (Lyman break). This elongated overdense area, marked by the +25% overdensity contour, extends both ~ 7 `arcmin` north-east and ~ 4 `arcmin` south-east of TN J1338–1942. The total angular length and width of this structure are approximately 10 `arcmin` $\times$ 3 `arcmin`, which corresponds to 21 `Mpc` $\times$ 6.3 `Mpc` at $z = 4$. The distribution of LBG candidates by Miley et al. (2004) and Overzier et al. are found to trace the smoothed density distribution in a relatively small area around TN J1338–1942. The distribution of LAEs by Venemans et al. (2002 + extension) in two larger areas do seem to trace the smoothed density distribution in the direct vicinity of TN J1338–1942, but there is less agreement in the area southeast of TN J1338–1942.

We conclude that there is non-random structure present in the spatial distribution of our LBG candidates. There is a moderate match between the smoothed surface density map of these LBGs and the positions of LAEs and LBGs from other authors in the vicinity of TN J1338–1942. The typical size of overdense areas is estimated at ~ 11.5 `Mpc`, which agrees with the size scales of clusters at low redshift and the proto-cluster size at $z = 4.9$ of Shimasaku et al. (2003).

5.3 Future Work

Some suggestions for future work are given to support and extend the work in this report. The smoothing method might be a powerful tool to find overdense areas at high redshift when the galaxy sample is obtained from a large FOV, is low on contaminants but not very complete. Most urgent, in our opinion, is verification of the observed surface density fluctuations. One approach is to verify the redshift of (a subset of) LBG candidates to evaluate how well the selection criteria worked. Another approach is to make deep wideband images (similar to the ACS images by Miley et al. 2004 and Overzier et al. in preparation), of several small areas in the TN J1338–1942 field, then extract a more complete catalog of LBGs and check that the observed over- and underdensities are consistent with our findings.

Based on the surface density maps in Figures 4.8, 5.3 and 5.5, interesting sites to study more closely for clustering are an overdense area just north-west of TN J1338–1942 ($13^h38^m00^s, -19°37'00''$) and an overdense area south of TN J1338–1942 ($13^h38^m30^s, -19°53'0''$). Both areas are labelled with a high confidence label, have surface density peaks above +100% and contain 1 or 2 brighter galaxies that might qualify as cD-predecessors.

It would be interesting to see if the LBG sample at $z = 4$ in the SDF and SXDF fields of Ouchi et al. (2004a) produces similar overdense structures. During the assigned observing time at the Subaru telescope, J.D. Kurk also made wide-FOV images of two other fields around the high redshift radio galaxies, namely 4C 4114 at $z = 3.80$ and TN J0924–2201 at $z = 5.2$. The work on the TN J1338–1942 field could be repeated on these fields as well, again making use of selection criteria from Ouchi et al. (2004a). This could extend the search for proto-cluster structure out to $z = 5.2$.

5.4 Acknowledgements

Appendix A

Subaru Telescope Observing Proposal

This section contains the original text of the observing proposal that was submitted by J.D. Kurk on May 6, 2002 to the time allocation committee of the Subaru telescope following a call for proposals for the second semester of 2002 (S02B). This proposal included observing the area around the high redshift radio galaxy TN J1338–1942. The proposal was rewarded with 2 nights observation time (January 31st and February 1st, 2003).

(Page 1)

| Subaru Telescope
National Astronomical Observatory of Japan | Semester **S02B**
Proposal ID **S02B–**
Received ___/___/___ |

Application Form for Telescope Time

1. Title of Proposal
A study of the highest redshift structures of galaxies known

2. Principal Investigator
Name: Kurk   Jaron   D

Institute: Sterrewacht Leiden

Mailing Address: P.O. Box 9513, 2300 RA, Leiden, The Netherlands

E-mail Address: kurk@strw.leidenuniv.nl

Phone: ++31-71-5275830   Fax: ++31-71-5275743

3. Scientific Category

☐ Solar System	☐ Normal Stars	☐ Extrasolar Planets	☐ Star and Planet Formation
☐ Compact Objects and SNe	☐ Milky Way	☐ Local Group	☐ ISM
☐ Nearby Galaxies	☐ Starburst Galaxies	☐ AGN and QSO Activity	☐ QSO Absorption Lines and IGM
☐ Clusters of Galaxies	☐ Gravitational Lenses	☐ High-z Galaxies	☐ Deep Surveys
☒ Large-Scale Structure	☐ Cosmological Parameters	☐ Miscellaneous	

4. Abstract *(approximately 200 words)*

We propose to exploit the unique wide-field imaging capabilities of Subaru to study proto clusters that we have discovered at $z \sim 4$ and 5, the most distant known structures of galaxies in the Universe. Deep imaging will be carried out using Suprime-Cam in 3 colours to detect Lyman break galaxies (LBGs) in these structures. Our first target field includes a luminous radio galaxy at $z = 4.1$ surrounded by 20 spectroscopically confirmed Lyα emitting galaxies whose narrow velocity distribution peaks at the redshift of the radio galaxy. Our second target is the most distant known radio galaxy at $z = 5.2$. Narrow band imaging of this field indicates that it is at least as rich as the field around our first target.

Our immediate goals are to (i) extend our 3×3 Mpc maps of the structures up to 20 Mpc, (ii) compare the properties (size, orientation and richness) of the structures defined by the LBGs and the Lyα excess galaxies, (iii) compare the properties of the different classes of proto cluster galaxies and field galaxies at the same distance, (iv) search for possible trends in galaxy properties across the structure and (v) determine whether the structures are isolated bound systems or part of a large scale filamentary web, as simulations of structure formation in the early Universe suggest.

5. Co-Investigators

Name	Institute	Name	Institute
Tadayuki Kodama	NAO, Japan		
Masami Ouchi	DoA, The University of Tokyo		
Bram Venemans	Sterrewacht Leiden, NL		
Huub Röttgering	Sterrewacht Leiden, NL		
George Miley	Sterrewacht Leiden, NL		

6. List of Applicants' Related Publications *(last 5 years)*

De Breuck C. et al. 1999, A&A 352, L51: *VLT Spectroscopy of the $z = 4.11$ Radio Galaxy TN J1338-1942*

Kodama T. et al. 2001, ApJ 562, L9: *The Transformation of Galaxies within the LSS around a $z = 0.41$ Cluster*

Kurk J.D. et al. 2000, A&A 358, L1: *A search for clusters at high redshift - I. Candidate Lyα emitters at $z = 2.2$*

Kurk J.D. et al. 2001, astro-ph/0110131: *A search for high redshift clusters associated with radio galaxies at $2 < z < 4$*

Ouchi M. et al. 2001, ApJ 558, L83: *Clustering properties of galaxies at $z \sim 4$ in the Subaru/XMM Deep Survey Field*

Ouchi M. et al. 2001, astro-ph/0109252: *Statistical properties of Lyman break galaxies at $z \sim 4$*

Ouchi, M. et al. 2002, astro-ph/0202204: *Luminosity Functions and Clustering Properties of Lyα Emitters at $z=4.86$*

Pentericci L. et al. 2000, A&A 361, L25: *A search for clusters at high redshift - II. A proto cluster at $z = 2.16$*

Pentericci L. et al. 2001, astro-ph/0110223: *The population of galaxies in the forming cluster around the radio galaxy MRC 1138-262 at $z = 2.2$*

van Breugel W. et al. 1999, ApJ 518, L61: *A Radio Galaxy at $z = 5.19$*

Venemans B.P. et al. 2002, ApJ 569, L11: *The most distant structure of galaxies known: a proto cluster at $z = 4.1$*

Last modified 02/20/02

Pages 2 and 3 will be used for technical review by observatory staff. Please provide here clear and detailed information for these purposes. The entire proposal including scientific justification will be passed to support astronomers for preparation of observations upon acceptance.

7. Title of Proposal
A study of the highest redshift structures of galaxies known

8. Observing Run

Instrument	# Nights	Moon	Preferred Dates	Acceptable Dates	Observing Modes
Suprime-Cam	3	Dark	March	March	B R_c r' i' z'

Total Requested Number of Nights [3]   Minimum Acceptable Number of Nights [2]

9. List of Targets *(Use an additional sheet if this space is not sufficient)*

[ ] I do **not** want observatory staff to see the target names for the technical review.

Target Name	RA	Dec	Equinox	Magnitude (Band)
LBGs near TN J1338-1942	13 38 26.06	-19 42 30.1	J2000.0	$i = 25 - 26$
LBGs near TN J0924-2201	09 24 19.92	-22 01 41.5	J2000.0	$i = 25 - 26$

10. Scheduling Requirements
Dark time is necessary in order to observe galaxies at $z > 4$. The field of 0924-22 is observable during the first and the field of 1338-19 during the second part of the night. We require 3 nights to achieve both our science goals (**A**), but are willing to concentrate our efforts on the field of 1338-19 only. In that case we will share the nights with another investigator (**B**). The minimum acceptable observing time for the field of 1338-19 is 10 hours, or 2 mornings, so we are willing to accept 2 shared nights as our minimum (**C**). In summary, we will accept the following possibilities in order of preference: (**A**) 3 full nights, (**B**) 3 half nights, (**C**) 2 half nights.

11. Instrument Requirements

12. Experience
MO and TK have extensive experience with imaging observations using Suprime-Cam on Subaru. In particular, MO and collaborators have carried out LBG and Lyα surveys at $z = 4 - 5$ in blank fields using a very similar technique during a Suprime-Cam commissioning run. Hence, they have good knowledge of data reduction/processing and analysis of surveys of LBGs based on Suprime-Cam data. Apart from earlier observations at various telescopes, JK, BV, HR and GM are involved in a large programme at the VLT to image high redshift radio galaxy fields in Lyα and are therefore very experienced in handling optical CCD data.

13. Backup Proposal in Poor Conditions *(specify object names)*
This program benefits from good seeing and high atmospheric transmission. We have to be able to accurately determine colours of very faint and small objects, so bad seeing ($\gtrsim 2.0''$) and heavy cirrus will compromise the science goals by decreasing the depth of the observations. Under extremely poor conditions we will revert to the imaging of $z \sim 1$ clusters around 3C radio sources. We do not need photometric conditions during the entire run, as long as we are able to do a photometric calibration afterwards, for example by a short exposure of the field and a standard star under photometric conditions. In case of severe winds from the direction of our objects (South) which inhibits pointing in the direction of our original objects, we propose to observe the fields of two other radio galaxies: 4C 41.17 at (06 50 52.17 +41 30 30.9) at $z = 3.80$ in the evening and 8C1435+63 at (14 36 38.96 +63 19 04.4) at $z = 4.25$ in the morning. Using exactly the same technique as for the original program we will search for LBGs at $z \sim 4$ in these fields.

Last modified 02/20/02

14. Observing Method and Technical Details

Describe your proposed observations. Please explicitly state the instrument configuration (filters, grisms, slit width, readout mode), intended exposure time, and required sensitivity to achieve your scientific goals. If you propose AO observations, please describe the nature of the target (extended or point source) as well as the guide star properties (separation, brightness, acceptable minimum Strehl ratio). Please read the Call for Proposals carefully.

Observing strategy

To carry out the analysis of the large scale structure at $z \sim 4$, we have to detect a sufficient number of galaxies in a proto cluster at $z = 4.1$. We propose to attain total exposure times of 4.7, 3.6 and 3.2 hours respectively in B, R_c and i', which according to the Exposure Time Calculator (ETC) for Suprime-Cam should result in 3σ limiting AB magnitudes of $B = 28.6$, $R = 28.0$, $i' = 27.1$ in $2''$ apertures and under good circumstances (clear sky, less than 4 days from new moon and $0.8''$ seeing). However, to determine the locus of the objects in the resulting colour-colour diagrams, we need small error bars on the magnitudes and we will therefore only use objects with $i' < 26.3$, which have a signal to noise (s/n) greater than 6 in i'. A typical Lyman break galaxy (LBG) with $i' = 26.3$ will have $R \sim 26.3$ and $B \sim 28.3$, resulting in s/n values of 14 and 4 respectively. Therefore, our faintest LBGs will have magnitude errors of ~ 0.2 in both $R - i'$ and $B - R$. Although the specified total exposure times will not result in equal s/n ratios in the three bands for our objects of study, it is a good compromise between s/n in the colour computation and time constraints. From our previous study in a blank field (the Subaru/XMM Deep Field, see Ouchi et al. 2001), we expect to detect about 50,000 objects in one FoV of Suprime-Cam. We will find at least 1500 LBG candidates satisfying the selection criteria: $B - R > 2.1$, $R - i' < 0.5$, $B - R > 5.4(R - i') + 0.9$. The 1500 LBGs will form a sample of galaxies at $3.3 < z < 4.3$ with a selection completeness of 90%, containing about 20% contaminants. These percentages are derived from our simulations based on the HDF-N galaxy catalog (Ouchi et al. 2001) and will give rise to about 20% uncertainty in the resulting value of the correlation length r_0, which is satisfactory for our proposed analysis. Raising the first selection criterium to $B - R > 2.5$ will predominantly select galaxies at $z > 3.8$ (see Fig. 1 in Ouchi et al. 2001, astro-ph/0109252). These observations fill 3 half nights, but we are willing to accept 2 half nights, in which we will reach $B = 28.4$, $R = 27.8$, $i' = 26.9$, resulting in at least 1000 LBGs at $3.3 < z < 4.3$.

Our second objective is to detect a number of LBGs in a proto cluster at $z = 5.2$. Exposure times of 3.8, 1 and 3.9 hours respectively in r', i' and z' should, according to the ETC, result in 3σ limiting AB magnitudes of $r' = 28.0$, $i' = 26.5$, $z' = 26.0$ in $2''$ apertures and under good circumstances. From a previous blank field study, we expect to find about 200 LBGs at $z \sim 5.2$ with $z' < 26$.

Technical implementation

We will carry out our observations with Suprime-Cam using standard broad band filters. Exposure times for the $z \sim 4$ source will be 20, 9 and 5 minutes in B, R_c and i' respectively per frame, repeated 14, 24 and 32 times to obtain a net exposure time per band of 5.1, 4.4 and 4.3 hours. Including readout times, the total amount of time required for this field is 13.8 hours. Our prime object, 1338-19, has an airmass less than 1.5 during 5 hours per night in March. Therefore, we need 3 nights to complete the program. During the first 4 hours of each night, we will observe the field of 0924-22. Exposure times for the $z \sim 5$ source will be 10, 6 minutes and 190 seconds in r', i' and z' per frame, repeated 23, 10 and 74 times to obtain a net exposure time per band of 3.8, 1.0 and 3.9 hours. Including the overheads, the total amount of time required for this field is 12.3 hours. In order to cover a contiguous region of the sky, taking into account the gaps between the CCDs, we will jitter the frames by about $20''$. This will also facilitate flat fielding and removal of cosmic rays.

15. Proposal Status

☐ The proposed observations are a continuation of previous observations or previously unsuccessful observations with

Proposal ID _____ and Title _____ .

16. Previous Use of Subaru Telescope

Please describe your previous use (in last 3 years) of Subaru Telescope and the status of data reduction/analysis and publication.

Year/Month	Proposal ID	PI name	Status: data reduction/analysis	Status: publication
2000/11	GTO	Okamura, S.	Reduced and analysed	Publ.: Ouchi et al. 2001
2001/01	S00-024	Kodama, T.	Reduced and analysed	Publ.: Kodama et al. 2001
2001/3-6	GTO	Okamura, S.	Reduced and analysed	Submitted to ApJ
2001/11	S01B-055	Shimasaku, K.	Data being processed	
2002/02	S01B-146	Nakata, F.	No data due to bad weather	
2002/04	S02A-128	Fukugita, M.	Data being processed	

17. Thesis Work

☒ This proposal is linked to the thesis preparation of B. Venemans _____

Galaxy structures at high redshift associated with radio galaxies _____

18. Subaru Open Use Intensive Programs

☐ This is a proposal for Intensive Programs.

1 INTRODUCTION

1.1 Proto clusters around high redshift radio galaxies

One of the most intriguing questions in modern astrophysics concerns the formation of structure in the early Universe (e.g. Bahcall et al. 1997). Low redshift ($z < 0.2$) galaxy surveys show that the distribution of luminous mass in the Universe is filamentary with voids on scales up to 200 Mpc. Simulations of the evolution of dark matter suggest that the most massive forming systems of galaxies will be found at the nodes of these filaments, but this cosmological web at high redshift has not been observed yet, apart from a possible linear structure of Lyα emitters at $z \sim 3$ (Møller & Fynbo 2001). Mapping high redshift clusters and their filamentary extensions would provide a fundamental test of the validity of the simulations and constrain both cosmological parameters and the mechanisms of galaxy formation.

The most luminous high redshift radio galaxies (HzRGs, $z > 2$) are unique probes of galaxy and cluster formation (reviews by e.g. McCarthy 1993; Röttgering & Miley 1996; Miley 2000). They are amongst the oldest and most massive galaxies in the early Universe (Fig. 1) and there is considerable evidence that they are forming galaxies at the centres of clusters or proto clusters. The HzRGs themselves have properties that would be expected of forming central cluster galaxies. Their extremely clumpy morphologies revealed by HST images are strikingly similar to simulations of forming brightest cluster galaxies, based on hierarchical models. Moreover, they are usually surrounded by giant Lyα emitting gas halos, whose sizes (~ 100 kpc) are comparable to those of cD galaxy envelopes (e.g. van Ojik et al. 1996, 1997; Pentericci et al. 1997).

1.2 Structures at z $\sim$ 2.2 to z $\sim$ 4.1

For the above reasons we initiated a large VLT program to find and study Lyα emitting galaxies in cluster sized regions around HzRGs. Deep narrow and broad band imaging are being used to locate candidate galaxies having excess Lyα emission and spectra are being taken subsequently. The program has been extremely successful to date. **The first 4 targets** that we studied with the VLT at $z \sim 2.2, 2.9, 3.1$ and 4.1 all have **spectroscopically confirmed substantial galaxy overdensities** with velocity peaks coinciding to within < 200 km s^{-1} with the velocity of the radio galaxies (Kurk et al. 2000, 2001, Pentericci et al. 2000, Venemans et al. 2002).

Our results have provided strong evidence that the most luminous radio sources at $z > 2$ are generally associated with the large bright galaxies, surrounded by substantial galaxy overdensities and strengthened our hypothesis that **HzRGs are massive forming galaxies at the centre of proto clusters**. The luminosity functions and lifetimes of luminous HzRGs are consistent with such objects being associated with the redshift *spikes* of samples of LBGs (Steidel et al. 1998).

1.3 1338 − 19: a proto cluster at z $\sim$ 4.1

Narrow band imaging followed by multi object spectroscopy revealed 20 Lyα emitting galaxies within 1.5 Mpc around the radio galaxy. Their velocity dispersion is ~ 350 km s^{-1} (Fig. 2) and the redshift distribution is centered on the radio galaxy. The galaxy overdensity of the emitters implies a mass for the structure of $1 - 2 \times 10^{15}$ M$_\odot$ (Venemans et al. 2002).

The establishment that galaxies occur in overdensities at such an early epoch in the history of the Universe is a fundamental result that must be explained by models of galaxy and cluster evolution. Our structure at $z \sim 4.1$ provides an important laboratory for constraining such models. The spatial distribution of the galaxies is inhomogeneous: most emitters are located in the south-east and none in the west (Fig. 2), but the size of the $7' \times 7'$ FORS field is insufficient to determine the size and shape of the entire structure. Recently, we have obtained additional FORS narrow band imaging to the south-east which indicates that the structure extends up to 5 Mpc. **The main goal of our present proposal is to use the uniquely large field provided by Suprime-Cam of $34' \times 27'$ to determine the boundaries of the structure at z = 4.1 and map the large scale structure on scales up to 20 Mpc.**

Although the discovery of the galaxy structure around 1338 − 19 is recent, this field has long been a target of the GTO of the Advanced Camera for Surveys on the HST. GM is a member of the ACS IDT and will have access to these data.

1.4 0924 − 22: a proto cluster at z $\sim$ 5.2?

The most distant known radio galaxy is 0924-22 at $z = 5.2$ (van Breugel et al. 1999). Radio and infrared observations indicate that this galaxy is in a dense environment and experiences vigorous star formation (200 M$_\odot$ yr^{-1}). Recently, we have obtained deep VLT narrow band (8h) and broad band I (2.5h) and V (1h) imaging of this field. Based on narrow band excess flux and V and I magnitude, we find > 25 bona fide candidate Lyα emitters. The number of candidates around 0924-22 indicates that the structure of galaxies at $z = 5.2$ is at least as rich as the structure around 1338-19. **Our second goal is therefore to find an additional population of star forming galaxies at z = 5.2.**

2 PROPOSED OBSERVATIONS AND ANALYSIS

The spectral energy distribution from young galaxies peaks in the UV, but intergalactic and interstellar neutral hydrogen cause a sharp drop in the light we receive from high redshift galaxies below 1216 and 912 Å (rest frame) respectively. This drop can be used to select star forming galaxies at $z > 3$ by multi colour imaging (Steidel et al. 1996). Our initial plan is to apply the Lyman break technique (see §14) to select objects with photometric redshifts consistent with $z = 4.1$ or 5.2 and to search for non-uniformities in the distribution of these in order to map the large scale structure. Additionally, we will compare the spatial and colour characteristics of the selected objects to other objects in the structures, to LBGs at $z \sim 4$ selected with the same criteria in a blank field and to simulations of structure formation in the early Universe.

3.1 Large Scale Structure around a z ~ 4 and z ~ 5 proto clusters

The large field of Suprime-Cam allows us to map the structure at $z \sim 4$ (5) up to ~ 20 (17) Mpc in projection. Investigation whether high redshift proto clusters are isolated bound structures or enhancements in a larger scale cosmic filamentary web provides important input to models of large scale structure formation. Using the Limber transformation (Peebles 1980), we will compute the spatial correlation function of the Lyman break galaxies. LBGs are a different but complementary population to Lyα excess galaxies (Steidel et al. 2000, Malhotra & Rhoads 2002). Lyα excess galaxies must have very little dust and may well represent an earlier evolutionary phase. A comparison of the extent and correlation functions of the structures defined by the Lyα and Lyman break galaxies in the region of overlap with the FORS fields can therefore provide constraints on the formation and evolution of galaxies and clusters.

3.2 Comparison with a blank field

We will compare the characteristics and especially the clustering properties of the detected LBGs with the about 1200 LBGs at $z \sim 4$ found in the Subaru/XMM-Newton Deep Field (Ouchi et al. 2001). They find a correlation length of 2.7 Mpc, which is twice as large as predicted for dark matter at this redshift but twice as small as that predicted for bright galaxies. By targeting our concentration of galaxies we should be able to significantly increase this signal and detect any differences with the *field* at this redshift.

3.3 Environmental variation of galaxy properties

The dependence of galaxy magnitudes and colours will be derived as a function of galaxy overdensity (rich cluster to field) and redshift and compared with models of galaxy and cluster formation and evolution (e.g. Kauffmann 1996). Trends in the above mentioned parameters will be examined for consistency with proposed scenarios for the nature of LBGs and their role in galaxy and cluster formation (e.g. Wechsler et al. 2001; Ellis 2001). Following Brinchmann & Ellis (2000), we shall attempt to use the colour statistics of this field combined with other data to constrain the rate at which galaxies are assembling and compare this to model predictions (e.g. Baugh et al. 1998). Within the cluster, difference in sizes and morphologies between the Lyα and LBGs and trends with location will provide direct information about galaxy evolution, particularly when complemented with differences in such parameters as a function of redshift.

The detailed study of the proto clusters at z = 4.1 and 5.2 will considerably advance our knowledge of high redshift galaxies and structure and will heavily constrain models of galaxy and structure formation.

Follow-up of the observations would comprise spectroscopy of the LBGs in the proto clusters at $z \sim 4$ and 5 to refine their redshifts and subsequently carry out a detailed analysis of their properties (e.g. spatial distribution, absorption lines).

References:

Bahcall et al. 1997, ApJ 485, L53
Baugh et al. 1998, ApJ 498, 504
Brinchmann & Ellis 2000, ApJ 536, L77
De Breuck et al. 2000, A&AS 143, 303
Ellis 2001, astro-ph/0102056
Kauffmann, 1996, MNRAS 281, 487
Kirshner et al. 1993, ApJ 415, 589
Kurk et al. 2000, A&A 358, L1
Kurk et al. 2001, astro-ph/0110131
Malhotra & Rhoads 2002, ApJ 565, 71

McCarthy 1993, ARA&A 31, 639
Miley 2000, in *The VLT Opening Sympo-sium*, Springer-Verlag
Møller & Fynbo 2001, A&A 373, 796
van Ojik et al. 1996, A&A 313, 25
van Ojik et al. 1997, A&A 317, 358
Ouchi et al. 2001, ApJ 558, L83
Peebles 1980, *The Large-Scale Structure of the Universe*, Princeton Univ. Press
Pentericci et al. 1997, A&A 326, 580

Pentericci et al. 2000, A&A 361, 25
Röttgering & Miley 1996, in *The Early Universe with the VLT*, Springer-Verlag
Steidel et al. 1996, ApJ 462, L17
Steidel et al. 1998, ApJ 492, 428
Steidel et al. 2000, ApJ 532, 170
van Breugel et al. 1999, ApJ 518, L61
Venemans et al. 2002, ApJ 569, L11
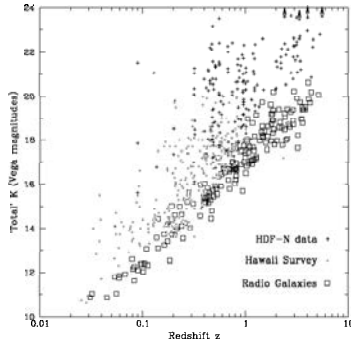Wechsler et al. 2001, ApJ 554, 85

Fig. 1. Composite Hubble K-z diagram of radio loud and quiet galaxies, indicating that HzRGs are amongst **the most massive galaxies** in the early Universe. The radio-loud galaxies trace the upper envelope of the K band luminosity distribution (De Breuck et al. 2000).
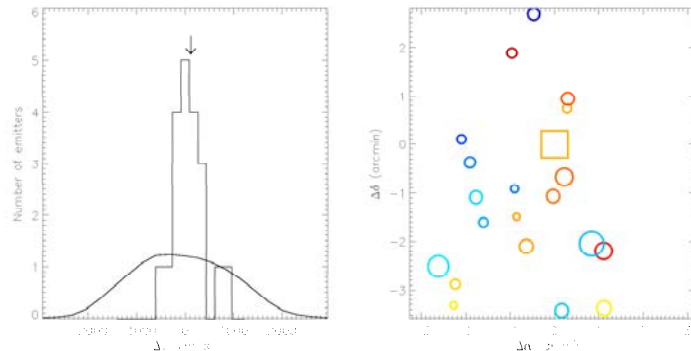

Fig. 2. *Left* Velocity distribution of the Lyα emitters around radio galaxy TN J1338-1942. The bin size is 175 km s^{-1} and the median of the velocities is taken as zero point. The velocity of the radio galaxy is indicated by an arrow. The transmission curve of the narrow band filter is also plotted. *Right* The spatial distribution of the emitters is indicated by circles (a square for the radio galaxy). The size of the circles is scaled according to their Lyα flux and their colour to their redshift: from blue ($z = 4.0882$) to red ($z = 4.1115$).

Bibliography

[1] Abazajian, K., et al., 2004, astro-ph/0410239

[2] Albrecht, M.A., et al., 1997, ADASS, **6**, 333

[3] Bahcall, N.A., & Soneira, R.M., 1983, ApJ, **270**, 20

[4] Bahcall, N.A., 1988, ARA&A, **26**, 631

[5] Barger, A.J., et al., 1999, ApJ, **518**, 5

[6] Barth, A.J., 2001, ADASS, **10**, 385

[7] Bertin, E., 1996, A&AS, **117**, 393

[8] Buttery, H.J., 2003, *New Methods for Detecting High-Redshift Clusters of Galaxies*, Ph.D. Thesis, Newnham College, Cambridge University, UK

[9] Colless, M.M., et al. 2003, astro-ph/0306581

[10] Cousins, A.W.J., 1978, MNASSA, **37**, 8

[11] De Breuck, C., et al., 1999, A&A, **352**, 51

[12] De Breuck, C., et al., 2000, A&AS, **143**, 303

[13] De Breuck, C., et al., 2002, AJ, **123**, 637

[14] Edge, D.O., et al., 1959, MNRAS, **68**, 37

[15] Efstathiou, G., et al., 1991, ApJ, **380**, 47

[16] Fanning, D.J., 2003, *IDL Programming Techniques, 2nd Edition*, Fanning Software Consulting, Fort Collins (CO), USA

[17] Fukugita, M., et al., 1996, AJ, **111**, 1748

[18] Giavalisco, M., 2002, ARA&A, **40**, 579

[19] Giuricin, G., et al., 1993, ApJ, **407**, 22

[20] Groth, E.J., & Peebles, P.J.E., 1977, ApJ, **217**, 385

[21] Hamilton, A.J.S., 1993, ApJ, **417**, 19

[22] Hewitt, P.C., 1982, MNRAS, **201**, 867

[23] Hickson, P., & Mulrooney, M.K., 1998, ApJ, **506**, 191

[24] Hogg, D.W., 2000, astro-ph/9905116, v4

[25] Huchra, J., et al., 1983, ApJS, **52**, 89

[26] Jacoby, G., 1998, AAS, **192**, 6101

[27] Johnson, H.L., & Morgan, W.W., 1953, ApJ, **117**, 313

[28] Joye, W. A., & Mandel, E., 2003, ADASS, **12**, 489

[29] Kauffman, G., 1999, AAS, **195**, 6701

[30] Kron, R.G., 1980, ApJS, **43**, 305

[31] Labbé, I., et al., 2003, AJ, **125**, 1107

[32] Landsman, W.B., 1993, ADASS, **2**, 246

[33] Landolt, A.U., 1992, AJ, **104**, 340

[34] Landy, S.D., & Szalay, A.S., 1993, ApJ, **412**, 64

[35] Levine, D.A., et al., 1998, ApJ, **504**, 64

[36] Limber, D.N., 1953, ApJ, **117**, 134

[37] Lonsdale, C.J., & Hacking, P.B., 1989, ApJ, **339**, 712

[38] Madau, P., 1995, ApJ, **441**, 18

[39] Matthews, T.A., et al., 1964, ApJ, **140**, 35

[40] Mazure, A., et al., 1996, A&A, **310**, 31

[41] Miley, G.K., et al. astro-ph/0401034

[42] Miyazaki, S., et al., 2002, PASJ, **54**, 833

[43] Möller, P., & Fynbo, J.U., 2001, A&A, **372**, 57

[44] Monet, D., et al., 1998, *The USNO-A2.0 Catalogue*, U.S. Naval Observatory, Washington DC, USA

[45] Ouchi, M., et al., 2001, ApJ, **558**, 83

[46] Ouchi, M., et al., 2004, ApJ, **611**, 660

[47] Ouchi, M., et al., 2004, ApJ, **611**, 685

[48] Overzier, R.A., et al., 2003, A&A, **405**, 53

[49] Palunas, P., et al., 2004, ApJ, **602**, 545

[50] Pavlovsky, C., et al., 2004, *ACS Instrument Handbook v5.0*, STScI, Baltimore

[51] Peacock, J.A., & West, M.J., 1992, MNRAS, **259**, 494

[52] Peacock, J.A., 2001, *Cosmological Physics, 3rd reprint*, Cambridge University Press, Cambridge, UK

[53] Peebles, P.J.E., 1973, ApJ, **185**, 413

[54] Peebles, P.J.E., 1980, *The Large Scale Structure of the Universe*, Princeton University Press, Princeton (NJ), USA

[55] Pilkington, J.D.H, & Scott, P.F., 1965, MNRAS, **69**, 183

[56] Pooley, G.G., & Kenderdine, S., 1968, MNRAS, **139**, 529

[57] Postman, M., et al. 1992, ApJ, **384**, 404

[58] Puschell, J.J., et al., 1982, ApJ, **257**, 57

[59] Reid, I.N., et al., 1991, PASP, **103**, 661

[60] Röttgering, H.J.A., et al., 1994, A&AS, **108**, 79

[61] Schlegel, D., et al., 1998, ApJ, **500**, 525

[62] Shectman, S.A., et al., 1996, ApJ, **470**, 172

[63] Shimasaku, K., et al., 2003, ApJ, **586**, 111

[64] Shimasaku, K., et al., 2004, ApJ, **605**, 93

[65] Smoot, G.F., et al., 1991, ApJ, **371**, 1

[66] Somerville, R.S., et al., 2004, ApJ, **600**, 71

[67] Songaila, A., 2004, AJ, **127**, 2598

[68] Spergel, D.N., et al., 2003, ApJS, **148**, 175

[69] Steidel, C.C., & Hamilton, D., 1993, AJ, **105**, 2017

[70] Steidel, C.C., et al., 1996, ApJ, **462**, 17

[71] Steidel, C.C., et al., 2000, ApJ, **532**, 170

[72] Strauss, M.A., et al., 2002, AJ, **124**, 1810

[73] Venemans, B.P. et al., 2002, ApJ, **569**, 11

[74] Yagi, M., 1996, *Data Reduction Software for Mosaic CCD Cameras*, Subaru Telescope Technical Report, **55**, 73

[75] Zabludoff, A.I., et al., 1993, AJ, **106**, 1273