Université de Liège
Faculté des Sciences
Département d'Astrophysique, Géophysique et Océanographie

A K_s–band selected, multi–wavelength survey for quasars in the XMM–LSS field

Theodoros Nakos

Doctoral Dissertation
January 2007

Pr. Jean–Pierre Swings	President of the jury
Pr. Jean Surdej	Thesis co–Adviser
Dr. Jon Willis	Thesis co–Adviser
Dr. Jean–François Claeskens	Examiner
Dr. Damien Hutsemékers	Examiner
Dr. Grégor Raw	Examiner
Dr. Stefano Andreon	External Examiner
Pr. Herwig Dejonghe	External Examiner

Caminante, son tus huellas
el camino, y nada mas;
caminante, no hay camino,
se hace camino al andar.

Antonio Machado
Canto XXIX, Proverbios y cantares,
Campos de Castilla, 1917

Marcheur, ce sont tes traces
ce chemin, et rien de plus;
marcheur, il n'y a pas de chemin,
le chemin se construit en marchant.

Acknowledgments

I would like to express my heartfelt thanks to all who contributed in any way to this work. First of all my thesis co–adviser, Jean Surdej, for his enthusiasm, support and stimulating interaction. A great thanks also goes to Jon Willis, who offered me the main bulk of data on which this work is based and provided valuable advice on many "technical" and scientific issues. Stefano Andreon should also be acknowledged, for providing me part of the material that played a major role in this work and for his suggestions during the course of our collaboration.

Starting from my home Institute, there are numerous persons from the University of Liège I am grateful to. First of all to Jean–Pierre Swings, for his advices concerning both administrative and scientific matters. An enormous thanks goes to Pierre Riaud, for all the things he taught me, the hours he spent discussing with me and for creating such an amazing atmosphere at the Institute. Olivier Garcet should also be acknowledged, for the productive discussions concerning the properties of the X–ray data and the work on the photometric redshifts. I am also indebted to Jean-François Claeskens, whom I have been "bombarding" with questions since so many years. I have learned a lot from him and I would like to express my gratitude once more. Finally, many thanks also go to Denise, Damien, Eric and Frederic for their advices and solutions regarding practical, administrative or scientific issues.

Going back in time, I would like to acknowledge the European Southern Observatory for the ESO studentship and especially Danielle Alloin, who supervised my work during my two–year stay at Santiago, as well as Alain Smette and Poshak Gandhi for the fruitful discussions.

Since I spent the early years of my thesis at the Royal Observatory of Belgium, I ought special thanks to a bunch of people there. First of all to Edwin van Dessel, who had been (and still is) supporting me for so many years. Many other colleagues, namely Marijke Burger, Patricia Lampens, David Duval, Jan Cuypers and Henri Boffin, should also be mentioned. Finally, the ex–director of the Observatory, P. Pâquet, has to be acknowledged, for his strong commitment in supporting me regarding financial and administrative issues.

There is a couple of persons I would also like to acknowledge from the National Observatory of Athens. First of all Dimitri Sinachopoulos, for being an excellent teacher and

a very good friend. I learned a lot thanks to him and I will always feel grateful. Secondly Panos Boumis, who contributed so much during the preliminary stages of my adventure in astronomy and, by now, ended up being a very good friend.

Finally, I would also like to thank Evanthia Hatziminaoglou for sharing with me her expertise during my two–week visit to Tenerife, for collaborating on the last chapter of the thesis and for running the code for modeling the infrared properties of the AGN in my sample.

Leaving the scientific environment, there are tens of persons who have not really contributed much that can be cited in this work, but they have surely made my life much more beautiful and helped me overcome the difficult moments. First of all my Greek friends in the homeland and abroad. They are numerous and I can not name them all. Nevertheless, special thanks should go to one of them, Foris, without whom things would surely be very, very different...Of course, Maria, Christina, Dimitris and Yiorgos also played a great role in this. Credits should also go to all the friends with whom I shared so many wonderful moments in Chile: "Daniel–son", Nuri, Claudito & Lu, Gael, Jon–Jon, Ivo and all the other "chiquillos".

Last, but surely not least, all my love and gratitude go to the persons who have been closest to me during all these years. My parents and my sister, for their everlasting love, support and understanding. Time has not only made us older, but wiser, and we have benefited so much from this... My daughter, Carolina, who never lets me forget that the happiness in life is found in the most simple things. Finally, Alejandra, for her unconditional love, who made me see the world with different eyes. This thesis is dedicated to them, they all deserve it...

Thodori Nakos

Preface

The phenomenon of Active Galactic Nuclei (AGN) is still puzzling astronomers, despite the fact that the very first discovery of an active galaxy took place a century ago (in 1908, to be more precise). The term AGN refers to a special type of galaxies, whose radiation is mainly produced by non–thermal processes in a small, compact region in the nucleus of the galaxy. Astronomers are now confident that this region, of the size of our solar system, consists of a super–massive black hole, constantly fed with matter from an accretion disk surrounding it.

The AGN "zoo" is populated by BL Lac objects, Type–1 and Type–2 quasars and Seyfert galaxies, etc. Although the AGN classification is mainly based on the amount of radiation emitted by the compact central source, the distinction between them is to some degree a matter of definition. The numerous sub–categories into which AGN have been classified is a self–proof on how incomplete is our knowledge regarding the processes taking place in the active nucleus. The "holy grail" for AGN astronomers, a unified model that will explain, using simple geometrical and physical solutions, the different features we observe in AGN, is far away from complete.

The quasars (quasi–stellar objects, or QSOs), one of the two main sub–categories of AGN, are among the most luminous and distant objects. The light of the most distant quasars registered on our detectors was emitted when the Universe was only a fraction of its current age. Thanks to their high luminosities, quasars serve as cosmological light-houses: even at the far edges of the Universe there are still some shinning, and capturing their light traces back to the very early cosmic history.

Once the importance of quasars was recognized, numerous QSO surveys were initiated for their discovery. However, their identification had to be based on some selection criteria. Objects not respecting these criteria were excluded, thus introducing a bias in the selected sample. The most striking example is the coincidental discovery of the first QSOs, as radio sources, in the late 1950s. We now know that the percentage of "radio–loud" quasars is only $\sim 5-10\%$ of the total QSO known population. Working on sub–samples of the parent population prevents us from building up a single model that can successfully describe the features we observe in the various sub–classes. For this reason it is essential to understand the selection effects related to each survey and how the observed populations fit in a more general AGN context.

One of the most commonly used optical techniques for the discovery of quasars is the so–called Ultra–Violet excess (UVX), implemented for the first time in the mid 1960s (Sandage 1965), and more systematically in the early 1980 (Schmidt & Green 1983). This technique is based on the fact that, in the $(U-B)$ versus $(B-V)$ color–color plane, quasars occupy a locus different from that of stars, because of their bluer colors. These colors are attributed (a) to the quasar blue continuum, and (b) to the Lya emission line (1216Å), which, for QSOs found until a given redshift limit, enters the U–band. The UVX method works quite well up to redshifts $z \approx 2$. At higher redshifts, however, the $U-B$ color starts becoming less effective in isolating quasars, and for redshifts higher than 2.2, when the Lya line is shifted to wavebands redder than the U–filter, the method fails in detecting high redshift quasars.

As mentioned before, each survey suffers from selection effects. Understanding the properties of a sample, selected using specific criteria, is fundamental for properly describing the parent population. Because of the filter combination, the UVX favors the selection of blue quasars. Hence, intrinsically red, and reddened, due to dust, QSOs are not selected by UVX. As a result, the question whether we are missing a "hidden" quasar population has been troubling astronomers since about a decade (Webster et al. 1995).

A new approach to tackle the weak points of the $UV-$excess was introduced by Warren et al. (2000). The K–excess (KX), similar in its concept to the UVX, suggested to use infrared filters (among which the K–band, at $\sim 2.2\,\mu$m) for detecting high$-z$, red and dusty QSOs. The quasar spectral energy distribution (SED) is described by a power–law, while stellar SEDs have a convex shape. Because of their flux–excess in the infrared, quasars should occupy a different locus with respect to stars in the infrared color plane, as they do in the UBV plane.

Due to the technological difficulties in constructing infrared detectors covering a large area on the sky, we have only recently started performing infrared surveys with ground–based facilities. In the following years, ambitious large–area IR programs, such as the UKIDSS, will have covered some thousands of square degrees. In this way it will be possible to complement our knowledge on the QSO population, to which deep optical (e.g. SDSS) and shallow IR (e.g. 2MASS) surveys have significantly contributed. On the other hand, current space–born facilities, like the Spitzer satellite, have clearly demonstrated the power of using infrared wavebands for probing the AGN population.

The work I did during my thesis has been laborious, in terms of data reduction, but modest in terms of surface, effort on exploring the possibility of discovering "red and reddened" quasars via the KX. The following chapters describe the data reduction, the selection process and the multi–wavelength properties of the K_s–selected quasar sample.

Abstract

My thesis project consisted in the data reduction and analysis of a set of K_s–band data. The near–infrared (NIR) observations were performed in the context of XMM–LSS, an X–ray survey with the main goal of studying the Large Scale Structures (LSS) of the Universe using ESA's satellite XMM–Newton. Multi–wavelength observations, coming from Legacy surveys associated to XMM–LSS, allowed not only to perform a detailed study of the properties of the selected population but also understand in–depth the selection process followed. The outline of the thesis is as follows:

Chapter 1 sets the scientific context of my research, describes the basic characteristics of active galaxies and quasars, their energy production mechanism and the merging of the various AGN sub–categories to a single scheme, via the Unification Model. It also introduces the concept of the classical method for discovering quasars in imaging surveys, the UV–excess (UVX). In the same chapter we also develop the problematic regarding a missing red or reddened, due to the presence of dust, quasar population, that would not be detected by UVX due to a selection bias. Finally, we introduce the concept of the K–excess (KX), which is meant to select the quasars missed by UVX using infrared color criteria.

Chapter 2 discusses the properties of infrared detectors, their differences compared to optical CCDs and the peculiarity of the IR sky–background. It presents the K_s–band observations and all the steps followed for reducing the data, in an extensive way. It describes the calibration procedures (photometric & astrometric), the construction of some mosaic images, the simulations that helped understand the properties of these mosaics and, finally, how the K_s–band catalog was built. Although I have to admit that it is a rather technical chapter, I hope it serves as a cook–book for those who are interested in doing a similar work.

Chapter 3 presents the properties of two photometric (R, z'–bands) and a spectroscopic (2dF facility) catalog, that played an essential role in the implementation of the KX technique. It describes the selection procedure, performed on the basis of $Rz'K_s$ colors and morphological criteria, and how it was refined using the spectroscopically identified galaxies and quasars together with some simulations.

Chapter 4 describes the multi–wavelength properties of the quasar candidate sample in the X–rays (XMDS survey), in the optical (using data from the CFHT Legacy survey)

and the near–to–mid infrared, thanks to data coming from the Spitzer satellite (SWIRE survey).

Chapter 5 presents the spectral energy distribution (SED) of the quasar candidates and how, on the basis of their SED it was possible to identify the false candidates (stars and galaxies) contaminating the QSO sample, from the true quasars. Additionally, due to the lack of spectroscopic data, we explain how we estimated the redshift of the KX–selected quasars using the available multi–wavelength photometry.

Chapter 6 deals with the evaluation of the KX–method, the origin of the contamination and the possibility of having missed any quasars or AGN, by comparing our detections to models available in the literature. We will describe the steps followed for discovering a population of Type–2 AGN, present their multi–wavelength properties and calculate their photometric redshift.

Chapter 7 gives an insight to the environment of the AGN and quasars found in our survey. Using a versatile torus model, that takes under consideration both the physical processes taking place in the torus' interior, attributed to dust, and the vast variety of possible geometrical configurations. we fit the spectral energy of the objects. This study has made it possible to investigate on the Unification Scheme hypothesis, that wants the observational differences among the AGN sub–classes to be attributed to a viewing angle effect.

Finally, in Chapter 8 we draw some conclusions and present the future perspectives.

Resumé

Notre projet de thèse a consisté en la réduction et l'analyse d'un ensemble de données obtenues dans la bande photométrique K_s. Ces observations dans le proche infra–rouge ont été réalisées dans le contexte du XMM–LSS, un sondage systématique d'une zone du ciel dont le but principal est d'étudier les Structures à Grandes Echelles de l'Univers à l'aide du satellite XMM–Newton de l'ESA. Des observations dans différents domaines de longueur d'onde obtenues par des programmes d'observations associés ont permis non seulement d'étudier en détail les propriétés de la population d'objets sélectionnée mais aussi de comprendre en profondeur le procédé de sélection appliqué aux données. Le travail s'articule comme suit:

Le Chapitre 1 fixe le contexte scientifique de notre recherche et décrit les propriétés fondamentales des galaxies actives et des quasars (le mécanisme de production de leur énergie ainsi que la description des différents types de noyaux actifs (AGN) par un seul Model d'Unification). Ce Chapitre introduit également le concept "d'excès UV" (UVX), qui est la méthode classique de recherche des QSOs dans les sondages du ciel en imagerie optique. Nous y développons aussi la problématique des objets manquants rouges ou rougis, c'est–à–dire non détectés par la méthode UVX, soit à cause de la présence de poussières, soit en raison de propriétés intrinsèques. Finalement, nous introduisons le concept "d'excès–K" (KX) dont le but est de sélectionner les quasars manqués par la technique UVX en utilisant des critères de couleur dans l'infra–rouge.

Le Chapitre 2 discute les propriétés des détecteurs infra–rouges, leurs différences comparées aux détecteurs CCDs dans le domaine optique, et les particularités du fond de ciel dans le proche infra–rouge. Il présente les observations obtenues dans la bande K_s et décrit en détail les différentes étapes de la réduction des données. Il décrit également les procédures de calibration (photométrique et astrométrique), la construction de mosaiques d'images, les simulations qui ont aidé à comprendre les propriétés de ces dernières et, finalement, comment le catalogue en bande K_s a été construit. Quoique ce Chapitre soit plutôt technique, nous espérons qu'il puisse servir de guide pratique à tous ceux intéressés par un travail similaire.

Le Chapitre 3 présente les propriétés de deux catalogues photométriques (en bande R et z') et d'un catalogue spectroscopique (du projet 2dF), lesquels ont joué un rôle essentiel dans l'implémentation de la technique KX. Ce Chapitre décrit ensuite la procédure de

sélection des candidates de quasars, réalisée sur base des couleurs $Rz'K_s$ et de critères morphologiques, et comment elle a été raffinée à l'aide de quasars et de galaxies identifiées spectroscopiquement, ainsi que de simulations.

Le Chapitre 4 décrit les propriétés multi–couleurs de l'échantillon de candidats quasars: dans le domaine des rayons–X (observations du XMDS), dans le domaine optique (données du "CFHT Legacy Survey") et dans l'infra–rouge proche et moyen (données du programme d'observations SWIRE avec le satellite Spitzer).

Le Chapitre 5 présente la distribution spectrale d'énergie des candidats quasars et comment, sur base de cette dernière, il a été possible d'identifier de faux candidats (étoiles et galaxies) contaminant l'échantillon. En outre, en l'absence de données spectroscopiques, nous expliquons comment nous avons estimé le redshift des candidats quasars sélectionnés par la méthode KX, en utilisant la photométrie multi–couleurs disponible.

Le Chapitre 6 traîte de l'évaluation de la méthode KX, de l'origine de la contamination et des éventuels quasars ou AGN manqués, en comparant nos détections aux prédictions des modèles disponibles dans la litérature. Nous y décrivons les étapes suivies pour découvrir une population d'AGN de Type 2, dont les propriétés multi–couleurs sont présentées et dont les redshifts photométriques sont calculés.

Dans le Chapitre 7 nous discutons quelques résultats se rapportant à l'environnement des AGN et quasars trouvés dans notre survey. En utilisant un modèle polyvalent pour le tore, prenant en compte à la fois les processus physiques à son intérieur, attribués à la poussière, et un grand nombre de configurations géométriques possibles, nous reproduisons et ajustons aux observations la distribution d'énergie spectrale venant de l'objet. Cette approche nous a permis de tester le modèle d'unification des AGN qui suppose que les différences observées entre les sous–classes d'objets sont essentiellement dues à leur observation le long de lignes de visée orientées différemment.

Finalement, dans le chapitre 8, nous tirons quelques conclusions et présentons les perspectives futures.

Contents

Chapter 1

An overview of Active Galactic Nuclei

1.1 Introduction

Active Galactic Nuclei (AGN) consist of a compact, extremely bright nucleus, embedded in an otherwise typical galaxy (also known as the host galaxy). Stellar nucleosynthesis fails to explain the amount of energy produced in the bright nucleus, because of its low efficiency ($\ll 1\%$). After more than four decades of observations, spanning from γ–rays to radio–waves, we are now convinced that the most plausible scenario is that of a black hole residing at the center of the host galaxy. The dense central galaxy provides material which accretes onto the black hole, releasing large amounts of gravitational energy with an efficiency of the order of $\sim 10\%$.

Studying AGN in a systematic way is fundamental for understanding their energy production mechanism. Observations in different wavebands probe different regions of the AGN powerhouse and the challenge is to assemble all parts of the puzzle together. There are several types of active galaxies, such as Seyferts, quasars, blazars, etc. Even though these objects look very different at a first glance, observations and theoretical models seem to converge to the conclusion that they are really the same thing, viewed from different directions (the so–called unified model, Antonucci (1993), Urry & Padovani (1995)).

Given the vast diversity of AGN, there is an extensive literature on the AGN phenomenon. For writing this chapter, I mainly consulted two excellent books by Krolik (1999) and Peterson (2001), the lecture notes of H. Netzer (I had the luck to follow his 30–hour course in Chile, in 2004) and his review article "AGN: Basic physics and main components" (to appear in the proceedings of the Santiago 2003 summer school on AGN), and the review article by Risaliti & Elvis (2004).

Band	Range (λ)	Range (ν)	Range (Energy)
Radio	3m$-$ 1 mm	$10^8 - 3 \times 10^{11}$ Hz	$4 \times 10^{-7} - 1.2 \times 10^{-3}$ eV
Submillimeter	$1000 - 150\mu$m	$3 - 20 \times 10^{11}$ Hz	$1.2 - 8.3 \times 10^{-3}$ eV
Far–IR	$150 - 40\mu$m	$2 - 7.5 \times 10^{12}$ Hz	$8.3 - 31 \times 10^{-3}$ eV
Mid–IR	$40 - 3\mu$m	$7.5 - 100 \times 10^{12}$ Hz	$0.031 - 0.41$ eV
Near–IR	$3 - 1\mu$m	$1 - 3 \times 10^{14}$ Hz	$0.41 - 1.25$ eV
Optical	1μm-3000 Å	$3 - 10 \times 10^{14}$ Hz	$1.25 - 4.16$ eV
UV	$3000 - 1200$ Å	$1 - 2.5 \times 10^{15}$ Hz	$4.16 - 10.4$ eV
EUV	$1200 - 12.5$ Å	$2.5 - 240 \times 10^{15}$ Hz	10.4 eV -1.0 keV
X–rays	$12.5 - 0.125$ Å	$2.4 - 240 \times 10^{17}$ Hz	$1 - 100$ keV

Table 1.1: Summary of the spectral ranges covered by the different regions of the electromagnetic spectrum (Risaliti & Elvis 2004).

In the following sections we will describe in detail the components of the active nucleus, the energy production mechanisms and the sub–categories into which AGN are classified. As the physical processes in the interior of AGN are associated with specific regions of the electromagnetic spectrum, we summarize the basic properties of the latter ones in Table 1.1.

1.2 An insight on the active nucleus

The nuclear environment in active galaxies consists of the following components:

- A **supermassive black hole (SMBH)**, with typical size $10^8 - 10^{10}\, M_\odot$, which is undoubtedly the most essential part of the active nucleus. It is still an open question how and when the BH formed (the highest redshift quasars are found at $z \sim 6.5$ (Fan et al. 2003)) and what is its relation to the host galaxy. The basic properties of a non–charged and non–rotating black hole can be described using its gravitational (r_g) and Schwarzschild (r_s) radii:

$$r_g = \frac{GM}{c^2} \text{ with } r_s = 2r_g \tag{1.1}$$

  Depending on the distance of the gas particles with respect to the black hole, they might be subjected to different types of forces and thus are not necessarily accreted onto the BH. The best example of a "non–AGN" case is our own Galactic Center, which does contain a $\sim 10^6 M_\odot$ black hole and accretes material, but does not show any visible AGN activity (Charmandaris (2006) and references therein). The motion, temperature and emission of the dust particles are linked to the conditions in the vicinity of the BH, namely the balance between gravity and the radiation

pressure force, the general processes of spherical accretion and the properties of the accretion disk.

- An **accretion disk**. The most efficient energy production processes are associated with accretion disks. Such systems are formed by gas that sinks into a central plane while conserving most of its angular momentum. Such disks, if dense and thick enough, can provide the necessary mechanism to transfer angular momentum out, and to allow gas in–fall into the vicinity of the BH.

  For a central source of mass M, emitting monochromatic luminosity L_ν, the radiation pressure force acting on gas particles at a distance r is given by:

$$f_{rad} = \frac{N_e \sigma_T}{4\pi r^2 c} L \tag{1.2}$$

  where L is the total luminosity, N_e the electron density and σ_T the Thomson cross section. The gravitational force is:

$$f_g = \frac{GM\mu m_p N_e}{r^2} \tag{1.3}$$

  where μ is the mean molecular weight (mean number of protons and neutrons per electron; about 1.17 for a fully ionized solar composition gas). Spherical accretion (i.e. the case where accretion starts within a large radius from the center of mass, where the gas is at rest) of a fully ionized gas can proceed as long as $f_g > f_{rad}$. From the requirement $f_{rad} = f_g$ we derive the so–called "Eddington luminosity", which corresponds to the upper–limit in the energy production rate needed to keep the system stable:

$$L_{Edd} = \frac{4\pi c G M \mu m_p}{\sigma_T} \simeq 1.5 \times 10^{38} (M/M_\odot)\,\mathrm{erg\,s^{-1}} \tag{1.4}$$

  The accretion rate onto the black hole is defined as $\dot{M} = L/\eta c^2$, where η is the efficiency of converting gravitational potential energy to electromagnetic radiation (of the order of $\simeq 10\%$). The Eddington accretion rate, $\dot{M}_{Edd}$, is the accretion rate required to produce L_{Edd}:

$$\dot{M}_{Edd} = \frac{L_{Edd}}{\eta c^2} \simeq \frac{0.3 M_8}{\eta}\ M_\odot\,\mathrm{yr^{-1}} \tag{1.5}$$

  where M_8 is the mass in unit of $10^8\,M_\odot$. The Eddington time, t_{Edd}, is the typical time associated with this accretion rate and is equal to:

$$t_{Edd} = \frac{M}{\dot{M}_{Edd}} \simeq 4 \times 10^8 \eta\ \mathrm{yr} \tag{1.6}$$

  For $\eta = 0.1$ we obtain an Eddington time of the order of $t_{Edd} = 40$ Myr, which is just a fraction of the life of a galaxy. Thus, even if every galaxy contained an active

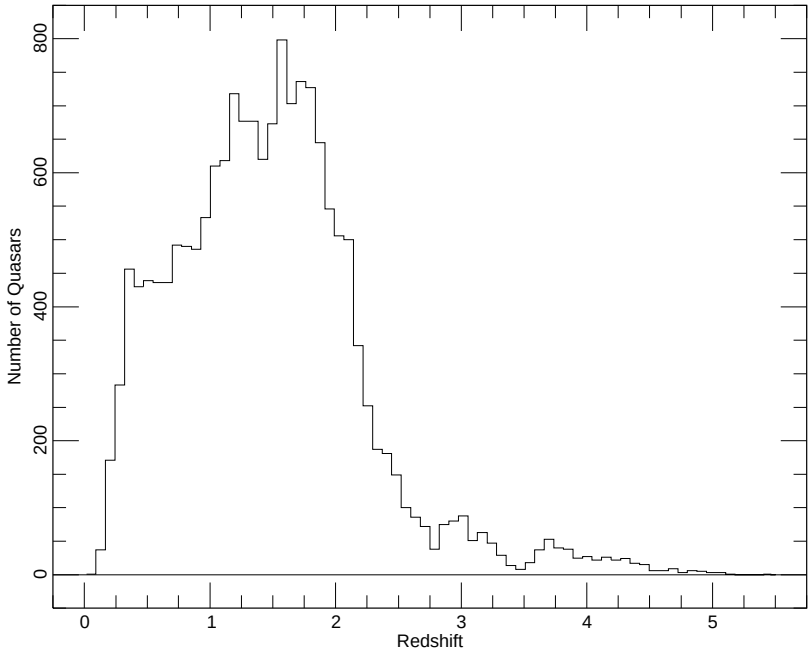

Figure 1.1: Redshift distribution for about 16,700 quasars observed with the Sloan Digital Sky Survey facility (SDSS, Schneider et al. (2003)).

nucleus, its detection would only be possible if the active phase coincided with the epoch of observations.

The redshift distribution for $\sim 16,700$ quasars taken from the Sloan Digital Sky Survey (SDSS, Schneider et al. (2003)) is shown in Fig. 1.1. The peak of the distribution is around $z \simeq 1.5$. For a flat universe, with $H_o = 70\,\mathrm{km/s/Mpc}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$, this redshift corresponds to 4.2 Gyr ($\sim 30\%$ of the age of the Universe). The most distant quasar resides at $z \simeq 6.5$, which is about 0.8 Gyr (6% of the age of the Universe). The question how super–massive black holes formed so early, what is the relation of the AGN to its host galaxy and the possible re–activation of the active nucleus during the life of the galaxy, are among the open issues in the AGN research.

It can be demonstrated that the disk temperature at all radii is:

$$T(r) = \left(\frac{3GM\dot{M}}{8\sigma\pi r^3} \left[1 - (\frac{r_{in}}{r})^{1/2} \right] \right)^{1/4} \tag{1.7}$$

Expressing the radius in units of r_g and normalizing the accretion rate in units of $\dot{M}/\dot{M}_{Edd}$ (i.e. for a given L/L_{Edd}) enable us to compare accretion disks around

various mass BHs. This gives:

$$T(r) \propto M_8^{-1/4} \left[\frac{\dot{M}}{\dot{M}_{Edd}}\right]^{1/4} \left[\frac{r}{r_g}\right]^{-3/4} \tag{1.8}$$

For a thin accretion disk around a BH, with $M_8 = 1$, $T(6r_g) \sim 10^5$K. Such disks emit most of their energy in the UV part of the spectrum. In a similar way, the maximum black body temperature associated with thin accretion disks around stellar–size BH is about 10^7 K, with most of the energy emitted in the X–ray part of the spectrum. Thus, accretion disks around massive BH are much cooler than stellar–size accretion disks.

- The **bipolar flows (jets)**, extended linear structures, which appear to originate from the central compact source. Their morphology suggests that they transport energy and particles from the black hole to the outer part of the nucleus. Their opening angles have a typical range between $\sim 30° - 100°$ and their size spans between 100–300 kpc.

  The ionization cones can be either single–sided or bi–conical structures, with the latter ones sharing a common axis. The cone(s) emerge(s) through obscuring material that blocks our direct view of the nucleus in the sources containing jets.

- The **corona**. Modeling of the energy processes in the active nucleus has shown that AGN accretion disks are "thin", that is, their vertical size is considerably smaller compared to their outer radius. However, thin accretion disks themselves fail to explain AGN emission over the full spectral range.

  We know that a large fraction of the AGN energy is emitted in the X–rays, yet a maximum disk temperature cannot exceed a few $\times 10^5$ Kelvin. Nevertheless, a hot dilute gas in the vicinity of the disk, where the soft disk–emitted photons undergo inverse Compton scattering, could allow them to gain sufficient energy for emitting in the X–rays.

  Accretion through the outer layers of the disk can result in a large energy dissipation and a sharp temperature rise. The outer skin expands and the disk develops a hot corona.

- The **broad–line region (BLR)**, which contains large column density ($\sim 10^{23}$cm^{-2}), high density ($\sim 10^{10}$cm^{-3}) dust clouds, situated at a location where $L/4\pi r^2 \simeq 10^9$ erg s^{-1} cm^{-2} ($\sim$1 pc for a very luminous AGN). As gravity completely dominates over the radiation pressure force, the clouds are bound. Their typical velocities will be of the order 3000 km s^{-1}, confirmed by the width of the observed emission lines in quasar spectra.

  The physical conditions are such that the illuminated surface of the clouds will be highly ionized. The most abundant ions in this part are He III, O IV–VI, C III-IV and the strongest predicted emission lines will be Lyα, C IV λ1549 and O IV λ1035.

Although the density is high, there are still some semi–forbidden lines surviving the conditions in the broad–line region. Strong predicted lines of this type are C III] λ1909 and O III] λ1663. Much of these large column density clouds must be partly neutral, since only X–ray photons can penetrate beyond a hydrogen column of $\sim 10^{22}\mathrm{cm}^{-2}$. These parts will produce strong lines of H I, Mg II and Fe II. The observed equivalent widths of the strongest lines will depend on the emissivity and covering factor, which for specified conditions will be of the order 10–100Å.

- The **narrow–emission line region (NRL)** is characterized by smaller column density ($\sim 10^{20} - 10^{21}\,\mathrm{cm}^{-2}$), medium density ($\sim 10^{4}\,\mathrm{cm}^{-3}$) clouds and is situated at a location where $L/4\pi r^2 \simeq 10^2$ erg $s^{-1}\,\mathrm{cm}^{-2}$ (about 3 kpc, for a very luminous AGN). The physical conditions in this larger region are considerably different from those in the BLR. Because of the smaller column density, radiation pressure force may be important. On average, optically thin gas is more ionized than an optically thick gas. The equivalent width of the emission lines is considerably smaller than that of the broad lines.

  The observed spectrum of this component includes intense forbidden lines, because of the low densities. This shifts the line cooling balance in such a way that the semi–forbidden and permitted lines are relatively weaker. Another group of lines that are predicted to be intense in the innermost part of this region are coronal lines, produced by fine–structure transitions and observed mostly in the infrared.

- The **obscuring torus**, which describes a region of size 1–100 pc located around the BH, with density of $10^4 - 10^6$ and column density varying from $10^{20} - 10^{25}$, or even higher. Far from the center, at 10 or even 100 pc, the gas is very optically thick and has a very low temperature, hence emitting mainly in the IR. Only the hard X–ray radiation can penetrate that far and even those photons are limited to a few Compton depths. The conditions must be similar to those in molecular gas in the galaxy. Such regions are likely to contain large amounts of dust. Their spectral signature is infrared emission and absorption with strong dust features.

  The inner part of such a structure is exposed to the central radiation field. The temperature and ionization of this part depend on the ionization parameter. Under certain conditions, strong, highly–ionized X–ray lines could form at this part of the torus. In particular, low, intermediate and high ionization iron K_α lines are expected to show a clear signature of such a structure. For high column densities, such a structure is a very efficient "X–ray reflector", which will scatter and reflect the incident X–ray continuum radiation. Finally, the inner parts of such a torus will probably serve as a gas reservoir for both X–ray driven winds and accretion onto the central BH.

- **Starburst regions** are normally considered as a different component of AGN. The typical dimension of such regions is several kpc and the typical density $10^0 - 10^3\mathrm{cm}^{-3}$. Such regions contain their own internal energy mechanism, supernova explosions and

	l_8 [pc]	$\theta_{\rm Gpc}$ [mas]	τ_c [yr]	$\tau_{\rm orb}$ [yr]
Accretion disk	$10^{-4} - 10^{-2}$	0.005	$0.001 - 0.1$	$0.2 - 15$
Corona	$10^{-3} - 10^{-2}$	5×10^{-3}	$0.01 - 0.1$	$0.5 - 15$
Broad–line region	$10^{-3} - 1$	0.05	$0.01 - 10$	$0.5 - 15000$
Molecular torus	>1	> 0.5	> 10	> 15000
Narrow–line region	>10	> 5	> 100	> 500000
Jet formation	$>10^{-3}$	$> 5 \times 10^{-4}$	> 0.01	> 0.5
Jet visible in the radio	$>10^{-2}$	> 0.005	> 0.1	> 15

Table 1.2: Typical scales of the main components in an active nucleus (Lobanov & Zensus 2006).

fast stellar winds. The typical X–ray signature is that of a multi–temperature hot plasma, i.e. strong collisionally excited emission lines.

The basic properties of the components of an active nucleus are summarized in Table 1.2 (Lobanov & Zensus 2006). Column 2 presents their corresponding linear scale, considering a black hole mass of $5 \times 10^8\,{\rm M}_\odot$. Column 3 presents the largest angular scale for an AGN located at 1 Gpc. Column 4 gives the rest frame light crossing time, while Column 5 gives the rest frame orbital period, for a circular Keplerian orbit. The main components of an active galactic nucleus (not to scale), as described in the previous section, are shown in Fig. 1.2 (taken from Urry & Padovani (1995)).

1.3 The AGN zoo

Active Galactic Nuclei are separated in two main categories, the "Radio Loud" (RL) and the "Radio Quiet" (RQ) ones. An AGN is considered as radio quiet if (Kellermann et al. 1989): $0.1 < \log(f_{radio}/f_{optical}) < 1$, with the monochromatic fluxes for the radio and the optical computed at 5 GHz (6 cm) and 4400 Å (680 THz), respectively. For the RL objects this ratio is found in the range $10 - 10000$. It has long been recognized (Katgert et al. (1973), Kellermann et al. (1989)) that Radio Loud sources correspond to only $\sim 10\%$ of the total AGN population.

Depending on their luminosity, the members of each category are split in two sub–classes, the low–luminosity and the bright–luminosity. For both Radio Loud and Radio Quiet AGN, the brightest members are called quasars. More analytically:

- For the Radio Quiet AGN, the low–luminosity members are known as Seyfert galaxies. Both the QSOs and Seyferts are split in two sub–classes, the Type–1 and Type–2 objects. Observations over many decades in wavelengths have revealed significant

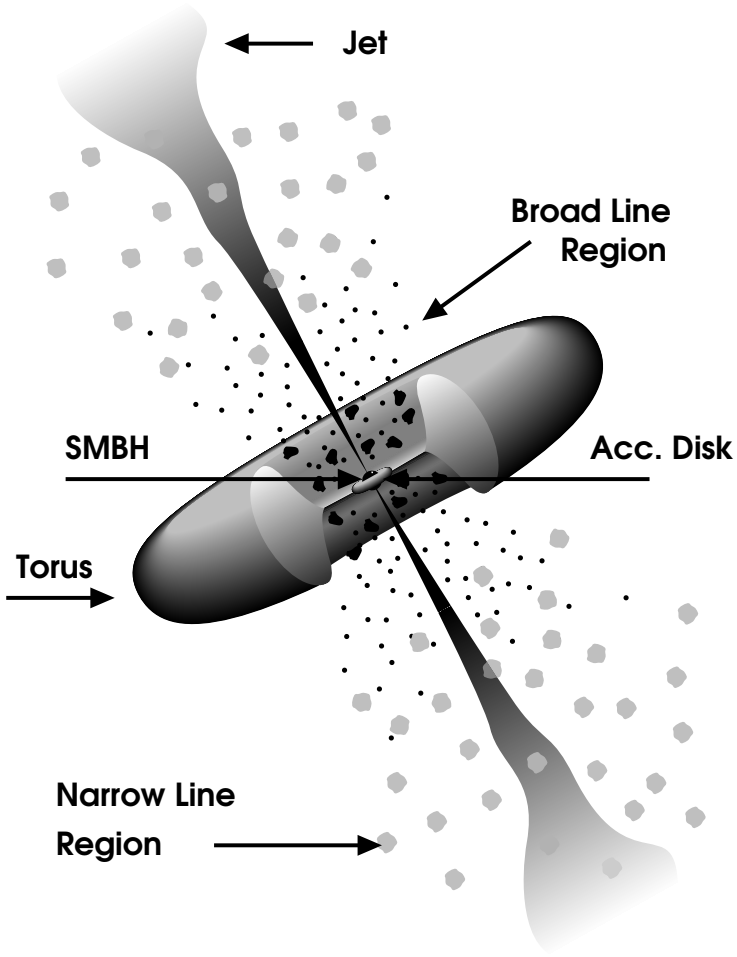

Figure 1.2: Schematic representation of our current view of an active nucleus (taken from Urry & Padovani (1995).

differences (e.g. X–ray emission, optical spectra) in the objects under study and such a distinction was more than obvious. According to the unified model (Antonucci 1993), these differences could be attributed to an orientation effect, with the Type–1/Type–2 populations being a face–on/edge–on version of the same type of objects, simply seen from a different viewing angle. Regarding the QSOs, there is also a third sub–class, the Broad Absorption Line (BAL QSOs, Turnshek (1984)), which corresponds to about $\sim 10\%$ of the RQ QSOs. Their emission lines are characterized by the presence of P–Cygni type absorption features (i.e. broad absorption lines blueward of the line center, in the quasar rest frame) extending anywhere from 2,000 to 60,000 km/s (Weymann et al. 1991).

- For the Radio Loud AGN, the low–luminosity members are known as Radio Galaxies. The latter ones are also split in two sub–classes, in a way similar to the Seyfert galaxies, the Broad Line and the Narrow Line Radio Galaxies (BLRG and NLRG,

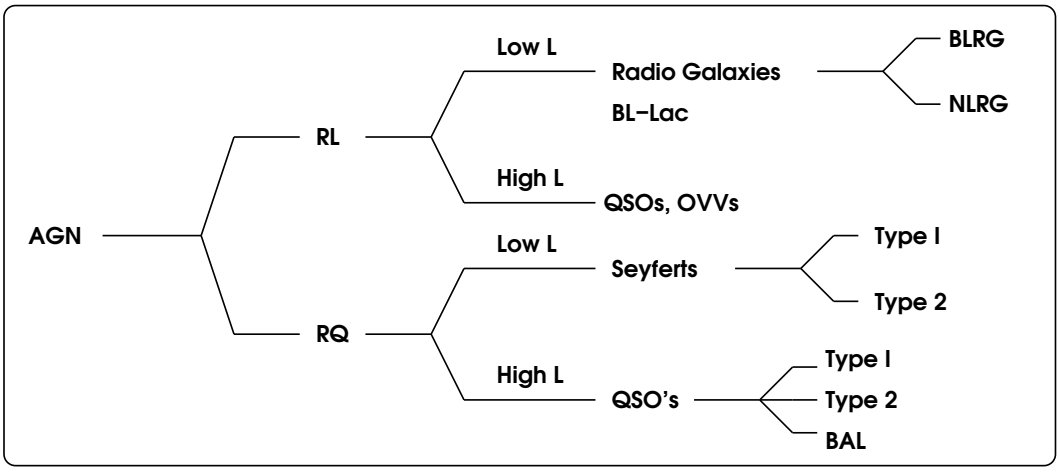

Figure 1.3: Classification of Active Galactic Nuclei (taken from Borguet (2006)).

respectively). BL–Lac and OVVs (Optically Violent Variable) are both face–on versions of radio sources, with the former corresponding to the low–luminosity and the latter to the most luminous sources.

A classification of the different sub–classes of active galactic nuclei is presented in Fig. 1.3. Although it is out of the scope of this introduction to describe the properties of each sub–class, we will briefly present the ones to which we will refer in subsequent sections. A full section will be dedicated to QSOs, since a significant part of this work is closely related to them.

1.3.1 Seyfert Galaxies

Seyfert galaxies (Seyfert 1943) are spiral galaxies with bright, tiny cores that fluctuate in brightness. The host galaxy, in which the active nucleus resides, is clearly visible. In terms of optical magnitudes, Seyferts are defined as the active galaxies with $M_B > -21.5 + 5 \log h_{100}$, where M_B is the absolute magnitude in the B–band (Vega system) and h_{100} the Hubble constant in units of 100 km s^{-1}Mpc^{-1}. Spectroscopic observations have revealed strong, high–ionization emission lines in their spectra. Seyfert galaxies do not have radio lobes and most of them are powerful sources of infrared radiation. In addition, some emit intensely in the radio, X–ray, and gamma ray regimes. Approximately 2% of all spiral galaxies are Seyfert galaxies.

There are two kinds of Seyfert galaxies, called Type–1 and Type–2. The nuclei of both types show emission lines, which is evidence of highly excited gas. The difference, however, arises in the shape of the emission lines. Type–1 Seyferts have very broad emission lines, suggesting gas velocities over 1000 km/sec. The emission lines of Type–2 Seyferts are much narrower, suggesting that the gas in these galaxies is moving much more slowly.

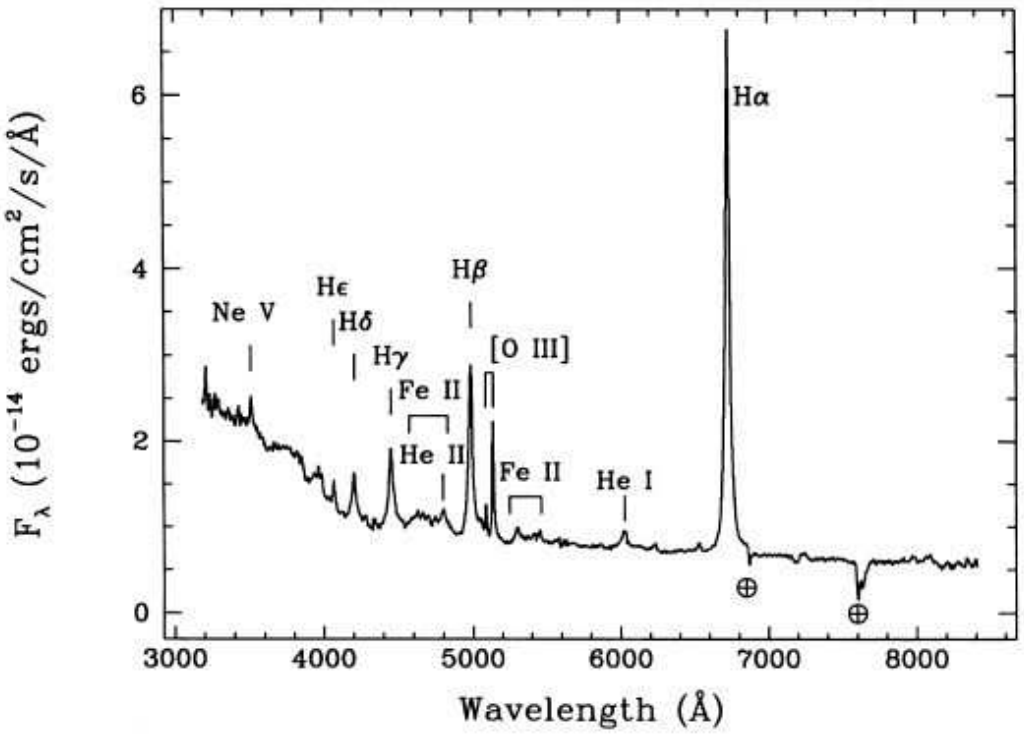

Figure 1.4: Optical spectrum of Markarian 335, a Seyfert–1 galaxy found at redshift $z = 0.026$. The intrinsic emission features are labeled. Two atmospheric absorption features are marked with an Earth symbol (Zheng et al. 1995).

One possible scenario for explaining the differences in the emission lines between the two categories is that of the viewing angle: in the case of Type–1 Seyferts we have an edge on view to the active nucleus, with the very hot, fast–moving gas generating the broad lines we see in Seyfert–1 spectra. Type–2 Seyferts are probably viewed almost through the torus and the light from the central black hole is significantly blocked. The narrower emission lines seen in Seyfert–2 spectra are generated by a slower–moving gas, farther from the black hole.

The optical spectrum of a typical Seyfert–1 galaxy, Markarian 335, is presented in Fig. 1.4 (Zheng et al. 1995). A summary of the most important emission lines, with their corresponding FWHM, is presented in Table 1.3.

The rise in the continuum for wavelengths shortward of 4000 Å is disentangled in two components:

- The "big blue bump", which has a thermal origin[1] and is associated with the physical processes taking place in the accretion disk. It starts at $\lambda \sim$4000 Å and possibly

[1]By thermal emission we refer to a particle velocity distribution that can be described by the Maxwell–Boltzman equation. An example of a non–thermal process is synchrotron radiation.

Line	λ (Å)	FWHM (km s^{-1})
Hδ	4101.73	1732 ± 145
Hγ	4340.46	1560 ± 450
O [III]	4958.90	598 ± 14
O [III]	5006.80	598 ± 14
He I	5875.70	3283 ± 45
Hα	6562.80	1006 ± 21

Table 1.3: FWHM for some of the most important emmision features in the optical spectrum of Markarian 335, shown in Fig. 1.4 (Zheng et al. 1995).

extends beyond EUV wavelengths.

- The "small blue bump", which extends between $2000 - 4000$ Å and is attributed to a combination of Balmer continuum emission and blends with Fe II emission lines produced in the broad–line region (Wills et al. 1985).

The presence of the blue bump consists a solid proof that in Type–1 AGN we have a direct view to the accretion disk. On the contrary, its absence in Type–2 spectra implies that the inner part of the active nucleus is hidden. NGC 1068 is a prototype Seyfert–2 galaxy located at $z = 0.0038$. In its optical spectrum, shown in Fig. 1.5, the lack of the blue bump is clearly visible (Risaliti & Elvis 2004).

Although the viewing angle hypothesis explains the differences in the emission lines and the presence (or lack) of the blue bump, there are still other features (e.g. the continuum shape, the slightly fainter magnitudes of Type–2 Seyferts compared to Seyfert–1), that the orientation scenario fails to explain. Hence, there is still a lot of work to be done before the unification scheme explains the diversity of features observed in AGN.

Osterbrock (1981) suggested to introduce the 1.5, 1.8 and 1.9 notation in the Seyfert classification, for differentiating Seyfert galaxies with small variations in their optical spectrum (larger number classification corresponds to weaker broad—line components relative to the narrow lines).

1.3.2 BL Lac Objects and Optically Violent Variables

Variability is a common property in active galactic nuclei. However, a special type of AGN, the optically violent variables ones (OVVs), show short time–scale variations (of the order of a day) with Δm in the visible larger than 0.1 mag. Additionally, their polarization is much higher than that of typical AGN, also varying in magnitude and position angle.

A sub–class of OVVs are the BL Lacertae objects (or BL Lac), named after the

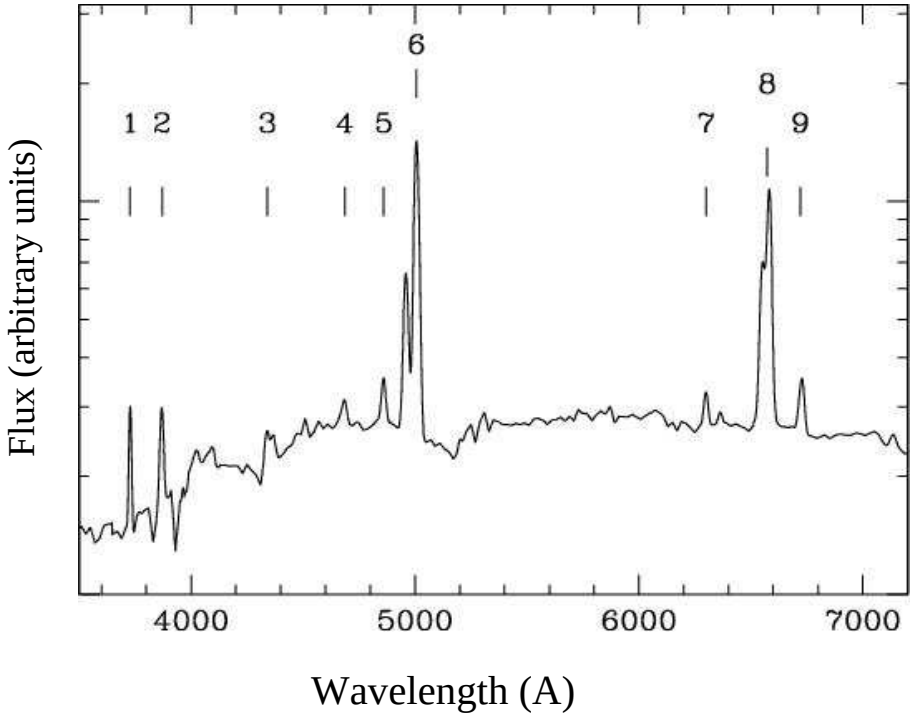

Figure 1.5: Optical spectrum of the Seyfert–2 galaxy NGC 1068. Numbers 1–9 designate the following emission lines: 1: [OII]λ3727Å 2: [NeIII]λ 3869Å 3: Hγ 4: HeII λ4687Å 5: Hβ 6: [OIII]λ5007Å 7: [OI]λ6300Å 8: Hα+[NII]λ6585Å 9: [SII]λ6732Å. The typical line widths are between 300 and 800 km s^{-1} (Risaliti & Elvis 2004).

prototype BL Lacertae, initially identified as a highly variable star. The lack of strong emission or absorption lines is one of the main properties of BL Lac objects.

Fig. 1.6 presents the results of a four–band, 20–year photometric monitoring of the BL Lac 2200+420, where the strong variability of the object (spanning several magnitudes) is clearly visible (Fan et al. 1998).

OVVs and BL Lac objects are collectively known as blazars. The main difference between blazars and Seyfert galaxies or quasars is that a blazar has one of its relativistic jets pointed toward the Earth, so that what we observe is primarily emission from the jet region. Synchrotron radiation, generated in the jet, is detected in the radio–waves, thus making blazars very bright radio sources.

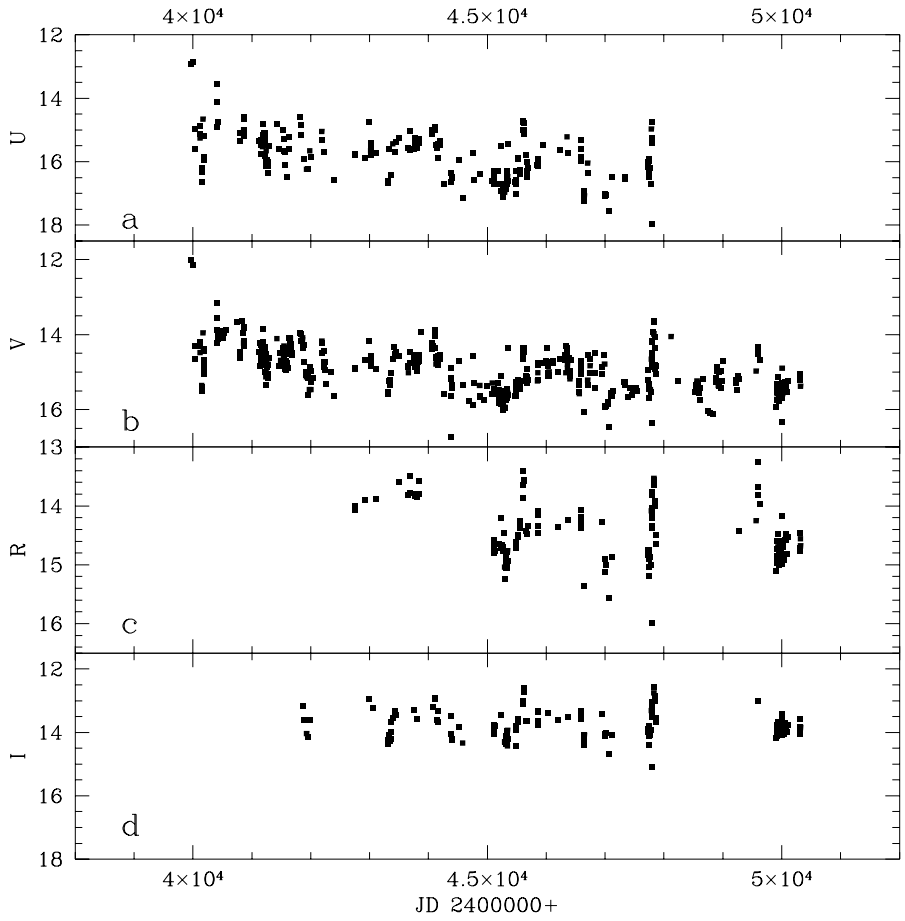

Figure 1.6: Results of the photometric monitoring of BL Lac 2200+420, spanning $\sim$ 20 years of observations (Fan et al. (1998) and references therein). The light–curves demonstrate a peak–to–peak variability for the U, V, R and I photometric bands of 5.12, 4.73, 2.59 and 2.54 mag, respectively.

1.4 QUAsi–StellAR objects

Quasars were first discovered in the radio waves in the early 1960s (Schmidt 1963). Some of the optical counterparts of the strongest radio sources were galaxies, but others appeared to be stellar–like, and this is how the acronym QUASAR (QUAsi StellAR radio sources) came from. Once the first redshifts were measured, an order of magnitude larger than those of the known Seyfert galaxies and among the largest ever measured at that time, the significance of quasars in cosmology was immediately recognized. Their extreme luminosities not only meant that quasars are unique tools for probing the early Universe, but also implied energy production mechanism of unprecedented efficiency. In terms of optical magnitudes, quasars are active galactic nuclei with $\mathrm{M}_B < -21.5 + 5 \log h_{100}$, where

M_B is the absolute magnitude in the B–band (Vega system) and h_{100} the Hubble constant in unit of 100 km $s^{-1}Mpc^{-1}$. Hubble Space Telescope images of the very first QSO ever–discovered (3C 273), revealing details about its environment, are shown in Fig. 1.7.

1.4.1 The importance of quasars in cosmology

The main open issues concerning quasars, both in terms of physical processes in their interior/close environment and as cosmological probes, are the following:

- Understand the AGN energy production mechanism and its relation to the various sub–classes of active galactic nuclei (unified model).
- Map the space density of quasars as a function of redshift (luminosity function).
- Study the quasar clustering and their relation with galaxy clusters.
- Study the relation between quasars and their host galaxies (understand their origin/evolution and their relation to starbursts).
- Study the inter–galactic medium through the absorption lines in quasar spectra (Lyman–α forest).
- Use the gravitationally lensed quasars as tools for studying the matter distribution and the geometry of the universe.

1.4.2 Quasar Spectral Energy Distribution

Quasars emit their energy over 20 decades in the electromagnetic spectrum. Unlike spectra of stars or galaxies, AGN emission can not be described using the black–body radiation model of a given temperature, nor as a composite over a small range in temperatures. The spectral energy distribution of a quasar continuum can be described by a power–law of the form:

$$F_\nu = C\,\nu^{-\alpha} \tag{1.9}$$

where α is called the power–law index, C is a constant and F_ν is the monochromatic flux, i.e. the flux per unit frequency interval (usually measured in $erg\,s^{-1}cm^{-2}\,Hz^{-1}$). A typical range for the power–law index when describing the QSO SEDs at large energy intervals is $0 \leq \alpha \leq 1$, but different values of α are found for different spectral ranges. According

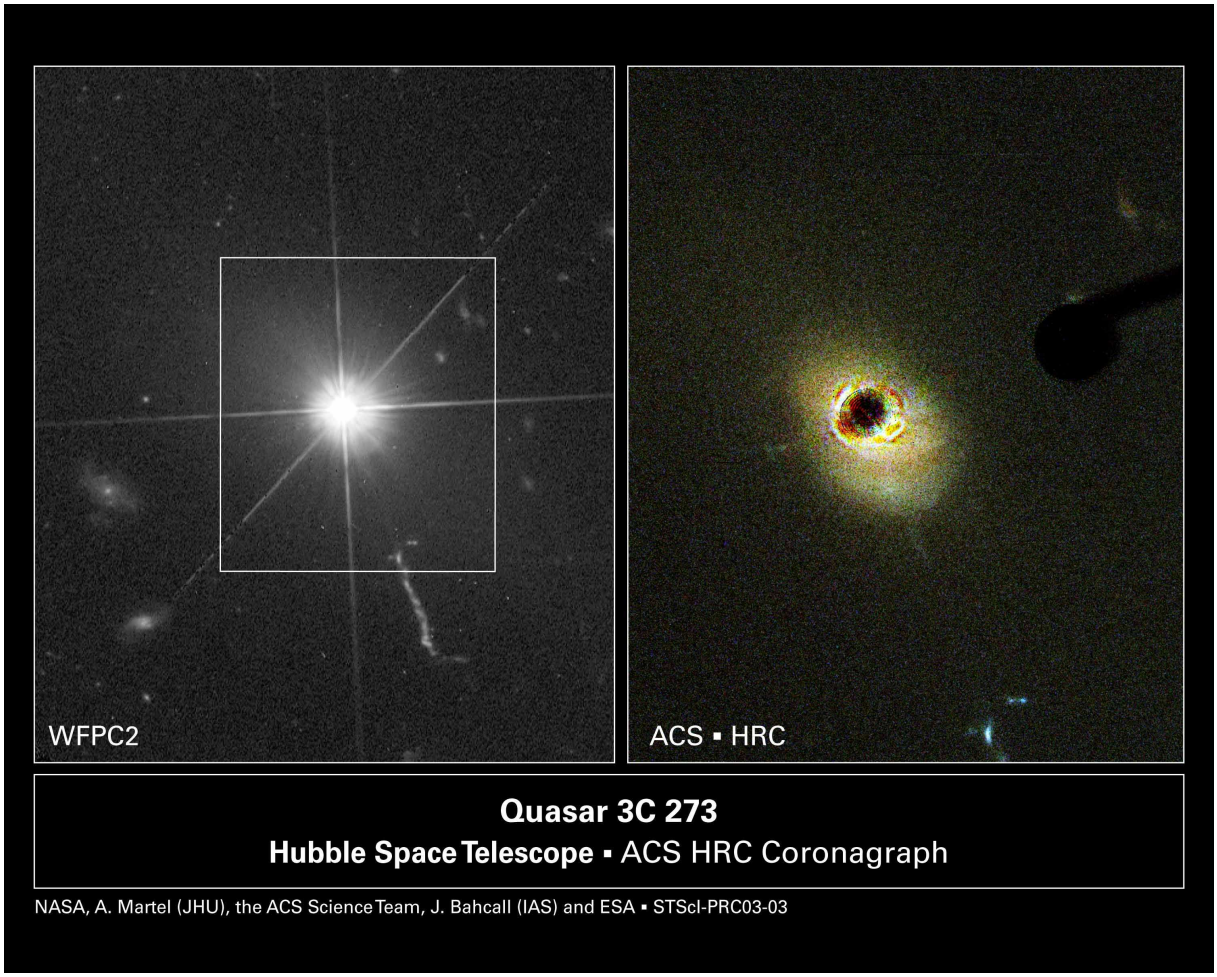

Figure 1.7: HST images of the most famous quasar, 3C 273 (Schmidt 1963), taken with the Wide Field Planetary Camera 2 (WFPC 2, left panel) and the Advanced Camera for Surveys (ACS, right panel). The WFPC 2 image shows the bright quasar but little else, with the diffraction spikes demonstrating the QSO's point–like morphology. Using the coronographic technique with ACS the quasar's light can be blocked, revealing details about its environment, such as a spiral plume wound around the quasar, a red dust lane and a blue arc and clump in the path of the jet blasted from the quasar. These details were never seen before. Previously known clumps of hot gas and the inner blue optical jet are now resolved more clearly. (Credit for WFPC 2 image: NASA and J. Bahcall (IAS); Credit for ACS image: NASA, A. Martel (JHU), H. Ford (JHU), M. Clampin (STScI), G. Hartig (STScI), G. Illingworth (UCO–Lick Observatory), the ACS Science Team and ESA.)

to Eq. 1.9, the total power received over a frequency range $\nu_1 - \nu_2$ is:

$$\begin{aligned} P(\nu_1, \nu_2) &= \int_{\nu_1}^{\nu_2} F_\nu d\nu = C \int_{\nu_1}^{\nu_2} \nu^{-\alpha} d\nu \\ &= \frac{C}{1-\alpha}(\nu_2^{1-\alpha} - \nu_1^{1-\alpha}) \quad (\alpha \neq 1) \\ &= C \ln\left(\frac{\nu_2}{\nu_1}\right) \end{aligned} \tag{1.10}$$

The case $\alpha = 0$ represents what we call a "flat spectrum" in a plot of monochromatic flux versus frequency, meaning that the source emits the same energy per unit frequency interval. The case $\alpha = 1$ corresponds to a source with the same energy per unit *logarithmic* frequency interval. By convention, we now plot the broad band SED of quasars as $\log(\nu F_\nu)$ as a function of $\log \nu$. In this case the power–law distribution becomes $\nu F_\nu \propto \nu^{1-\alpha}$, so the case $\alpha = 1$ represents a horizontal line in the log–log plot and the $\alpha = 0$ spectrum has a positive slope as the frequency increases.

A closer look at a quasar spectral energy distribution reveals several features, such as emission lines, bumps and dips. Therefore, depending on the part of the SED we are looking at, the power–law index –expressed by Eq. 1.10– may take a different value. The observed features are associated with the physical processes taking place in the quasar's active nucleus. Depending on the case, these mechanisms are related to thermal, or non–thermal emission. The spectral energy distribution of a quasar (and any active galactic nucleus, in general) consists of the following features:

- A soft X–ray excess, an emission feature observed below $\sim 1\,\mathrm{keV}$.
- An X–ray continuum, which covers energies from $\sim 1 - 100\,\mathrm{keV}$.
- The big blue bump, which covers the wavelength range from $10\,\text{Å}{-}4000\,\text{Å}$.
- The near–infrared inflection, between $1\,\mu$m–$2\,\mu$m.
- The infrared bump, which extends from 1 to $100\,\mu$m.
- The submillimeter break, which marks a sharp drop in the AGN emission.

The spectral energy distribution (continuum emission) of a typical AGN, where all the above features are clearly visible, is shown in Fig. 1.8 (Manners 2002). The gap between $\sim 100\,\text{Å}$ and $912\,\text{Å}$ is due to the lack of observational data (that would allow to investigate on the SED evolution) since our galaxy is practically opaque in the specific wavelength range, due to absorption by neutral hydrogen.

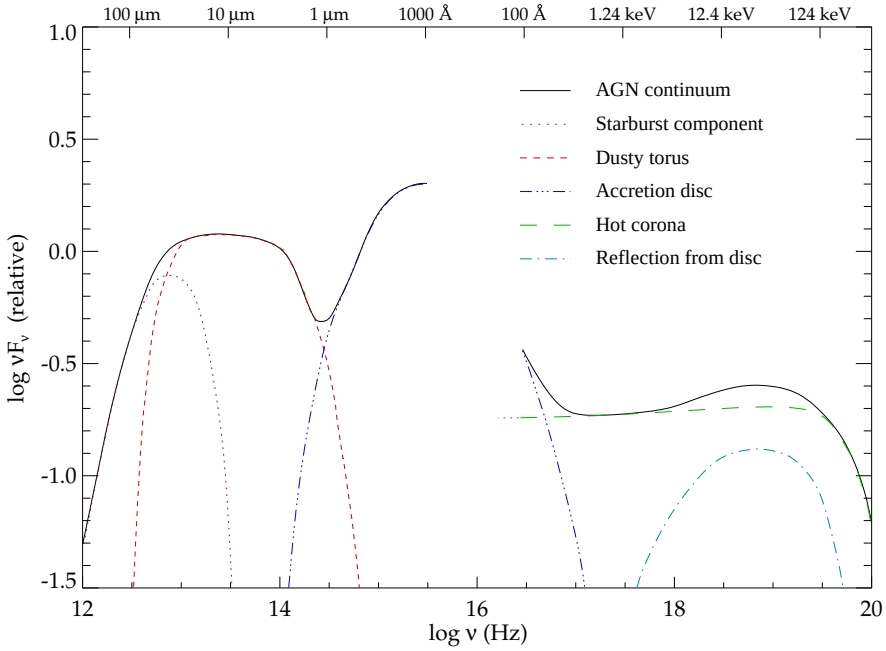

Figure 1.8: Spectral Energy Distribution of a typical active galactic nucleus (taken from Manners (2002)). The contribution of the different components is marked using different line types.

1.4.2.1 X–rays

The fact that all AGNs are luminous X–ray sources is probably the only characteristic that active galaxies have in common. The X–ray emission requires energy production mechanisms rarely found in the Universe, thus observing in the X–rays is a very efficient way for identifying AGNs (and more specifically, quasars). Thanks to their extreme energies, electromagnetic radiation emitted in X–rays is practically unaffected by line–of–sight absorbers between the X–ray emitting source and Earth, thus giving an insight to the physical processes taking place in the active nucleus, free of any observational selection effects.

In section 1.4.2 we saw that quasar SEDs are fitted by power–laws, where Eq. 1.9 describes the emitted flux per unit frequency interval. In X–ray astronomy the convention is to fit SEDs with power–laws expressed in photons per second, per energy interval, since this is closer to the observed quantity, counts per second. Following this concept, X–ray power–law fits are of the form:

$$P_E(\text{photons s}^{-1}\ \text{keV}^{-1}) \propto E^{-\Gamma} \propto \nu^{-\Gamma} \tag{1.11}$$

Since the energy flux F_ν is proportional to P_E, we have:

$$\begin{aligned} F_\nu &\propto P_E(\text{photons s}^{-1}\ \text{keV}^{-1}) \times h\,\nu\,(\text{ergs/photon}) \\ &\propto \nu^{-\Gamma+1} \propto \nu^{-\alpha} \end{aligned} \tag{1.12}$$

where the last analogy comes from Eq. 1.9, where we have defined the index α. In order to distinguish the latter from $\Gamma(=\alpha+1)$, the index α is generally called energy, or spectral index, while Γ is known as photon index.

The X–ray emission from quasars extends from the Galactic absorption cut–off at $\sim$ 0.1 keV up to $\sim$ 300 keV. It can be disentangled in two components, the *primary emission* and the emission due to *reflection*.

1.4.2.2 Optical/UV

The UV to optical emission of quasars is characterized by the "big blue bump", as the maximum of the emission is usually found in this interval (see also paragraph 1.3.1). The peak energy is around the Lyman edge ($\lambda = 1216$ Å) and the spectrum can be well approximated with a power law both at lower and higher frequencies. The optical–UV emission originates from an optically thick and geometrically thin accretion disk, where fast–moving ionized clouds produce a large number of broad and blended emission lines.

A quasar spectrum, spanning from the UV to the optical, is displayed in Fig. 1.9. The part redward of Lyα is a composite spectrum generated from some 2200 spectra of SDSS quasars (Vanden Berk et al. 2001). The slope of the optical continuum varies between $-1 < \alpha < 0$ (considering that $F_\nu \propto \nu^\alpha$), with a median slope of $\alpha = -0.46$. The continuum blueward of the Lyman edge comes from HST/FOS observations of some 200 quasar spectra (Zheng et al. 1997). At about ~ 1050 Å there is a break in the optical continuum slope, with $\alpha \approx -1.8$ in the far–UV. In the same figure, the "small blue bump" (see section 1.3.1) is also indicated.

1.4.2.3 Infrared

The integrated IR emission of quasars (1–150 μm) corresponds to $\sim$ 30% of their bolometric luminosity, with values in individual objects ranging from $\sim$ 15%–$\sim$ 50%. The spectral shape is characterized by:

- A well defined minimum, placed at $1-2\,\mu$m (Neugebauer et al. 1987), associated with thermal emission mechanisms. A primary UV–optical thermal continuum emission heats the torus which absorbs, scatters and re–emits the incoming radiation. The torus mainly consists of two astronomical dust components, silicate and graphite grains. These grains can not survive the strong radiation field, unless they are far

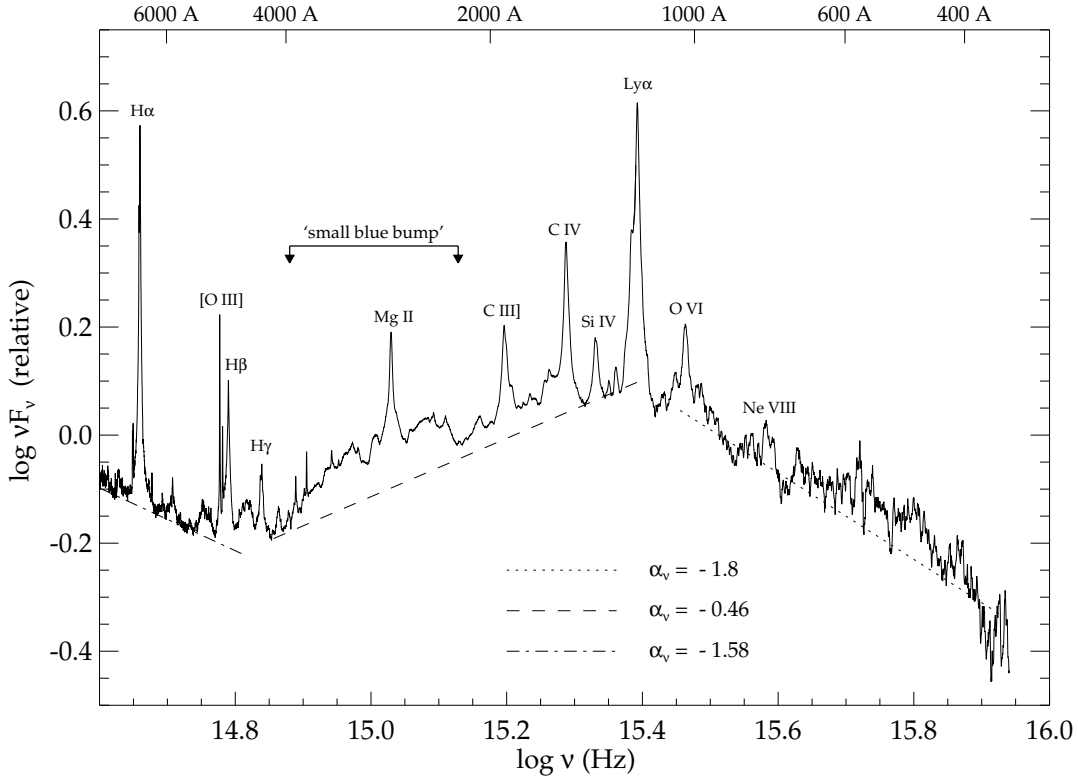

Figure 1.9: QSO composite spectrum spanning from UV (Zheng et al. 1997) to optical (Vanden Berk et al. 2001) wavelengths (see text for more details). Figure taken from Manners (2002).

away from the central source ($\sim 0.1 - 1$ pc), where the temperature has dropped below their sublimation[2] temperature ($\sim$ 1450 and $\sim$ 1700 K for the silicate and graphite grains, respectively). Hence the IR inflection at $\sim 1.5\,\mu$m represents the meeting point between the Wien tail of a $\sim$ 2000 K black body (describing the accretion disk thermal emission) and the hot–dust torus emission.

- The "IR bump", typically at 10–30 μm, due to the thermal emission of dust at a temperature of about $\sim$ 100 K. The width of the bump depends (a) on the physical properties of the torus, namely its optical depth, and (b) its geometrical properties, namely the ratio of the outer radius of the torus over its height (Granato & Danese 1994).

- A steep decline ($a < -3$) at large wavelengths, typical of the low energy spectrum of the gray emitter.

[2]The sublimation temperature is the maximum temperature at which dust of a certain composition can survive to the energy emitted by a central radiation field, without evaporating.

1.4.2.4 Submillimeter

The submillimeter emission in quasars is due to optically thin, cool dust. Consequently, in the context of the unified scheme, the quasar submillimeter emission is orientation–dependent. Nevertheless, intrinsic processes in the QSO host galaxy, namely stars in starbursts, will also contribute in the sub–mm emission. Therefore, it is essential to study the properties of AGNs in this wavelength range for disentangling the origin of the emission.

The quasar SED experiences a significant drop in the submillimeter band, known as the "millimeter break". The size of the drop depends significantly on the object. For Radio Loud QSOs, this drop is of the order of two decades, while for the Radio Quiet sources it is more than five, or even six decades (Kellermann et al. (1989), Elvis et al. (1994)). The striking difference in the radio emission between Radio Loud and Radio Quiet AGNs is demonstrated in Fig. 1.10 (Elvis et al. 1994).

The Atacama Large Millimeter Array (ALMA)[3], with a spatial resolution $\sim 10-100\,\mathrm{pc}$ and mJy sensitivity, will be able to probe the structure of the molecular tori, the starburst activity and study the kinematics in the vicinity of the black hole[4].

1.4.2.5 Radio

The radio emission in Radio Loud quasars is mainly associated with relativistic jets, formed as electrons are accelerating under the influence of the magnetic fields which develop in the vicinity of the black hole.

The VLBI and the VLA[5] facilities, with observations at $4.997\,\mathrm{GHz}$ ($\lambda = 6\mathrm{cm}$), have resolved the central engine of the Radio Loud AGNs, revealing both small–scale (i.e. sub–parsec, (Lal et al. 2004)) and long–scale (kiloparsec, (Kharb et al. 2006)) structures. These observations led to the conclusion that these radio structures do not always have the same orientation (Baum et al. 1993), implying that more complex processes are responsible for the radio emission in RL AGNs (Kukula et al. 1996). A representative example of the multi–scale features observed in Radio Loud AGNs is illustrated in Fig. 1.11, with VLA, VLB and MERLIN[6] observations of Markarian 6, an early–type (S0a) galaxy containing a compact, variable Seyfert nucleus (Kharb et al. 2006).

[3]ALMA is an array of some tens of 12–m submillimeter antennas, with baselines of several kilometers. It will be operational in about 2010.

[4]http://www.eso.org/projects/alma/science/cosmo.html

[5]VLA: Very Long Base Line Interferometer, VLA: Very Large Array

[6]Multiple Element Radio Linked Interferometer Network

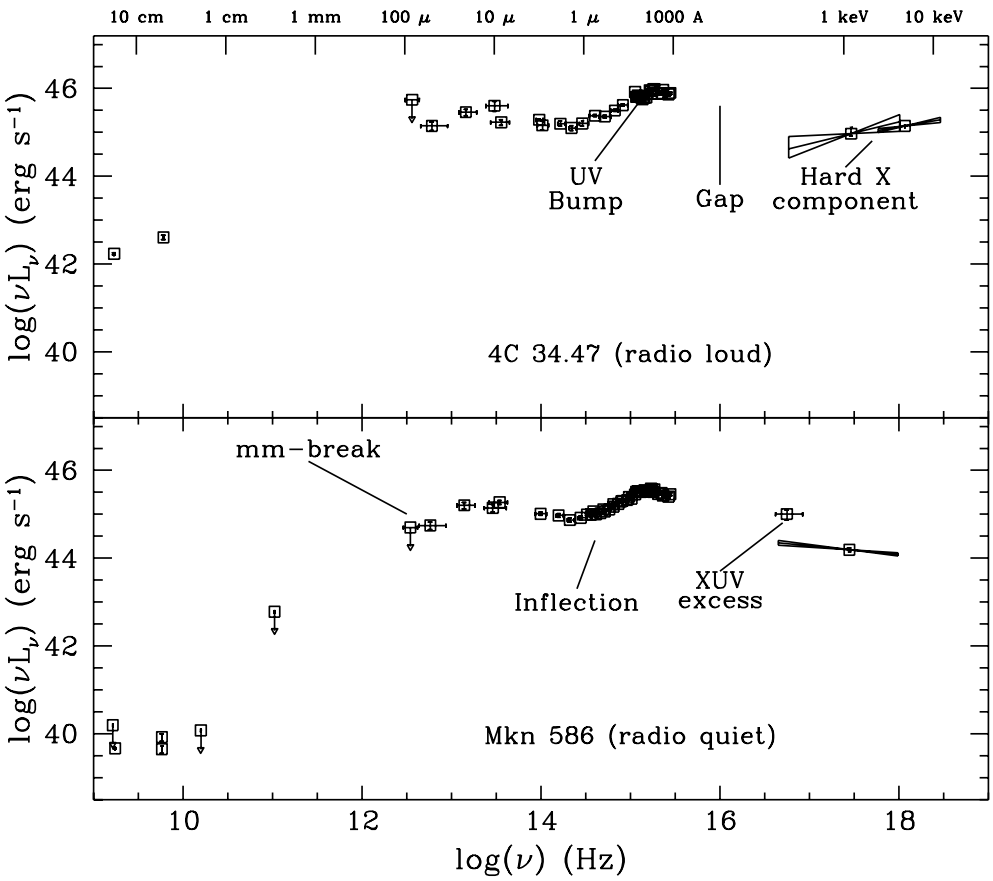

Figure 1.10: Rest–frame SED for a Radio Loud (upper panel) and a Radio Quiet (lower panel) quasar. The X–ray to far–IR part of the SEDs is quite similar for both objects, with the basic features (i.e. blue bump, IR inflection, IR bump) being easily identified. At about $\sim 100\,\mu$m the RQ quasar SED experiences a much steeper decline, known as the millimeter break. (Elvis et al. 1994)

1.5 Type–II Quasars and the unified model

Type–2 quasars are the luminous analogs of Seyfert–2 galaxy nuclei (see also paragraph 1.3.1). The former were predicted on the basis of the unification scheme of active galactic nuclei. According to this picture, Type–1 and Type–2 AGN are essentially the same objects viewed at different inclination angles. Type–2 quasars are thought to be powerful objects whose central regions, emitting continuum light and broad emission lines, are shielded from the observer's line–of–sight by large amounts of gas and dust confined in a torus (Antonucci (1993), Urry & Padovani (1995)). By analogy with Seyfert–2 type galaxies, Type–2 quasars should therefore only display in their spectra narrow emission lines with high-ionization ratios. According to this orientation-dependent obscuration, Type–2 quasars should not be easy to find.

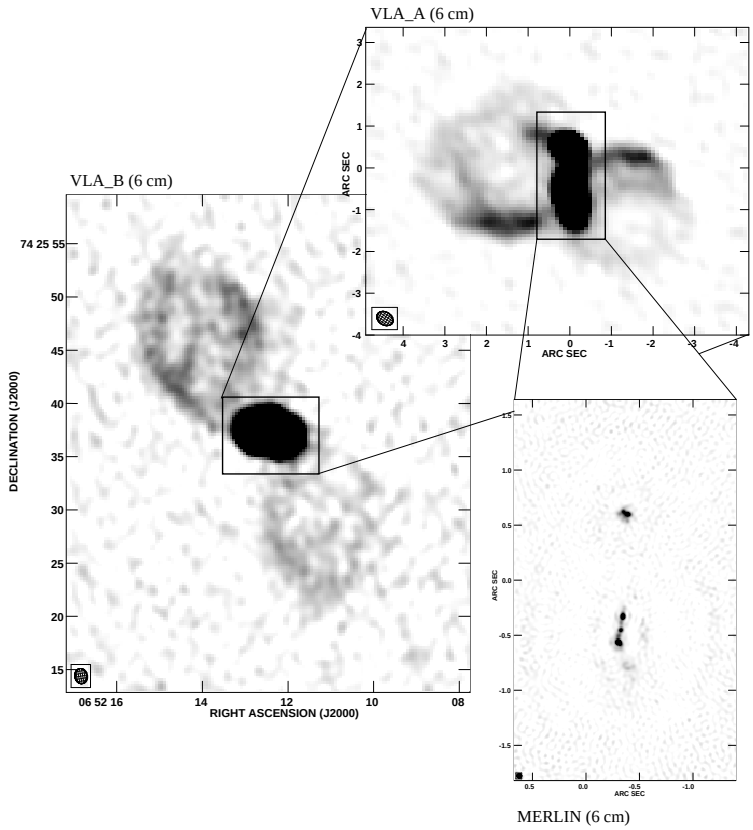

Figure 1.11: Mosaic image of the radio emission at $\lambda = 6$cm (4.977 GHz) of Markarian 6, an active galaxy classified as Seyfert 1.5 (Osterbrock & Koski 1976). The radio observations have revealed structures at three scales. First, there is a radio jet of extent 1 kpc ($3''$) extending in the north-south direction, at the active galactic nucleus. Second, there are a pair of inner bubbles of size 1.5×1.5 kpc ($4'' \times 4''$) each, extending east-west and nearly perpendicular to the central jets. And third, there are a pair of outer bubbles of dimensions 4.5×7.5 kpc ($12'' \times 20''$) each, extending northeast-southwest. (The VLA images are from Kharb et al. (2006) and the MERLIN from Kukula et al. (1996)).

Radio observations of AGN often show powerful jets, streams of particles coming from the central source. Charged particles are accelerated to nearly the speed of light in these jets. In the unified view of active galaxies, high–energy quasars are being viewed with the jet pointed towards us which allows us to see the resulting energetic radiation. With Seyfert galaxies, we are viewing from the side and do not see the very high-energy radiation which is traveling down the jet.

The present research trend is towards a unified model for the central engine with different "ambient" conditions. However some AGN with strong radio emission (radio–loud: radiogalaxies and blazars) produce relativistic jets powering the extended radio lobes, while in others (radio quiet: Seyfert galaxies, QSOs) this phenomenon is not observed. The distinction between radio loud and radio quiet AGN may represent a fundamental

difference within AGN.

It is now largely accepted that absorbed AGN play a significant role in the production of the Cosmic X–ray Background (CXB) above 2 keV and should be the site where a large fraction of the energy density of the universe is generated. However, many important questions are still unsolved. For example, the relationship between narrow and broad line AGN in the optical regime and absorbed and unabsorbed AGN in the X–ray regime is still unclear, and doubts on the very existence of high–luminosity Type–2 AGN are often cast.

Previous searches for Type–2 quasar candidates have reported a few tens of objects at a variety of wavelengths (optical, infrared, radio) but the multi–wavelength properties of these heterogeneous samples do not always show clear trends. For instance, narrow line radio galaxies have been known since a few decades and powerful central engines have been proposed to explain their narrow line luminosities (McCarthy 1993). However, significant obscuration in the X–ray has only been found for some (Carilli et al. 2002) but not for all narrow line radio galaxies (Barcons et al. 2003).

Following the discovery that the X–ray background is largely due to obscured AGN at redshifts near $z = 1$ (Barger et al. 2003), the interest in Type–2 quasars was revived and several candidates were identified. In the X–rays, these objects are expected to be highly luminous ($> 10^{44}$ erg s^{-1}) and absorbed by column densities $> 10^{22}$ cm^{-2}. However, a straight connection between X–ray selected obscured AGN and optically identified narrow emission line AGN remains unclear as some sources, which are classified as Type–1 AGN in the optical, turn out to be absorbed in the X–rays (Matt 2002). To further complicate matters, most Compton thick sources ($N_H > 10^{23.5}$ cm^{-2}) are very faint in the [2-10] keV band and therefore are totally missed in X–ray selected samples. Furthermore, more than one third of Seyfert–2 galaxies have $N_H > 10^{24}$ cm^{-2} (Maiolino et al. 1998) but this fraction is totally unknown for objects with quasar luminosities.

Type–2 AGN are not easily identified in optical surveys. Unlike Type–1, Type–2 quasars lack the UV–excess signature and distinctive colors. For a given intrinsic luminosity, Type–2 AGN are also much fainter in optical than their Type–1 analogs and therefore much deeper surveys are mandatory in order to find them. Recently, Zakamska et al. (2003) have reported a sample of nearly 300 Type–2 AGN candidates based upon the identification of narrow, high-ionization emission lines detected in SDSS spectra redshifted in the range $z = 0.3 - 0.8$. They used the [OIII] 5007 emission line as a tracer of the nuclear activity (cf. X–ray luminosity), assuming that this nebular emission arises from an unobscured, extended narrow line region. They report that up to 50% of their objects qualify as Type–2 quasars, but their X–ray properties are still mostly unknown.

As a conclusion, searches for type II quasars at different wavelengths do not retrieve the same population of objects and it is not clear at the moment how different selection techniques relate to each other. Also questionable is the use of the [OIII] luminosity (Zakamska et al. 2003) or the X–ray-to-optical ratio (Fiore et al. 2003) to infer the X–ray

luminosity of the Type–2 AGN candidates. It implies that the contribution from obscured AGN to the black hole census, their luminosity function and their fraction to the whole AGN population remain poorly known. It is not clear either whether the unification scheme first proposed for low luminosity Seyfert galaxies equally applies to quasar luminosities.

The relation between the AGN taxonomy and the viewing angle, that could possibly explain the diversity of features observed in active galaxies in the context of the unification scheme, is illustrated in Fig. 1.12.

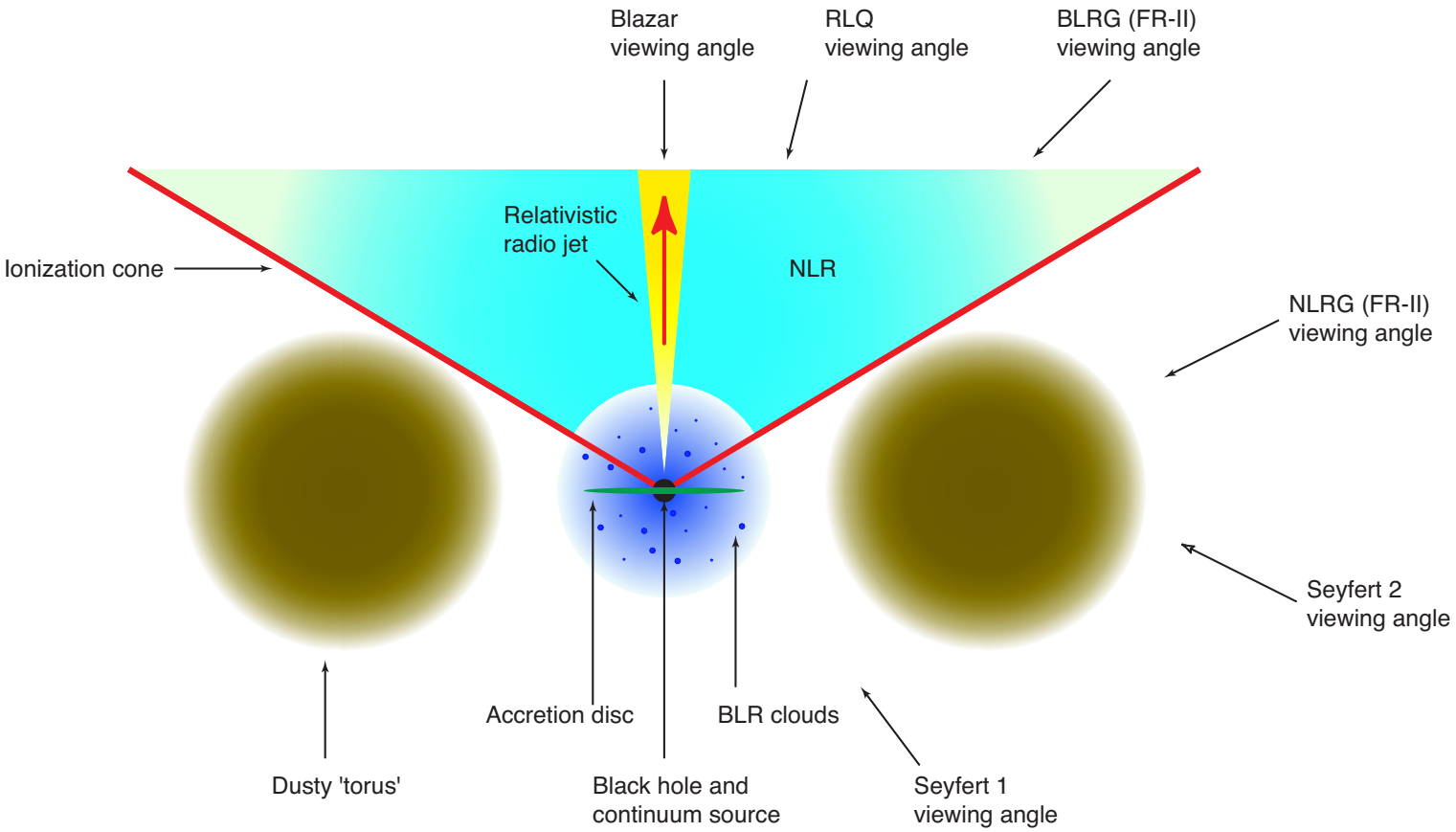

Figure 1.12: Schematic representation of the Unified Model (not to scale). Credit: A J Young. Reproduced with permission.

1.6 Quasar selection techniques and selection biases

1.6.1 The UV–excess

The "classical" method for discovering quasars in optical imaging surveys is the so–called ultraviolet excess (UVX, Sandage (1965), Schmidt & Green (1983)). This technique consists in the fact that stars and quasars, the two main classes of objects with point–like morphology found in any survey, have quite different spectral energy distributions. As a result, in a $(U-B)$ versus $(B-V)$ color–color diagram, the majority of the objects with stellar morphology and $(U-B) < -0.2$ (i.e. the objects are brighter in the U–band) will probably be quasars. The UVX has proved to be the most efficient method for obtaining large quasar samples with $z \leq 2.2$.

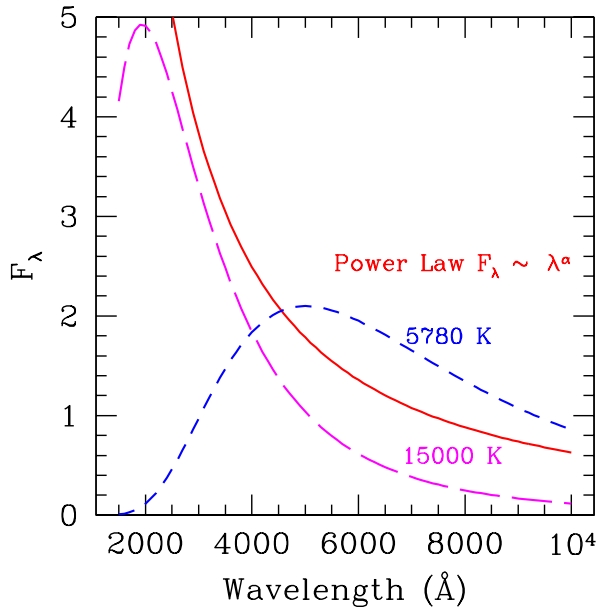

Figure 1.13: Comparison of a power–law curve, describing the quasar spectral energy distribution, and that of two black bodies at different temperatures. Based on the flux differences in the U and B filters between quasars and stars, it is possible to identify the different populations in the $(U-B)$ versus $(B-V)$ color–color plane.

The idea of describing QSO SEDs with a power–law has already been introduced in paragraph 1.4.2. The difference between stellar and quasar SEDs is schematically represented in Fig. 1.13, where the solid line shows a power–law (with index $\alpha = -0.50$ and $F_\nu \propto \nu^\alpha$) and the two dashed lines two black body (BB) spectra at different temperatures. Although the three SEDs have been arbitrarily re–scaled for visualizing purposes, the plot describes successfully the UVX concept, as the UV–excess only accounts for the *differential flux*, via the color indexes, and *not* for magnitudes in specific energy bands.

Let us first compare the power–law to the low–temperature BB spectrum. In the UV waveband ($\sim 3100 - 4000$Å, Bessell (1990)), the BB flux experiences a significant drop, due to its convex shape. As a consequence, the ultraviolet flux–excess of quasars will make them look much bluer, hence they will occupy a different area in the UBV plane. If we now compare the quasar SED to that of the hottest black body, it appears that they have about the same slope in the ultraviolet. Nevertheless, thanks to its power–law shape, the quasar SED experiences a smoother drop with respect to the stellar one in the blue waveband ($\sim 3500 - 5200$Å), counterbalancing the latter effect. As a result, the $U - B$ quasar color index will still look bluer with respect to the stellar one, allowing to identify QSOs in the color–color diagram.

Let us know have a more detailed view on the UV–excess. Fig 1.14 shows the monochromatic flux as a function of wavelength (in the observer's plane) for a composite QSO template (Vanden Berk et al. 2001), at three different redshifts (0 – lower panel, 1.5 – middle panel, 2.5 – upper panel). Overplotted are the transmission curves for the U, B and V Bessel filters (Bessell 1990), together with the identification of some of the most prominent emission lines in QSO spectra. As it was also shown in Fig. 1.13, low–redshift quasars (lower panel) clearly show an ultraviolet excess in their continuum emission, which makes their identification relatively easy with respect to stars. As we go to higher redshifts (middle panel), emission lines shift through the U–filter and the UV–excess method still differentiates quasars from stars. Nevertheless, for redshifts higher than 2 (upper panel) the Lyα line gradually exits from the U–filter and the QSO continuum emission drops significantly. As a result, the method fails in isolating high–z QSOs, with $z > 2.2$.

1.6.2 Red and Reddened Quasars

Our view of the extragalactic universe is extremely limited at visible wavelengths. Dust particles, found in the objects themselves or in clouds between the background source and Earth, attenuate the emitted light with an absorption efficiency which is decreasing very rapidly with increasing wavelength. By comparing the optical and infrared apparent magnitudes of samples containing objects selected using wavelengths unaffected by dust (e.g. in the radio), we can deduce the corresponding light losses.

For studying the effect of dust on quasars, Webster et al. (1995) studied a subset of radio–loud quasars from the Parkes catalog (Wright & Otrupcek 1990). The sample consisted of 323 objects, $\sim 70\%$ of which were spectroscopically identified quasars, $\sim$15% galaxies, BL–Lac objects and planetary nebulae, while $\sim 15\%$ had no spectroscopic identification. When comparing the $B_J - K$ color (where B_J denotes for the photographic blue apparent magnitude) to their infrared K–band magnitude, the scatter in the $B_J - K$ color was enormous, with $1 < B_J - K < 8$, much larger compared to that of optically selected quasars.

The rising question was whether the observed scatter could be interpreted as intrinsic

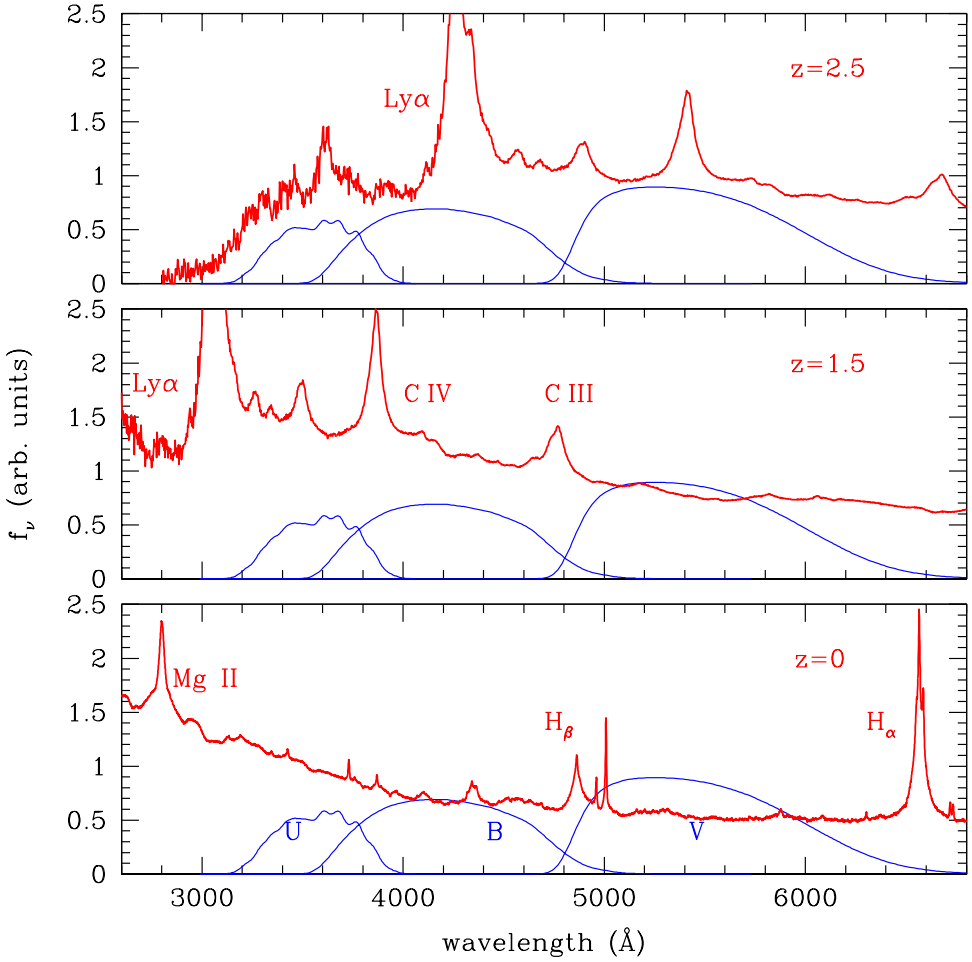

Figure 1.14: Schematic representation of the concept used for identifying quasar via the UV–excess technique. A typical quasar spectrum (Vanden Berk et al. 2001) has been shifted at three different redshifts. Overplotted are the typical transmission curves for the U, B and V Bessel filters (Bessell 1990). The method succeeds in separating QSOs from stars, up to redshifts $z \simeq 2.2$, since the integrated total flux (continuum emission or continuum + emission lines) in the U and B–bands make quasars occupy a different locus in the UBV color–color plane.

light absorption (the dust being distributed throughout the quasar host galaxy or in the dusty torus) or due to dust somewhere along the line of sight. A third explanation, already proposed by Wills et al. (1992) when studying the properties of the IRAS 13349+2438 quasar, suggested that the large scatter in the quasar colors was actually a geometrical effect: the redder quasars were being viewed through the dusty torus and their optical light was experiencing a much higher absorption. The arguments given by Webster et al. (1995) converged towards the dust extinction scenario, on the basis of spectral features of the reddest quasars and simulations on the luminosity function of absorbed quasars.

The Webster et al. (1995) study was performed on radio–loud quasars, which only constitute $\sim$ 10% of the total quasar population. If indeed dust is responsible for the reddening, this will have a significant impact on our current knowledge on the QSO population, since we will have to revisit the QSO luminosity function, in case radio–quiet quasars are affected in a similar way by dust. Benn et al. (1998) argued against the conclusions about a dusty population, on the basis of other effects affecting the extremely red $B_J - K$ quasar colors. Some of the arguments were related to physical processes, such as contamination by starlight from the host galaxy and by optical/infrared synchrotron radiation associated with flat radio spectra, while other suggested observational effects, such as variability between the epochs at which the B_J and K magnitudes were obtained, or the significance of the photometric errors.

Although Francis et al. (2000) confirmed the basic result of Webster et al. (1995), Richards et al. (2003) demonstrated that, for the majority of the SDSS quasars, redder colors can be explained by intrinsic processes in the AGN. These quasars are "red" in the sense that their continuum emission does not decline as steep as that of the classical, "blue" quasars. Additionally, the major difference between red and reddened QSOs is that the spectra of the former ones do not show a differential reddening (more in the shorter wavelengths than in the longer), due to dust extinction, like the latter ones do. Richards et al. (2003) also concluded that features accounting for dust extinction in QSO spectra corresponded to about 6% of the total SDSS quasar population. According to their estimations, the number of missed red and reddened quasars in surveys like the SDSS (different from typical optical surveys because of the filter selection) comes down to "only" 15%.

The poor understanding of the properties of red/reddened quasars comes from the fact that we are trying to study this population using non–IR surveys, since technical limitations have hampered, until recently, this approach. Nevertheless, the impact of intervening dust in the foreground galaxies in lensed quasars (Falco et al. 1999) and reddening because of (a) damped Lyα systems (Ellison et al. 2005) or (b) the cosmological evolution of the host galaxies (Wyithe & Padmanabhan 2006), can not be explored unless we study quasar samples at wavelengths less affected by dust, spanning a wide redshift range. Additionally, IR emission in quasars, a signature of X–ray absorption due to dust surrounding the AGN power house, can not be studied in detail unless we work on IR–selected samples.

1.6.3 The K–excess

Because of their power–law nature, quasars show an excess in their flux – with respect to stars – both in the UV and IR (Fig. 1.13). Therefore, in the near–infrared color plane, they should also occupy a distinct region with respect to stars. This infrared analog to the UVX method, where the K–band filter (2.2 μm) plays the role of the U–band, is known as K–excess (KX, Warren et al. (2000)). Thanks to the use of at least one infrared filter,

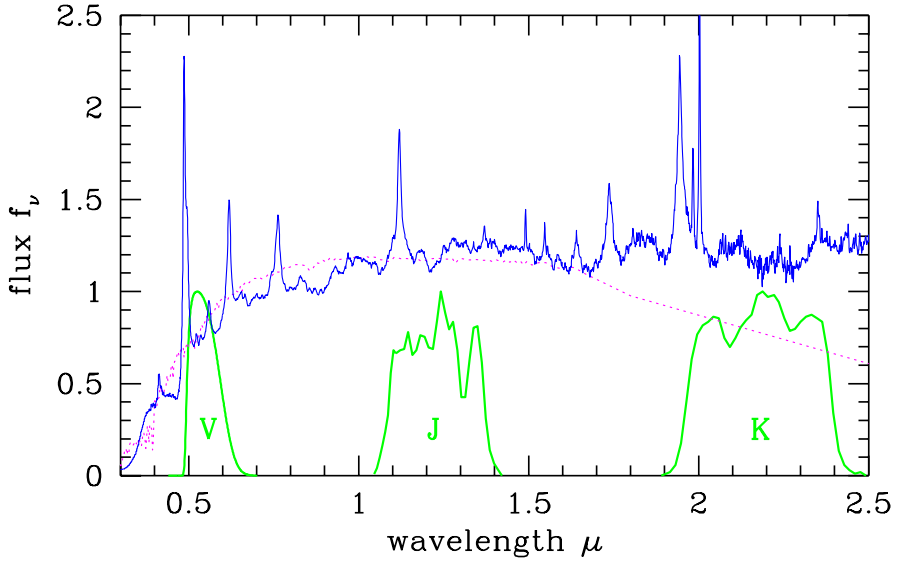

Figure 1.15: Comparison of the spectral energy distribution of a quasar (solid line) and a star (dotted line), demonstrating the principle of the KX method. The composite QSO SED is redshifted to $z = 3$ and the stellar spectrum corresponds to an early K–type star. The transmission curve for the V, J and K–band filters is also plotted. The redshifted quasar and the star have the same $V - J$ color. However, the amount of flux entering the K–band is higher for the QSO, producing the K–excess effect. The flux is in arbitrary units. (Figure taken from Warren et al. (2000))

the method cancels the biases associated with the UVX, being capable of identifying both quasars at $z > 2.2$ and quasars which are red or reddened. Fig. 1.15 demonstrates the effectiveness of the KX method, by comparing a synthetic quasar spectrum (solid line) to the SED of an early K–type star (dotted line). The optical LBQS quasar spectrum (Large Bright Quasar Survey, Francis et al. (1991)) was redshifted to $z = 3$. The stellar SED comes from Buser & Kurucz (1992). From the plot we conclude that although both the quasar and the star will have about the same $(V - J)$ color index, the former one will have a much redder $(J - K)$ color thanks to its K–band excess.

1.6.4 Selection biases and the K–excess

As we saw in section 1.6, the UVX technique suffers from two selection biases, one against high$-z$ and one in favor of "blue" quasars. With their paper, Webster et al. (1995) raised the issue of a third selection effect, which is not just related to the UVX but to all surveys using optical selection criteria for identifying quasars. This selection bias is related to the effect of dust, which is two–fold: dust makes objects look redder (reddening), but also fainter (extinction). The former effect means that, in a color–color diagram, quasars experiencing reddening will not occupy the same locus as the ones that do not. On the

other hand, light extinction (a) in quasars found in a dusty environment, or (b) due to line–of–sight dust, will make quasars have fainter apparent magnitudes. As a result, reddened quasars will be missed in a flux–limited survey, optimized to detect "classical" (i.e. blue) QSOs.

As a general rule, every survey will select objects respecting its own selection criteria (e.g. color, flux limit, etc.). Objects violating these criteria will not be selected, hence bias–free surveys do not exist. For this reason, it is essential to study any object population using complementary surveys, since this is the only way to ensure that the diverse properties characterizing the different populations will be compatible with the selection criteria of at least one of the surveys.

The idea introduced by Warren et al. (2000) comes as a solution to the problem of selection biases from which optical imaging surveys suffer, since it is designed to detect high$-z$, red and reddened QSOs. Nevertheless, until recently, the implementation of the KX has been limited by a single (but pathological) "drawback": the need for observations in at least one infrared band. Observing in this part of the electromagnetic spectrum encompasses some complications, such as:

- The small field of view, as IR technology is not as advanced as the one used for constructing optical detectors.
- The more complex electronics and infrastructure, as IR detectors have to be cooled down to much lower temperatures (77 K).
- Due to its nature, the data reduction of NIR data requires a more demanding and detailed handling of the information registered on the detector.

The trade–off between pencil–beam deep surveys, or large–area shallow surveys, imposed by the small size of IR detectors, reduces the probability of discovering "exotic" objects and high–z QSOs, unless huge amounts of time and large–aperture telescopes are allocated. Finally, the difficulty in obtaining simultaneous (or, at least, in close by epochs) telescope time, for performing the optical and IR observations, is another element that has to be taken under consideration. The conjunction of the above factors is probably the explanation for the very few publications that have used the K–excess concept until now.

Barkhouse & Hall (2001) studied the properties of about 2800 QSOs from the Véron-Cetty & Véron (2000) catalog, having a counterpart within a $4''$ radius in the 2MASS survey[7] (Skrutskie et al. 1997). Their objective was not to detect new quasars but rather study the selection effects when implementing infrared, or optical–infrared color criteria to their large statistical sample. Due to the lack of V–band photometry they used the B filter for studying the properties of the quasars in the BJK color–color diagram. Their

[7]Two–Micron All–Sky Survey

main conclusion was that this filter combination is more sensitive to detection of QSOs affected by dust.

Croom et al. (2001) implemented the KX in the $48\,\mathrm{arcmin}^2$ covered by the ESO Imaging Survey – Chandra Deep Field South (EIS – CDFS, formerly known as AXAF field, (Rengelink et al. 1998)[8]). They studied the $(V-J)$ versus $(J-K)$ colors of sources having $\mathrm{K}_s < 19.5$ and isolated 54 QSO candidates, having $(J-K)$ colors redder than those of the stars. For 17 of the candidates they managed to obtain good quality spectra, but only 3 sources were found to be QSOs, with the majority of the remaining objects being galaxies. Although their study was based on a quite small area of the sky, the relatively faint magnitude limit of the survey and the available multi–wavelength information of the EIS – CDFS survey allowed them to conclude that the KX–selected sources were also showing $UV-$excess, with $(U-B) < -0.2$. Hence, according to Croom et al. (2001), a KX–selected sample will include all objects that could be selected by the UV–excess, although the inverse is not valid.

Sharp et al. (2002) implemented the KX in three different fields, covering about $0.7\,\mathrm{deg}^2$ in total. Although their approach was based on the concept of Warren et al. (2000), the decided to use the gzH filters. The main advantage of this combination is that infrared observations had to be performed in only one filter, with observations in the H–band being equivalent to observing in the K, in terms of dust extinction effects and practical observation requirements. Although Sharp et al. (2002) report on the detection of 22 spectroscopically confirmed quasar candidates, selected on the basis of their $(g-z)$ versus $(z-H)$ colors, they do not provide any information regarding the initial candidate sample. Additionally, due to the incomplete follow–up observations in the g and z–bands, it was not possible to discuss about of the completeness of their survey and to properly evaluate the efficiency of their selection process.

Gregg et al. (2002) report on the detection of "The reddest quasars", as they found two QSOs with $B-K \approx 8.4$ and $B-K \geq 10$. Although the quasar selection was not based on the KX technique, the Gregg et al. (2002) paper is worth–mentioning both for the selection strategy and the interesting results. The QSO selection was based on a correlation of objects with radio emission in the FIRST[9] catalog (Becker et al. 1995) with objects in the 2MASS catalog (Kleinmann et al. 1994), with the latter ones having no optical counterpart in the Automatic Plate Measuring catalog (APM, McMahon & Irwin (1992)). Apart from the extreme $B-K$ colors, it was even more intriguing to find out that the reddest of the two QSOs is gravitationally lensed (Refsdal 1964) by a foreground galaxy acting as a lens (Winn et al. 2002). According to the gravitational lensing theory, the light of the background quasar is amplified by the lens galaxy (Refsdal & Surdej 1994). As the amplification effect is achromatic (Falco et al. 1999), the lensed quasar is expected to have bright magnitudes in all bands. The fact that, despite its bright K–band magnitude ($K_s = 13.35$), the lensed quasar is too faint to be included in

[8]This draft–paper never made it to a publication

[9]Faint Images of the Radio Sky at Twenty cm

the APM catalog, implies that the amplification of the optical light is canceled, due to extinction by dust in the lens galaxy.

Since May 2005, an infrared survey, as ambitious as the SDSS in optical wavelengths, was initiated. The UKIDSS[10] is a five broad–band (Hewett et al. 2006) survey, covering the spectral range $0.97-2.2\,\mu$m. The observations are performed using the Wide Field Camera (WFCAM, Casali et al. (in preparation)), mounted on the 3.8 m UKIRT telescope. One of UKIDSS' scientific drivers is the implementation of the KX and two of its photometric bands (Z (0.84–0.93 μm), Y (0.97–1.07 μm)) have been designed to optimize this scientific driver of the survey (Warren & Hewett 2002). Detailed information on UKIDSS can be found at: http://www.ukidss.org/

1.7 Implementing the KX in the XMM–LSS field

The XMM Large Scale Structure survey (XMM–LSS) is an X–ray survey, covering several square degrees on the sky, conducted using ESA's XMM–Newton satellite. The center of the field is at $\mathrm{RA} = 2^h\,18^m\,00^s$ $\mathrm{Dec} = -7^\circ\,00'\,00''$ (Equinox J2000). A full description of the survey's design, its science goals and the first results are given in Pierre et al. (2004).

The XMDS (XMM Medium Deep Survey,) is an X–ray survey which lies in the heart of the XMM–LSS field, covering a contiguous area of approximately 2 deg^2. Thanks to the Legacy surveys associated with XMDS, multi–wavelength information spanning from the X–rays (0.5 keV) to the mid–infrared (24 μm) is available for this region. Within this context, R, z' and K_s–band data were obtained, giving the opportunity to implement the K–excess in approximately 1 deg^2 in a sub–region of the XMDS.

Chapter 2 deals with the reduction of the K_s–band data.

[10]United Kingdom Infrared Telescope (UKIRT) Infrared Deep Sky Survey, Lawrence et al. (2006)

Chapter 2

Observations and data reduction

2.1 Introduction

In 2002, within the context of XMDS (see section 1.7), an observing proposal was submitted to the Las Campanas Observatory (LCO), Chile, asking for five nights in order to image 0.75 deg^2 in the near–infrared (K_s), using the 100 inch Du Pont telescope. The primary investigator of the proposal was Dr. J. Willis from the Pontificia Universidad Católica (PUC, Santiago, Chile), with co-investigators professor H. Quintana (PUC) and Dr. M. Pierre (CEA, Saclay, France) on behalf of the XMM–LSS consortium. The scientific objectives of the proposal were two–fold:

- To identify galaxy clusters on the basis of their X–ray emission
- To detect intrinsically red and reddened (i.e. due to dust extinction) QSOs, via the KX technique (see subsection 1.6.3), by combining already existing optical (Rz') with near–infrared (K_s) data.

Fig. 2.1 shows the optical and near–infrared coverage with respect to the XMM pointings. Each circle represents a $20'$ core diameter of the $30'$ field of view of XMM–Newton's European Photon Imaging Camera (EPIC). Each letter corresponds to various observing programs: "G" – guaranteed time, "B,C" – pointings subject to subsequent observing periods. The typical exposure time for the "G" and "B/C" pointings is of the order of 20 ksec and 10 ksec, respectively. The "SDS" identification corresponds to the Subaru Deep Survey. The large squares indicate available CTIO–4 m/MOSAIC II R, z' data. The cross-hatched area indicates the field that the proposed K_s–band observations would cover. The application was accepted and observational time was allocated to the project in October 2002.

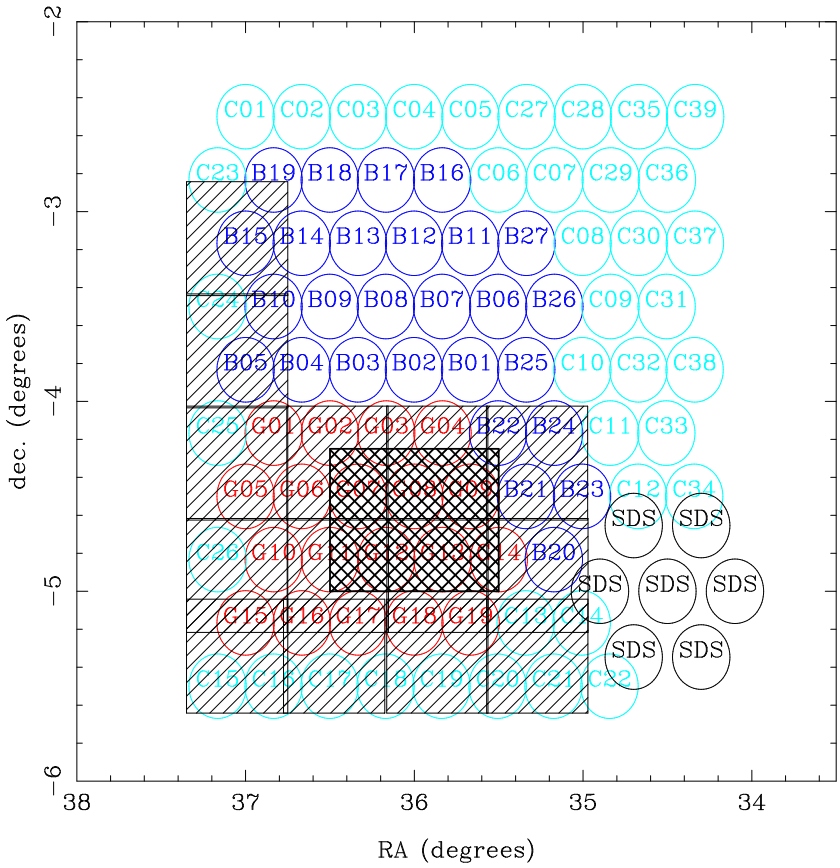

Figure 2.1: Optical/NIR coverage with respect to the XMM–LSS pointings. Each circle represents 20′ core diameter of the 30′ field of view of the XMM Epic camera. Each letter corresponds to various observing proposals: "G" – guaranteed time, "BC" – pointings subject to subsequent observing periods, "SDS" – Subaru Deep Survey. The large squares indicate already available Rz' data, which will play a key role in the present work. The cross-hatched area indicates the K_s Las Campanas observations (An extensive review on the XMM–LSS project can be found in Pierre et al. (2004)).

The scope of this chapter is to describe the steps followed for the reduction of the K_s–band data, the photometric and astrometric calibrations used to construct a mosaic image of the area covered, the simulations through which we succeeded in defining the properties of the objects in our field and, finally, the building of the K_s–band catalog.

2.2 The K_s–band observations

The observations were carried out at the 2.5 m (100 inch) Du Pont telescope, during the period $24-28$ October 2002, by Dr. G. Galaz (PUC) and myself. The Las Campanas field was observed in the K_s–band only, using the Wide Field Infrared Camera (WIRC, Persson

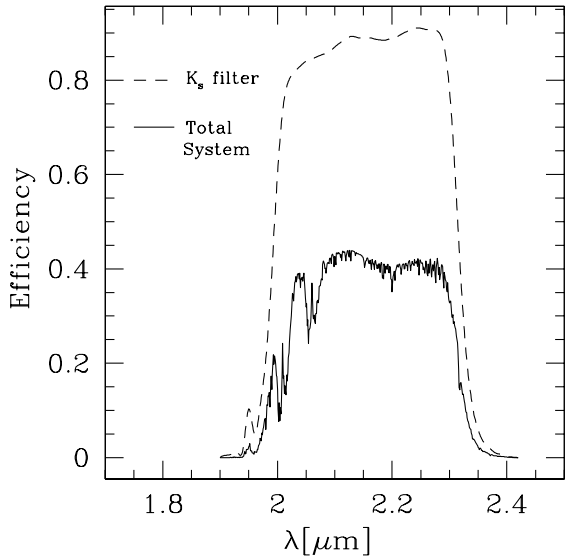

Figure 2.2: Transmission curve of the K_s–band filter (solid line) used during the Las Campanas October 2002 observing run. The total system throughput, including the contribution from the filters, telescope, camera optics, the detector and the average atmospheric transmission at the LCO observatory is shown as a dashed line.

et al. (2002)). Fig. 2.2 shows the transmission curve of the K_s filter used during the October 2002 observing run (solid line). The dashed line represents the final product of all the elements related to the observations (telescope optics, detectors, average atmospheric and filter transmission). The first night was lost due to bad meteorological conditions, whereas the nights of 25 and 27/10 were of mediocre quality. On the contrary, the observing conditions during the nights of 26 and 28/10 were photometric. Dark frames and dome flat fields were obtained at the end of each night. For some of the nights, sky flats were also acquired before starting the observations.

The majority of the following sections will be dedicated to the description of the steps implemented for reducing the K_s–band data. Nevertheless, before presenting the pipeline, we will introduce the operating principles of infrared detectors and will compare them to those of CCDs, the detectors used in most applications in optical astronomy. Since we are more familiar with the latter ones, such a comparison will help evaluate the particularities of the infrared detectors and, hence, the complexity of a pipeline meant to reduce infrared data. This comparison will then be followed by a description of the characteristics of the WIRC camera. Finally, we will introduce the concept of the dithering mode, which is indispensable when performing observations in the infrared. With all this information in hand we will then be ready to get into the details of the custom–made pipeline used for reducing the Las Campanas data.

2.2.1 Optical and Infrared Detectors

Charge Coupled Devices (CCDs) are continuous sheets of silicon (Si) covered in electrodes. They are structured in such a way that the light sensitive region is also where the accumulated charges are transported across the CCD, as part of the readout process. The charge collected under each electrode during the exposure is progressively passed, through adjacent collections sites, to the output register, as part of the parallel transfer process. The output register is then clocked in a similar way so that the charge is progressively passed from element to element of the output register and then to the output amplifier.

The physical operating principles of infrared (IR) detectors are similar to those used in CCDs, but the construction and operation differ in several significant aspects. First of all, the materials used for manufacturing IR arrays (e.g. PtSi, HgCdTe, InSb) are much more sensitive, in energy terms, than those used for CCDs. Electrons can be excited not only by photons, but by thermal energy as well. Unfortunately, the energy region at which the observing elements (i.e. telescope optics, filters, gratings) radiate coincides with the region at which infrared detectors are sensitive. Therefore, in order to minimize the noise budget, IR detectors are usually built into their associated instrumentation and the whole system is operated at very low temperatures (typically 77 K for the near–infrared detectors).

A second difference between CCDs and infrared detectors is the fact that, for both physical and technical reasons, the materials used for manufacturing IR detectors can not function properly as shift registers and amplifiers, at the same time. In other words, the arrays used in infrared astronomy are *not* CCDs, but rather consist of individual amplifiers for each pixel. As a result, each element is entirely independent, and the signal from a given pixel goes nowhere near the components that make up any other pixel. A detailed review on the technical aspects of IR detectors and their applications in astronomy is given by Joyce (1992).

The technological developments mentioned above are very recent and it is only since a few years that infrared astronomy is experiencing significant progress. Still, IR arrays are smaller in format than CCDs. The only way to cover a large area on the sky when performing IR observations is the "mosaicking", i.e. the composition of a single mosaic image from a number of individual images. However, imaging programs with objective to cover large areas using this technique require a well–planned observing strategy and a careful, multi–level data reduction.

2.2.2 The Wide InfraRed Camera (WIRC)

WIRC (Persson et al. 2002) is a near–infrared ($1.0 - 2.5\,\mu$m) wide–area camera, which consists of four Rockwell HAWAII HgCdTe arrays of 1024×1024 pixels each. With a pixel size of 18.5 μm, the scale of the f/7.5 Ritchey-Chrétien configuration is $0.2''$/pixel. As a result, each chip covers a 3.4×3.4 arcmin^2 area at the focus of the telescope.

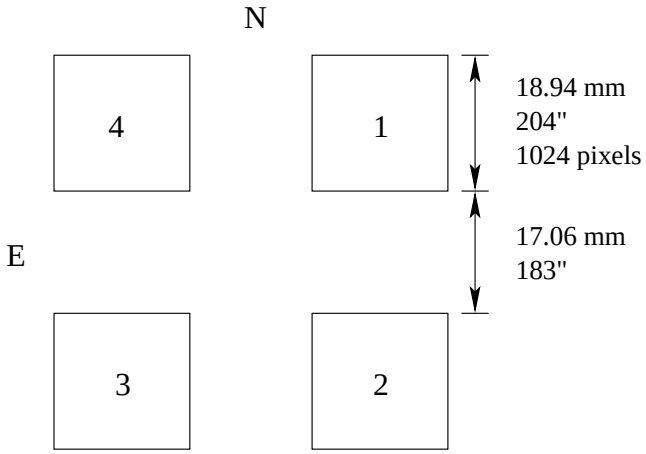

Figure 2.3: Schematic representation of the WIRC camera. It consists of four identical HAWAII detectors. The sizes (physical and angular) and spacings are indicated. The gap between the detectors is $\sim 90\%$ of their size.

Multi–array cameras are mainly used for performing wide–area surveys. Experience has shown (Persson et al. 2002) that for producing nearly uniform signal–to–noise ratio mosaics the arrays have to be spaced by either a very small number of pixels or by (roughly) the width of an array. The manufacturers decided that the latter configuration would best serve WIRC's science goals, with the four arrays being separated by $\sim$90% of an array dimension. Fig 2.3 shows the configuration of the camera and its orientation on the sky. According to their design, each chip is split in four quadrants, of 512×512 pixels each. Consequently, during the data reduction process, each of the 16 elements of the WIRC camera can be treated separately. Table 2.1 provides the most important technical specifications for the Rockwell detectors.

Due to the high brightness of the near–infrared sky ($\sim 13 - 13.5\,\mathrm{mag/arcsec^2}$), it is impossible to perform long exposures when observing in this part of the electromagnetic spectrum. To overcome this limitation it is necessary to acquire many short exposures, remove the sky background and then combine (e.g. average, coadd) the sky–cleaned frames. A commonly used strategy for removing the sky is the "dithering", i.e. point the telescope at a given position, take NDIT short exposures of DIT integration time each, offset the telescope to a new position, execute a new NDIT cycle, and so on. Since the dithering step is a small fraction of the detector's dimension, a high percentage of the area on the sky covered by adjacent images will be common.

2.2.3 Observational Strategy - Definitions

Choosing the dithering step is always a trade–off between covering a large, shallow area or a smaller but deeper one. For the scientific objectives of the Las Campanas mission a dithering step of $20''$ was considered to be a good compromise between the two options.

Parameter	Value
Manufacturer	Rockwell
Type	HAWAII
Material	HgCdTe
Spectral response	$0.85 - 2.5\mu$m
Pixel size	18.5 μm
Capacity	$> 60{,}000\ e^-$
Operating Temperature	77 K
Dark current	$< 0.15\ e^-\ s^{-1}$ at 77 K
	$< 0.01\ e^-\ s^{-1}$ at 60 K
Quantum Efficiency	70% at 2.2 μm
	60% at 1.2 μm
Bad-pixel count	<0.02%

Table 2.1: Parameters of the WIRC detectors

Each dithering cycle consisted of four exposures (NDIT = 4) of 28 seconds integration time each (DIT = 28). These four frames will be considered as one "unit", i.e. it will be only at the very early stage of the data reduction that they will be treated independently. Unless otherwise stated, they can simply be considered as a single frame of 112 seconds (4×28) total integration time. A **pointing** will represent a dithering cycle, that is four images (one per chip) of 112 seconds exposure time each. The dithering cycle was repeated nine times, producing a **cube** of nine images per chip (in total 36×112–second frames).

Fig 2.4 shows what is called an **exposure map**, i.e. a FITS file that contains the integration time for every pixel of its associated science frame. The nine frames of each cube are combined together, resulting in an image that covers an area larger than the individual images (how large depends on the dithering pattern). The arrows enclosed in the solid line box indicate the size of a single frame (1024×1024 pixels). The telescope moves according to the positions indicated by the nine numbers, and at the end of the nine dithering cycles the telescope will have observed a region on the sky at different depths. The common area observed by all nine frames will have a total integration time of $9 \times 112 = 1008$ seconds (white region). Regions on the sky observed by less frames, will have a smaller integration time. Fig. 2.5 shows what in this work is defined as a **tile**, i.e. the area on the sky covered by the four cubes (one per detector) after nine dithering cycles have finished. Looking at the WIRC configuration (Fig. 2.3), the most obvious way for filling the gaps between the four detectors would seem to produce a mosaic image of four adjacent tiles, by sequentially moving the telescope toward three directions in the sky (e.g. Tile #2 south–wards with respect to Tile #1, Tile #3 east–wards with respect to Tile #2 and Tile #4 north–wards with respect to Tile #3). However, this is a time–expensive and relatively complicated observing strategy, for covering a wide area on the sky. Besides, not only the observing overheads are large, because of the continuous moves

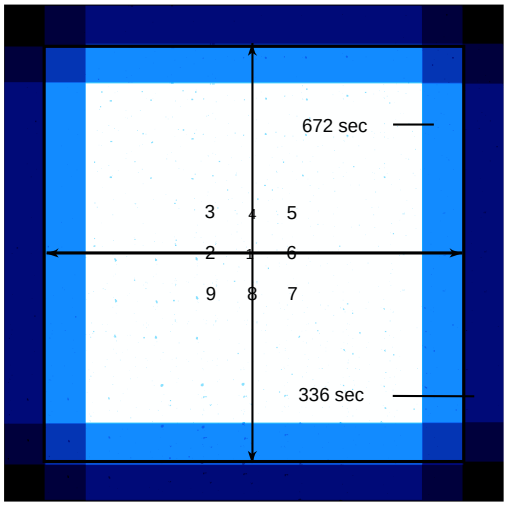

Figure 2.4: Exposure map of a sequence of nine dithered frames, each one having a total exposure time of 112 seconds. The arrows enclosed in the solid line box indicate the size of a single frame. The nine numbers correspond to the central coordinates of the nine dithered images.

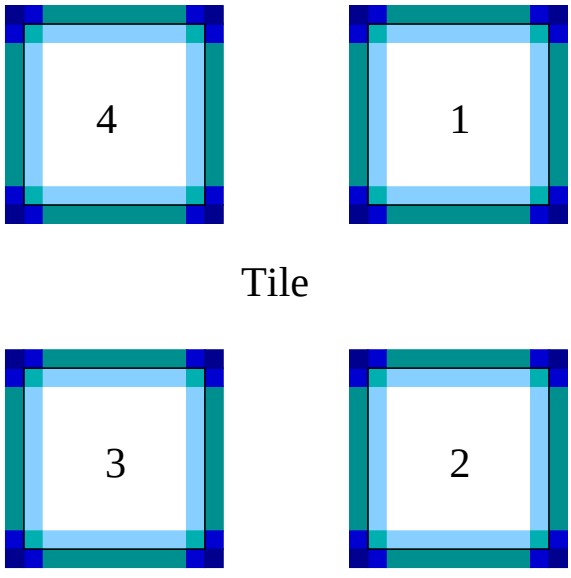

Figure 2.5: A tile is defined as the area on the sky covered by the four WIRC detectors, after nine dithering cycles have been performed. In terms of data reduction a tile will consist of four nine-frame cubes, that will be reduced independently. The central regions, marked in white color, indicate areas with a total integration time of 1008 seconds. Same orientation on the sky and detector identification as in Fig. 2.4.

in both Right Ascension and Declination, but also the risk of the telescope pointing to a wrong position is also higher. A rather simple and time–efficient methodology is to cover the area of the survey in stripes, by mainly moving the telescope toward a single direction in the sky (e.g. North–South or East–West). Starting from the lower (or upper) border of the area, successive tiles are observed, stepping each tile in Declination until the field limit is reached. The Las Campanas area was covered by moving the telescope in the South–North direction. In this way, chips 1–2 and 3–4 were "associated" to each other, covering the survey's area in stripes. Once the area covered by chips #1,#4 or #2,#3 has reached the northern border of the survey, the telescope is re–directed to the southern border, slightly shifted toward the east and the observations restart for covering a new stripe.

Because of the $\sim$ 90% gap between the detectors, the same region on the sky can be observed by different tiles. These regions are formed at the border of the combined images and have less exposure time with respect to the central region (see Fig. 2.4). In a mosaic image these regions will serve as "nodes" for photometrically and astrometrically associating different tiles among them. Thus, the mosaicking technique favors the overlapping regions, increasing their total exposure time and improving their signal–to–noise. Fig. 2.6 summarizes all the above in a single graph. For clarity reasons, only four tiles (with identification 9–12) are used. Detectors 1,2,3 and 4 are plotted in black, red, green and blue color, respectively. The southernmost part of the area on the sky is covered by detectors 2 and 3 belonging to tile number 9. Two stripes on the sky are formed when moving the telescope toward the north: one by chips 1–2, the other by chips 3–4. The gap between chips 1–2 (4–3) of pointing No 9 is covered by chip No. 2 (3) of pointing No. 10, forming overlapping regions. In this way, the gaps are filled and overlapping regions are formed, as shown in the graph.

The weather conditions on the night of October 24 forced us to stop the observations after the first four pointings. The poor quality of the data did not allow to perform any kind of reduction and therefore that region of the sky was completely lost. Table 2.2 provides the identification of the tiles observed each night.

Night	Tile Identification
Oct. 24	–
Oct. 25	05–19
Oct. 26	37–48
Oct. 27	49–63
Oct. 28	55–61

Table 2.2: Log–book of the 2002 Las Campanas observations. The first four tiles were completely lost, due to bad weather conditions.

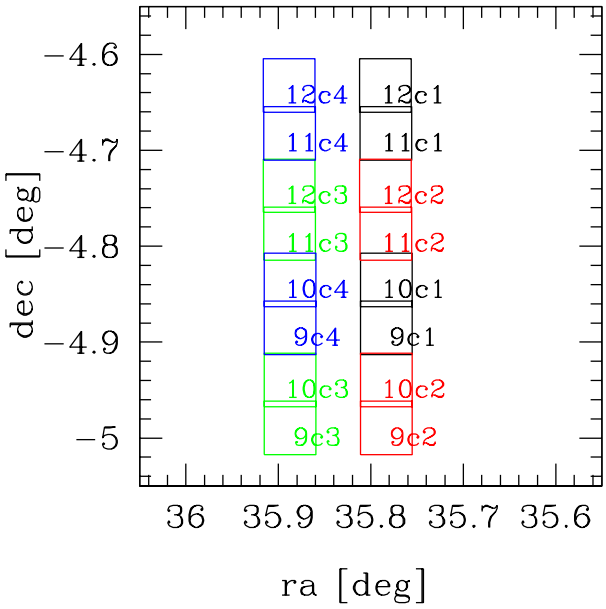

Figure 2.6: Schematic representation of the observing strategy followed for covering the Las Campanas field. For simplicity reasons, only four pointings (9–12) are plotted. In the color version of the plot, black, red, green and blue correspond to detectors 1,2,3 and 4, respectively. By moving the telescope in the South–North direction, the observations "scan" the sky in stripes, produced by chips 1–2 and 3–4. Sequential pointings help fill in the gaps between the detectors and at the same time form overlapping regions.

2.3 Data Reduction

2.3.1 The basic tools: IRAF & PHIIRS

The tools that were extensively used for the data reduction were mainly IRAF[1] and a package called PHIIRS, acronym for "Pat Hall's Infrared Imaging Reduction Software" (Hall et al. 1998). PHIIRS is an astronomical package written in IRAF's command language, specialized in the reduction of near–infrared data. The basic image manipulation (e.g. arithmetic operations, combination of images and statistics) was performed using some of IRAF's general tasks, such as imarith, imcombine and iterstat. However, the complete reduction of the Las Campanas data was performed by means of a pipeline written by the author, which consisted of a combination of IRAF & PHIIRS commands. The pipeline was built gradually, by testing different sets of parameters for each individual task. The best combination was decided on the basis of statistics, performed on each task's output images. Once the performance of all tasks was optimized, the tasks were put together

[1]IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the National Science Foundation

and the pipeline was re–executed over the whole data set.

Before presenting the pipeline's core we will briefly describe imcombine, an IRAF task that we extensively used throughout this work, both directly and indirectly (i.e. called internally by other tasks). The correct choice of imcombine's parameters was crucial for performing an optimized data reduction and in this work we will frequently have to justify why one set of parameters was chosen instead of another. Since this is only a brief description, we advice the reader to consult the corresponding IRAF help pages for more details. The purpose of the task is to combine a set of images to an output image, using a weighted average or median. Pixels may be rejected by using pixel masks, threshold levels and rejection algorithms. The images may be scaled multiplicatively or additively based on image statistics, image header keywords, or text files, before rejection. After the scaling, the images may be spatially combined using different types of coordinate systems, to produce a single image of bigger size than the input images. imcombine's most important parameters are:

- project: a parameter controlling the number of output images. If set to *no*, the planes of an image–cube are collapsed to a single frame.
- offsets: defines the type of coordinate offsets to be applied to the individual images before producing the output files. If the parameter is set to *none* then the output image(s) will have the same dimensions as the individual input images and the (x,y) pixel in any of the input frames will be found in the (x,y) position in the final frame(s).
- scale: multiplicative image scaling to be applied, to bring the input images to the same level, using the *mode, median*, or *mean* value of the individual image pixel value distributions.
- reject: type of rejection operation performed on the pixels remaining after offsetting, masking and rescaling. Two of the more commonly used criteria are the *n–low/n–high* pixel rejection ("*minmax*") and the *sigma–clipping* algorithm.
- weight: type of weights to be applied when producing the final image. The main choices are *none, mode, median*, or the *mean*.
- combine: type of combining operation performed on the final set of pixels (after offsetting, masking, rescaling and rejection). The choices are average or median.

2.3.2 Overview of the pipeline

An overview of the steps followed for reducing the K_s–band data has as follows:

- PRE–REDUCTION:
  - Produce a master dark frame (MDF) per night, per chip.
  - Remove the MDF from the individual flat–field (FF) frames, make a normalized flat–field (NFF) per night.
  - Using the flat–field frames, construct a Bad Pixel Mask (BPM) per chip.
  - Generate 112–sec dithered frames.
- FIRST LEVEL REDUCTION:
  - First–pass sky subtraction on dithered frames.
  - First–pass subtraction of sky residuals:
    - Split frames in quadrants.
    - Remove X–direction sky residuals.
    - Reassemble quadrants to generate dithered frames.
  - Compute the dithering pattern for every nine–frame dithering cycle.
  - Coadd the dithered frames of the same cube, generate coadded image.
  - Generate individual–image object masks (OM) out of coadded image.
- SECOND LEVEL REDUCTION:
  - Second–pass sky subtraction on dithered frames.
  - Second–pass subtraction of sky residuals.
  - Perform destriping.
  - Identify cosmic rays (CR).
  - Identify additional outliers (O).
  - Generate combined pixel mask (CPM) for each dithered frame (CPM = BPM + OM + CR + O).

With the construction of the combined pixel mask the complete study of the properties of the dithered frames has arrived to its end. After the photometric calibrations, some of the steps of the data reduction were repeated, for generating (from the dithered frames) the final–pass coadded frames, from which the Las Campanas mosaic image was built. In the subsequent paragraphs we will describe in more details the individual steps of the pipeline.

2.3.3 Preliminary data reduction

2.3.3.1 Master Dark Frame

15 dark frames (per chip) were taken at the end of each night, of 8 seconds exposure time each. A *Master Dark Frame* per chip and per night was created, using IRAF's imcombine task. All 15 dark frames were scaled to the same level using the mode of the pixel value distribution of the individual frames. The two lowest and two highest pixel values were rejected ("minmax" rejection), and the mean of the 11 remaining values was used for computing the intensity for each pixel of the Master Dark Frame.

2.3.3.2 Flat fields

As for the case of the dark frames, 15 dome flat fields were obtained at the end of the night (having the same exposure time as the darks). From the individual flat field frames, the corresponding Master Dark Frame was subtracted (same chip, same night) and then a combined flat field frame was created, using the same parameters as for the Master Dark frame. By dividing the combined image with its median value a Normalized Flat Field (NFF) frame was produced, per chip, for each night. The pixel distribution of a typical flat field can be fit by a Gaussian distribution of mean $\mu = 1$ and σ equal to the RMS of the pixel value distribution, according to the following formula:

$$G(x) = \frac{1}{\sqrt{2\,\pi}} \exp\left[\frac{-(x-\mu)^2}{2\,\sigma^2}\right] \tag{2.1}$$

A histogram of the pixel values of a typical flat field, together with a Gaussian distribution (long–dash line) are shown in Fig. 2.7. The pixel–value distribution in the flat–fields is normal and largely (but not totally) explained by random photon noise.

2.3.3.3 Bad pixel mask

Bad pixels are the array elements which produce a grossly inaccurate signal, outside the detector's predicted behavior. As their presence is due to imperfections during the fabrication process, bad pixels occupy fixed positions on a specific array. Because of their erroneous signal, they can influence (a) the image statistics and (b) the photometry of the objects to which they might be associated. Therefore, generating a bad pixel mask (BPM) helps to better control the sources of photometric uncertainty, as it avoids the use of problematic pixels during the various steps of the data reduction.

Since the BPM is not time dependent, any set of normalized flat fields can serve for this study. As it has been shown in paragraph 2.3.3.2, the pixel value distribution of any flat field can be considered to be a Gaussian. By (a) subtracting from every normalized flat field NFF_i its mean value m_i and (b) dividing by the RMS of the pixel distribution

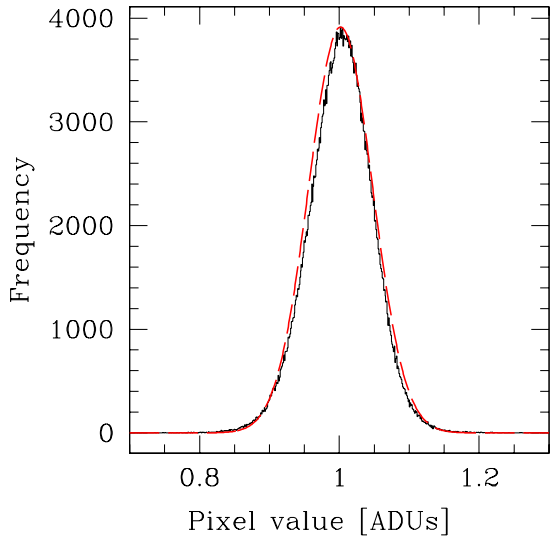

Figure 2.7: Typical histogram of a normalized flat field (solid line). The largest part of the pixel values follows a Gaussian distribution (marked with the long–dash line).

σ_i, the pixel value distribution of any flat field can be studied in terms of σ. Since, for a Gaussian distribution, the expected number of values at $N\sigma$ away from the mean is known, it is straightforward to identify the outliers in the flat fields. For example, the probability of a value to be outside the $[-3\sigma, +3\sigma]$ limit is $P = 1 - 0.99730 = 0.27\%$. Therefore, for the 1024×1024 pixel values of a flat field, there should be ~ 2800 pixels outside the 3σ limit. Fig. 2.8 shows the pixel value distribution of the normalized flat field presented in Fig. 2.7, in terms of σ. A Gaussian distribution of mean $\mu = 0$ and $\sigma = 1$ is plotted as a red long–shade line. The narrow peak on the left of the main distribution has to be related to bad pixels on the infrared detector, since the pixel rejection algorithm (used when constructing the flat fields) did not succeed in rejecting[2].

For building the bad pixel mask, IRAF's task imedit was used to set to unity all pixel values within a certain range and to zero the value of the outliers. It should be noted, however, that one must be very careful with the pixel characterization as as good or bad. For example, in Fig. 2.8, it is clearly seen that the pixels at the wings of the histogram follow a certain distribution, even if it is not a Gaussian one. Trying to be conservative, we considered as bad all pixels outside the $[-19\sigma, +13\sigma]$ range. The name of the bad pixel mask was added to the image headers of the science frames, since this information has been used during later stages of the data reduction.

[2]The explanation for this is that the flat–field frames are not dithered. As every bad pixel lies in the same (x,y) array position for all flat–field frames taken with the same chip, imcombine's statistical rejection algorithm fails to identify, and hence, reject them.

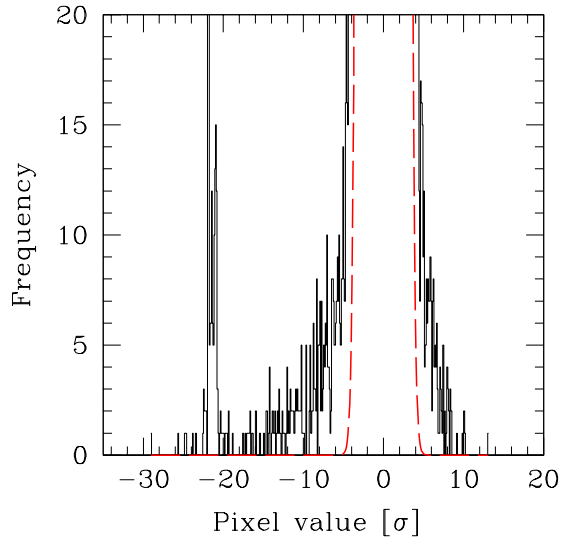

Figure 2.8: A close–up at the "wings" of the pixel distribution of Fig. 2.7. The pixel values have been re–calculated, in terms of σ. The outliers can be directly identified when comparing to a Gaussian with mean $\mu = 0$ and $\sigma = 1$ (plotted as a long–dashed line).

2.3.3.4 Construction of the dithered frames

As explained in section 2.2, four exposures of 28 seconds each were taken at each pointing. The four frames were coadded (no rejection algorithm was applied) and the output frame was flat–fielded with the corresponding normalized frame (chip, observing night), producing a single *dithered frame*, of 112 seconds total integration time.

Contrary to the flat–field frames, the four science frames are coadded without having previously removed the sky background from them. As it will be explained in more details in the following paragraph, the spatial and temporal variations of the near–infrared sky make the sky subtraction a delicate task. Hence, for taking advantage of the information contained in a statistically significant sample and decided to perform the sky subtraction once the dithered frames of the same "cube" have been generated.

The construction of the dithering science frames is the last of the series of operations related to the preliminary data reduction. The flow chart presented in Fig. 2.9 summarizes the steps described in paragraphs 2.3.3.1–2.3.3.4.

2.3.4 The near–infrared Sky

In contrast to the optical region of the electromagnetic spectrum, the transmission of the Earth's atmosphere has strong absorption bands in the near–infrared, is primarily due

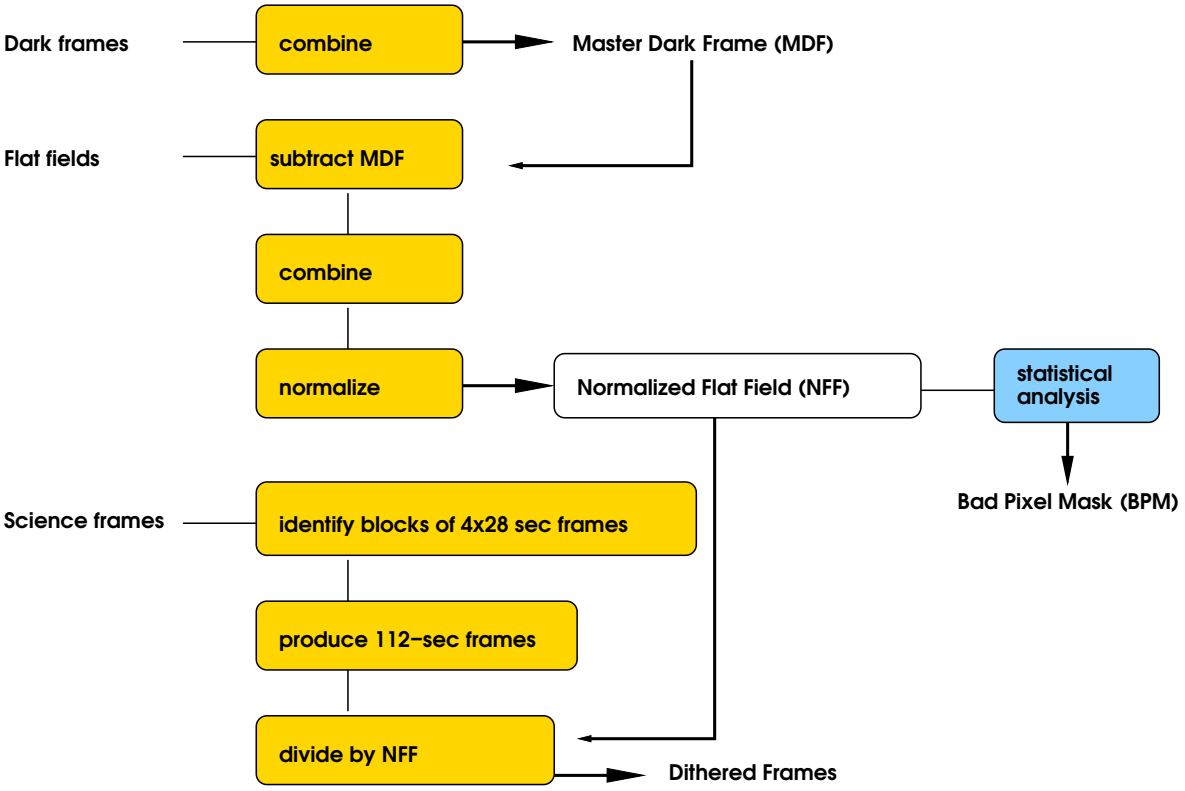

Figure 2.9: Flow chart describing the steps of the pipeline related to the preliminary data reduction.

to unresolved line emission from H_2O and CO_2. The intervening, relatively transparent regions, have been designated as photometric bands identified by letters, similarly to the visible UBVRI bands. For the near–infrared part of the electromagnetic spectrum these bands are the J ($1.1 - 1.4\,\mu$m), H ($1.5 - 1.8\,\mu$m), K ($2.0 - 2.4\,\mu$m) and L ($3.0 - 4.2\,\mu$m).

In addition to the absorption features of the Earth's atmosphere, the radiation emitted from the sky is much higher in the IR than in the visible. At wavelengths shortward of $2.3\,\mu$m this bright background primarily rises from OH^- and O_2 molecules, excited by solar radiation during the day. The typical sky brightness for the V–band is roughly 22 mag/arcsec^2, while for the K–band it is $\sim 13 - 13.5$ mag/arcsec^2. As it can be easily understood, extracting information for sources thousands of times fainter than the background can be a challenge. Besides the very bright IR sky background, the atmospheric layers formed by the OH^- and O_2 molecules display time variations. The typical time scale for the ($2.0 - 2.4\,\mu$m) region (K–band) is of the order of 20 minutes. As a consequence, associating the correct sky background with a certain group of observations is essential, especially for the detection of faint sources.

2.3.5 Background subtraction

2.3.5.1 Sky removal

Sky subtraction was performed using the PHIIRS's irsky task. The task works in a way similar to imcombine (section 2.3.1): the nine dithered frames of each cube are combined to produce nine sky images. Each sky frame is removed from its corresponding dithered image, producing a set of sky–cleaned frames. Similarly to imcombine, the combination of the frames is mainly based on three parameters, scale (that scales the images to the same level), reject (that performs a pixel rejection based on user–defined criteria) and combine (that combines the images for producing the sky frame). The selection of the best irsky parameters was based on numerous tests, as a three–step process:

- The nine frames were brought to the same scale using either the mode or the median value of the pixel distribution.
- The bad pixel map (stored in the header descriptor of each frame) was used to remove bad pixel values, and then the 3σ and the minmax value rejection algorithms were applied when combining the frames.
- The sky background was computed as the average of the remaining pixels.

The parameter choice had to be based on a sufficiently large statistical sample and the 540 frames obtained on October 25 well served for this purpose. The methodology followed was to experiment with the first of the parameters (keeping all the rest fixed), study the statistical properties (i.e. mean, median, mode and RMS) of the output frames, decide on the value of the first parameter, experiment with the second and so on. Through this iterative process, all parameters were optimized.

Fig. 2.10 shows the statistical results for the sky–cleaned frames of October 25. The left and right panels correspond to two different parameter sets, A and B respectively. For both sets the 3σ rejection algorithm was used and the frames were combined using the average option. On the left graph (set A) the frames were scaled to the same sky level using the median, on the right graph (set B) they were scaled using the mode. The upper panels show the RMS for the sky–cleaned frame. In the lower panels the mode, mean and median are plotted as green triangles, red circles and blue dots respectively. A statistical analysis of the results showed that the properties of the sky–cleaned frames were slightly better (i.e. lower rms, less scatter between the average, median & mode) in the case of median scaling.

Having decided on the scaling parameter, we experimented with two rejection criteria: the rejection of the two lowest and two highest values (criterion one) or the rejection of values outside the $3\,\sigma$ range (criterion two). The RMS for the two sets of sky–cleaned frames are shown in Fig. 2.11, from where it is obvious to see that the min/max rejection criterion produces frames with lower RMS.

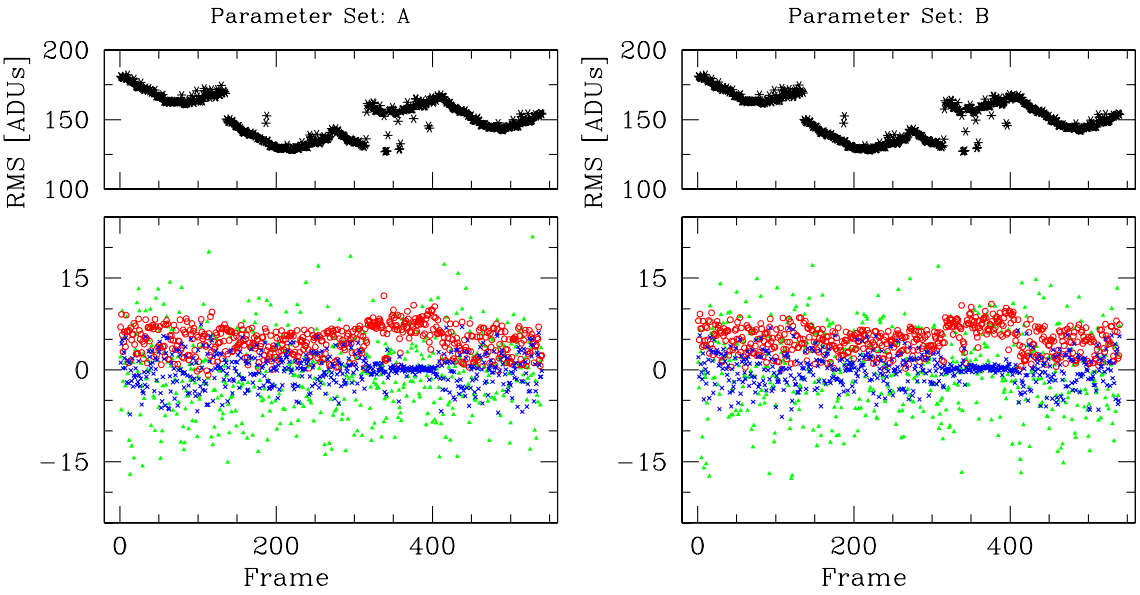

Figure 2.10: Statistics for 540 frames, sky–subtracted using the same rejection algorithm but different sets of parameters. *Left panel*: Median scaling, 3σ rejection, average of rest (Param. Set "A"). *Right panel*: Mode scaling, 3σ rejection, average of rest (Param. Set "B"). In the lower panels, the mode, mean and median are plotted as green triangles, red circles and blue dots, respectively.

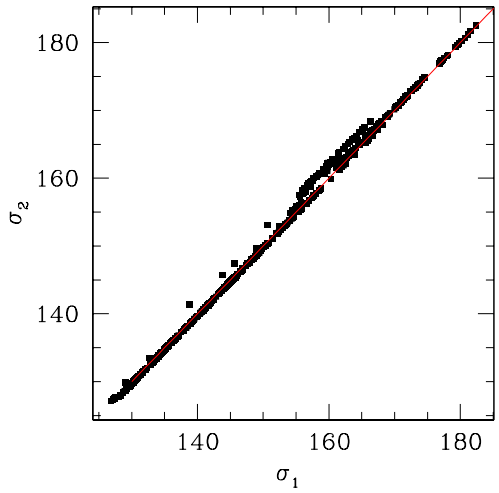

Figure 2.11: RMS versus RMS diagram, applied to the same data set as in Fig. 2.10, used for choosing between two rejection algorithms. The nine dithered frames of each cube were brought to the same level using the median, before the rejection algorithm was applied. X–axis: rejection of the lowest/highest pixel value. Y–axis: rejection of values deviating more than $3\,\sigma$ from the median. The solid red line corresponds to $\sigma_1 = \sigma_2$.

For about 98% of the data the sky background was successfully removed. However, a systematic pattern was present on the sky–subtracted frames of tile #22: it consisted of "hills and valleys" with peak–to–peak variations of the order of 20%, forming stripes along the diagonal direction of the frames. The width of the high, low intensity regions was of the order of 35 and 20 pixels, respectively. The presence of such a pattern demonstrated that the sky–frames produced by irsky did not well represent the background of the dithered frames, probably because the time–scale of the sky variations was shorter than the elapsed time corresponding to the frames used for producing the sky frames. Despite the fact that different solutions were tested (e.g. different combination of irsky parameters, fitting the residuals with a polynomial) we failed in removing the sky pattern. As a consequence, the sky–subtracted images of the four cubes of tile #22 were $\sim 15\%$ noisier than the typical RMS value of the rest of the sky–subtracted frames.

Despite its higher noise properties, we decided to fully process the tile #22 frames. This decision was taken on the basis that, at this stage, it would be rather premature to reject noisier frames, as this would have an impact on the area coverage of the K_s–band observations. Nevertheless, it is obvious that, at later stages of our analysis, a thorough study of the properties of the Las Campanas mosaic will have to be performed, because of the inhomogeneity of the frames participating in the mosaic building.

2.3.5.2 Removal of the background residuals

What in infrared astronomy is called sky background is the result of various additive factors: that of the bias level of the detector, the dark current contribution, the IR emission of the telescope configuration (optics, detectors, filters) and the sky illumination. However, each of the above factors has a different contribution in the total background, not only in terms of flux, but also in terms of contribution as a function of time. As a result, removing one of them does not necessarily entail the successful removal of the other. Although the parameters used for removing the sky background were optimized, a statistical analysis of the sky–cleaned images indeed revealed fine structures at a few percent level (for a sky intensity of 40,000 ADUs a 2% residual would still correspond to 800 ADUs).

Fig. 2.12 shows the statistics of four sky–subtracted frames, randomly chosen from the Las Campanas data set. Each panel corresponds to a frame obtained with one of the WIRC detectors (lower right with chip #1, lower left with chip #2, upper left with #3 and upper right with #4). The horizontal axis represents the 1024 columns of the frames. The perpendicular axis shows the mean intensity corresponding to each column, averaged over all 1024 lines. The graphs clearly demonstrate that a repetitive pattern is present on the sky–subtracted frames, whose characteristics are directly associated with the detector with which each frame was obtained.

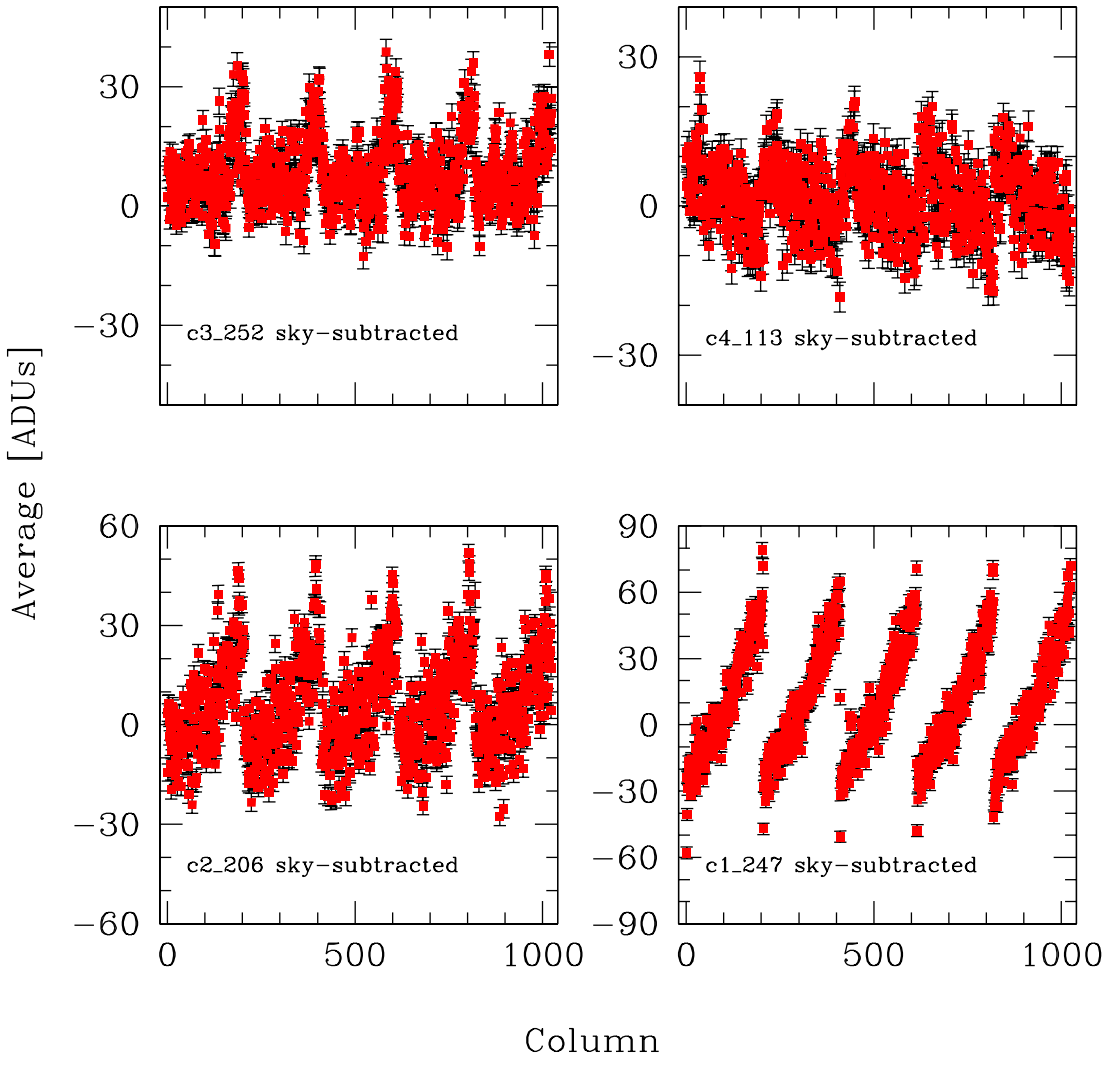

Figure 2.12: Statistics on sky–subtracted individual frames. Each panel corresponds to a frame obtained with one of the WIRC detectors (lower right – chip #1, lower left – #2, upper left – #3 and upper right – #4). The sky residuals form a repetitive pattern, representative of the detector with which each frame was obtained.

Through the use of some IRAF tasks[3] we gradually removed the Y and X–direction residuals. The results are shown in Figs. 2.13 and 2.14 respectively. Fig. 2.12 can serve as a reference, since in all thee figures the same panel corresponds to the same sky–clean, residual Y–clean and residual X–clean frame, respectively. A comparison of the three figures demonstrates that the removal of the Y–direction residuals has practically no effect on the repetitive pattern and it is only the removal of the X–direction residuals

[3]The tasks used are: imcopy, imtranspose, imarith and imcombine.

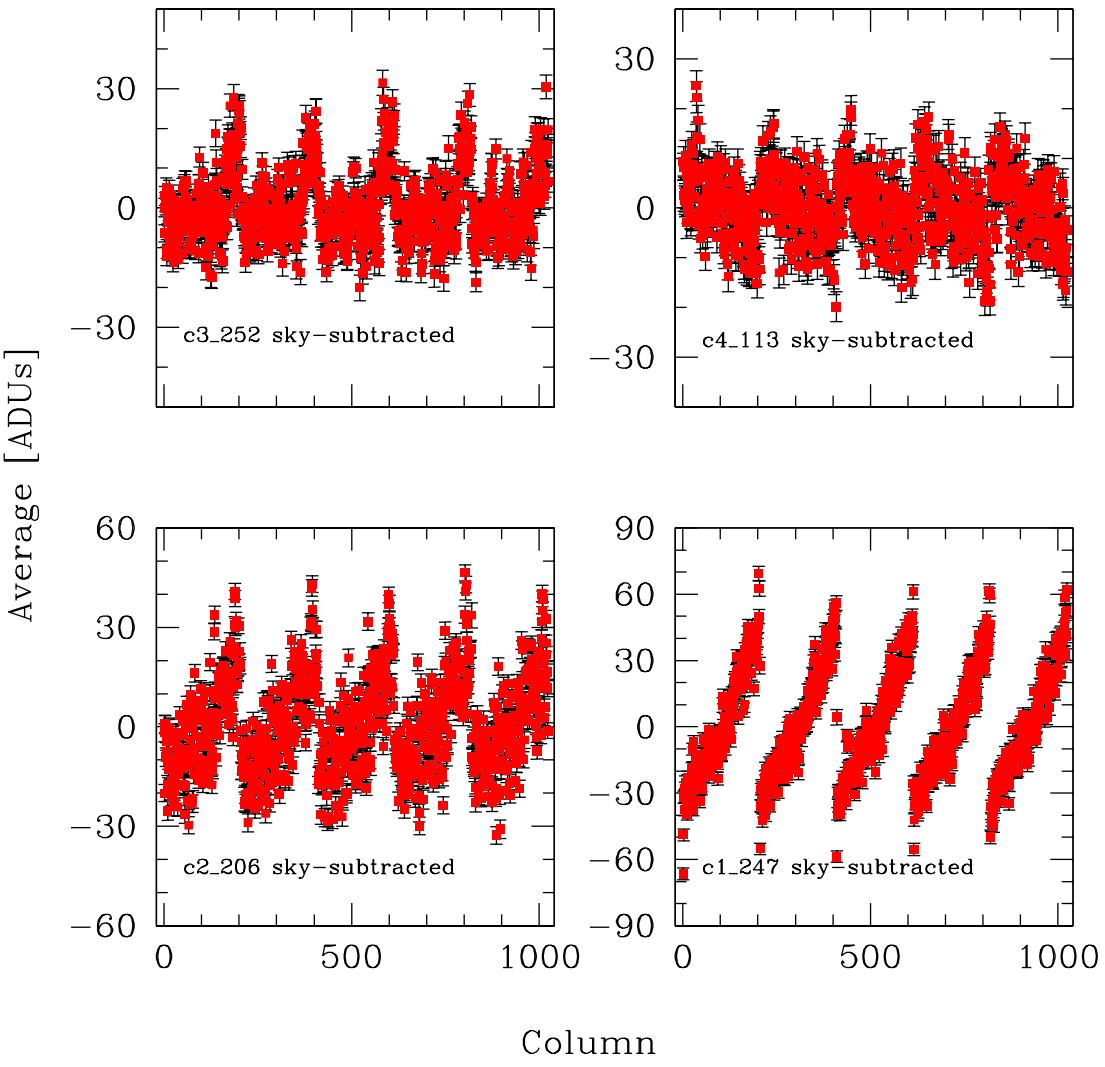

Figure 2.13: Statistics on four sky–subtracted individual frames, having the Y–direction residuals removed. The panel identification is the same as in Fig. 2.12.

that produces pattern–free frames.

As mentioned in section 2.2, each of the HAWAII detectors consists of four quadrants[4]. During the observing run some of the quadrants behaved in an unstable manner, resulting in much noisier frames. We decided to study the impact of the noisier quadrants on the sky residual subtraction using a set of $\sim$ 140 frames obtained with chip #3 on October 25. Each sky–subtracted frame was introduced to the pipeline twice, once treated as a single

[4]We name as No. #1 the quadrant at the lower left part of the detector, counting follows the clock-wise orientation

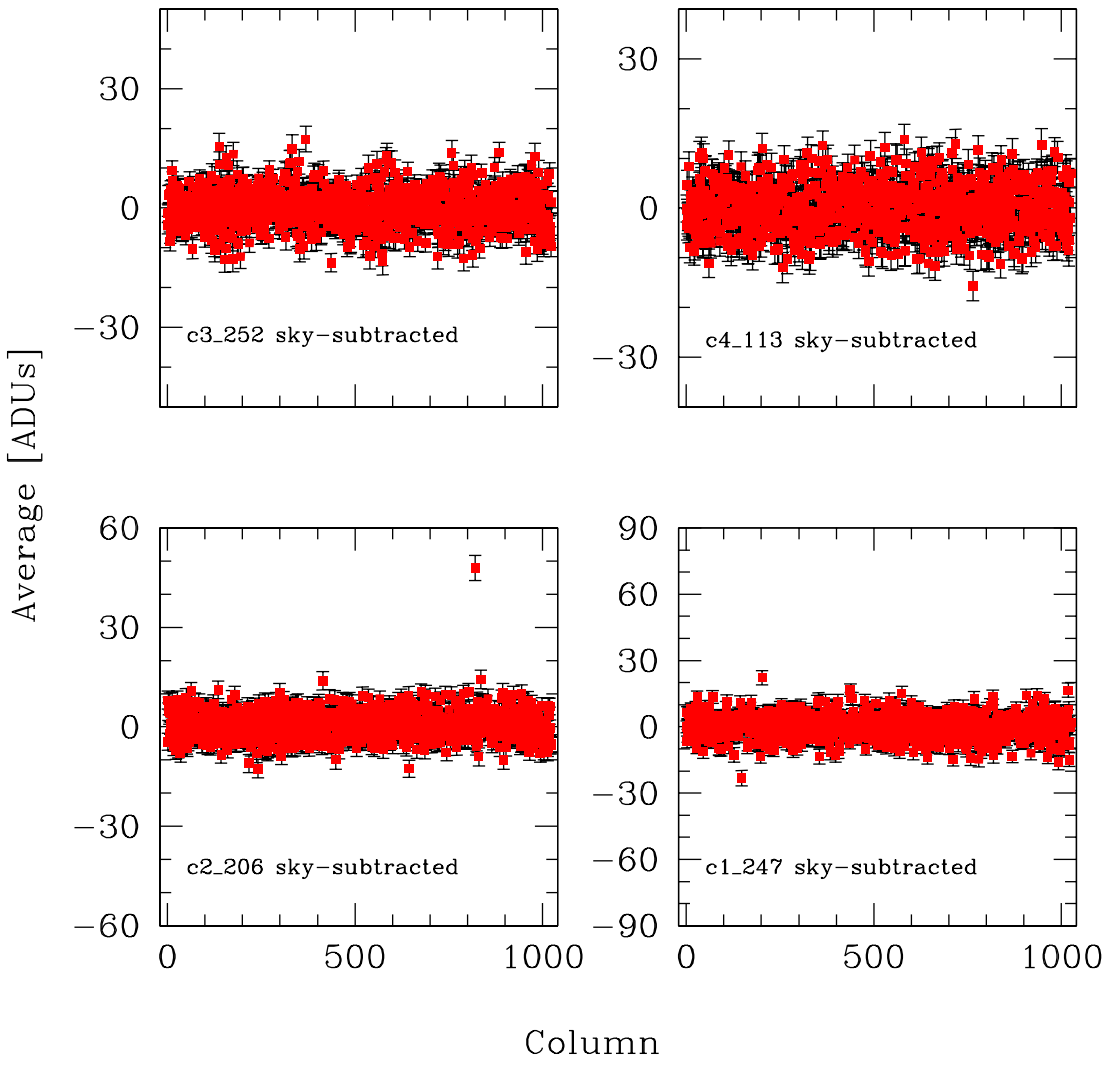

Figure 2.14: Statistics on four sky–subtracted individual frames, having the Y and X–direction residuals removed. The repetitive pattern, characteristic of the detector with which each frame was obtained, has been successfully removed. The four panels correspond to the same frames as in Figs. 2.12, 2.13.

1024×1024 pixel array and once treated as four independent quadrants of 1024×256 pixels each. Fig. 2.15 shows the mean value, averaged over the 1024×1024 pixels of each frame, as a function of frame number. Fig. 2.16 shows the results obtained for the same set of frames, having every sky–cleaned frame split to four quadrants before being reduced by the pipeline. By comparing the two graphs we conclude that a) when reducing the images in quadrants the statistics show a smaller scatter, and b) any possible source of higher RMS related to the electronics' unstable behavior only affects the specific quadrant.

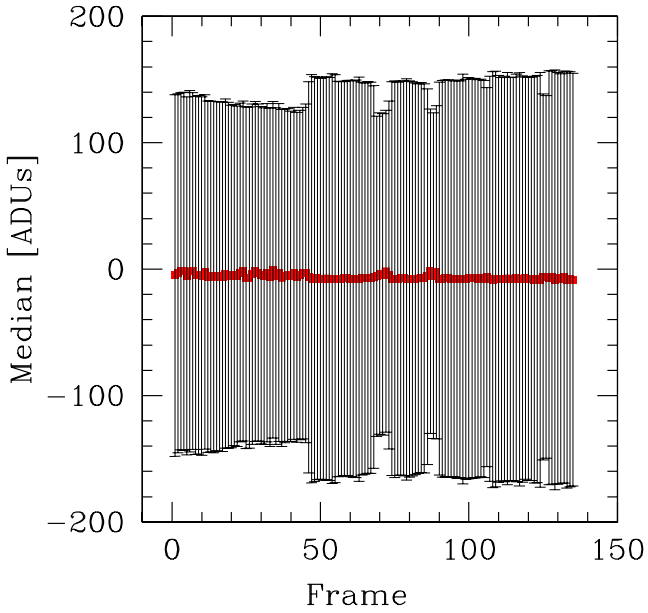

Figure 2.15: Statistics on approximately 140 sky–subtracted and X residual–free frames. Every frame was treated by the pipeline as a single 1024×1024 pixel array.

This can be understood as the additional electronic noise associated with the unstable quadrant is no longer propagated to the unaffected quadrants in the same chip.

2.3.6 Object Masking

Subtraction of the sky background and of its residuals has to be as precise as possible. Any over–estimation of their values will propagate to the photometry of the astronomical objects and it can be crucial for the fainter ones. The presence of the objects themselves in the images is the main cause for over–estimating the sky and its residuals, since they bias the statistics toward higher values.

Using the PHIIRS pipeline it was possible to identify the pixels associated to objects and produce a binary object–mask (OM) than can be used for excluding them from the creation of sky background frames. Once the object detection has been accomplished, one can associate an object–mask to its dithered science frame and repeat the subtraction of the sky and its residuals.

For ensuring a complete identification of the objects, it is preferable to perform the object detection on the coadded image corresponding to a nine–frame dithering cycle, since it has a higher signal–to–noise ratio. Initially the user needs to indicate at least one object common in all dithered frames. Based on their position, PHIIRS calculates the relative shift between them, in order to combine them and produce a coadded frame. From

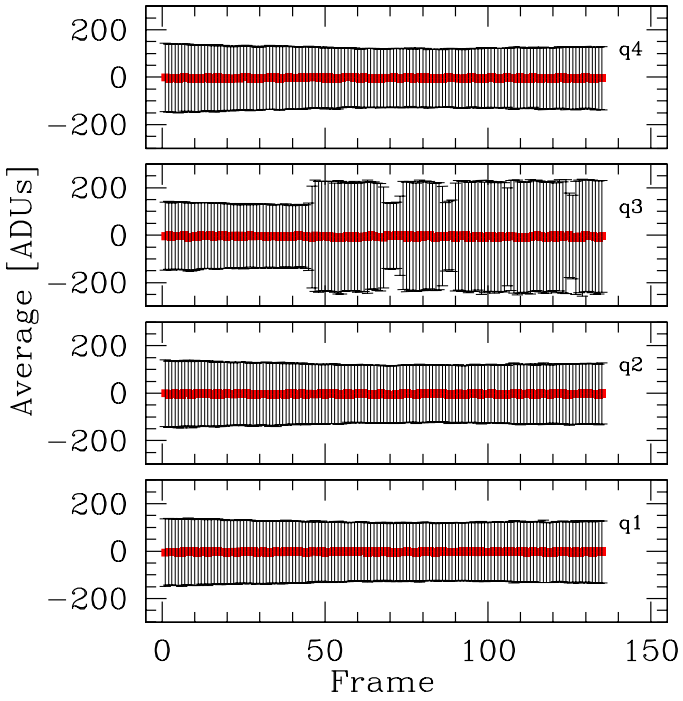

Figure 2.16: Statistics on the same set of sky–subtracted and X residual-cleaned frames, as in Fig. 2.15. The sky–cleaned frames were split into four quadrants of 1024×256 pixels each, before having their X–direction residuals removed. On the upper right corner of each panel the quadrant identification is indicated. The additional electronic noise has been identified and isolates from the analysis.

the objects detected in the coadded frame, it uses the relative shifts of the dithered frames for generating an individual object–mask for each of the nine dithered frames of the given cycle. The object masking was performed through the use of three tasks, namely irshift, ircoadd and mkobjmask, which are described in more details in the following paragraphs.

2.3.6.1 Computing the dithering pattern (task irshift)

The task irshift is used for computing the dithering pattern among the nine images belonging to the same cube. The user has to visually identify and mark with the cursor at least one object common in all frames. From the (x_i,y_i) position of the object in the i-th frame, the task computes the relative shift $(\Delta x_{ij},\Delta y_{ij})$ of the frame i with respect to a (user selected) reference frame j. The results are stored in a two–column (Δx,Δy) ASCII file with nine lines, the j-th line containing the $(\Delta x_{jj},\Delta y_{jj}) = (0,0)$ relative shift of the reference frame with respect to itself.

The dithering pattern is produced by offsetting the telescope between consecutive pointings. Since all four detectors are simultaneously moved, the relative shift $(\Delta x_{ij},\Delta y_{ij})$

should be independent of the chip with which the i and j frames have been obtained. In other words, assuming that for a given tile it is possible to execute irshift on one chip only, the measured dithering pattern should equally be describing the dithering pattern for the three other chips.

The above statement is valid only in the most simple case, where all four detectors are perfectly aligned. In a more general case, where a detector is rotated with respect to the rest, more complicated transformations have to be used. The position of an object in any of the arrays can be described by a pair of (x,y) pixel coordinates. These coordinates form part of a grid defined by a Cartesian system of the form (OX^iY^i) ($i = 1, 2, 3, 4$), whose starting point O coincides with the (0,0) pixel at the lower left corner of the i detector and the axes OX^i and OY^i are perpendicular to each other. If the four Cartesian systems, describing the WIRC detectors, have exactly the same orientation, then indeed the dithering pattern measured in one of the chips will represent the move of a star in any of the other three detectors. For a more general case, however, where any detector i is rotated by an angle θ with respect to a reference detector j, it is necessary to calculate the transformation from one Cartesian coordinate system to the other. Since in all four arrays the reference point (0,0) falls on the same position, we can consider that the Cartesian systems are not translated and that there is only a rotation term to be calculated. Considering that detector #1 is the reference system, the position (x′,y′) of an object in the Cartesian system (OX^iY^i) ($i = 2, 3, 4$) will be related to its position (x,y) in the system (OX^1Y^1) according to the following relation:

$$x' = x\,cos\theta + y\,sin\theta \tag{2.2}$$

$$y' = -x\,sin\theta + y\,cos\theta \tag{2.3}$$

Based on the above equations, the angle θ is given by:

$$\theta = arcsin\left(\frac{y\,x' - y'\,x}{x^2 + y^2}\right) \tag{2.4}$$

To measure the rotation of detectors 2, 3 and 4 with respect to detector 1 we used the irshift results for the cases where it was possible to run the task on all four cubes of the same tile.[5] The above study was performed for 48 tiles, for which it was possible to run irshift on all four cubes. The average rotation angles (in rad) calculated are:

$$\theta_{21} = +0.0015 \pm 0.0027$$
$$\theta_{31} = -0.0195 \pm 0.0023$$
$$\theta_{41} = -0.0074 \pm 0.0032$$

[5]Since the reference image j of cube 1 did not necessarily coincide with the reference image k of cube M (M=2,3,4), the relative shifts (Δx_{ik},Δy_{ik}) for chips #2, #3 and #4 were properly recomputed in order to bring them to the same reference image j of chip1 #1.

The large error for chip's 2 rotation angle suggests that chips #1 and #2 have the same orientation, i.e. measuring the dithering pattern in one of the two chips is equivalent to knowing the pattern of the other. In cases where it was not possible to run irshift on cubes originating from chips #3 and #4, due to the absence of at least one common object in all dithered frames, a small program written by myself was used to compute their dithering pattern, starting from the results obtained with one of the first two chips.

2.3.6.2 Preliminary co-addition of the dithered images (task ircoadd)

As soon as the dithering pattern has been measured, one can use the task ircoadd in order to produce a coadded image. The input to this task is a catalog with the names of the nine sky–subtracted and residual–cleaned images and the name of irshift's output. Because of the centering of the nine–step dithering pattern around the first image (see Fig. 2.4), ircoadd will produce a final image (approximately) homogeneously "expanded" toward all directions with respect to the first of the dithered images. For a dithering step of $20''$ the size of the coadded frames was $\sim 1200 \times 1200$ pixels.

2.3.6.3 The object masks (task mkobjmask)

The purpose of mkobjmask is to identify and mask the objects present in the coadded images. At a first step the task identifies the pixels with intensities exceeding the background level in the coadded frame by a given threshold[6]. It then produces a binary mask, setting to zero all the pixels above this threshold. The task masklet uses this binary mask and the ASCII file with the relative shifts produced by irshift and, starting from the coadded frame, calculates the position of the objects in the dithered frames. In this way nine 1024×1024 pixel binary masks, each one associated with a dithered frame, are produced. The name of the associated mask is written to the image header of the *initial* frames, i.e. the flat–fielded frames with 112 seconds total integration time. Once this step has been performed, the user has to re–execute the steps described in paragraphs 2.3.5.1 and 2.3.5.2. In this way a more precise subtraction of the sky background and its residuals will be performed, free from the presence of any object that might bias the statistics.

A flow chart of the first four steps of the main data reduction is presented in Fig. 2.17. The dithered frames are sky–subtracted, split in quadrants, the quadrants are cleaned for the X–direction sky residuals, reassembled to form a single 1024×1024 pixel frame, masked for objects and re–inserted in the pipeline. The dashed line separates the logical steps of the reduction (left part), from the IRAF/PHIIRS tasks used (right part).

[6]As the background in the sky–subtracted frames is approximately zero, the threshold is based upon some multiple of the *rms* noise in the coadded frame

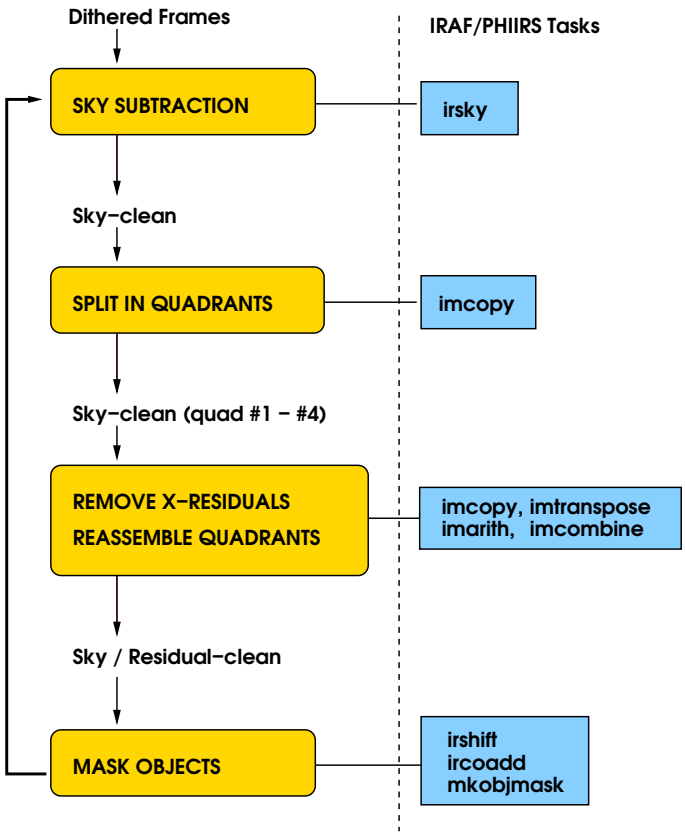

Figure 2.17: Flow chart describing the first part of the data reduction pipeline. Leftward to the dashed line the logical steps of the pipeline are shown, rightward the IRAF/PHIIRS tasks used.

2.3.7 Striping

When the number of electrons, produced by the in–falling photons, exceeds the detector's capacity[7], pixels experiencing this effect will be saturated. As a result, during the sky–residual cleaning (paragraph 2.3.5.2), the excess in counts of these pixels will bias the statistics, producing low–level artifacts.

As in IR detectors each array element is read individually (see paragraph 2.2.1), the cause of the striping is not similar to that of CCDs, where adjacent pixels are flooded by the excess in electrical charges in saturated pixels. Nevertheless, our study showed that, depending on the magnitude of the objects, the artifacts produced by striping can exceed the dimensions of a single quadrant. Therefore, for treating the sky and background–residual clean frames for this effect, it was necessary to work on the 1024×1024 frames, instead of the quadrants.

The destriping was performed by fitting a cubic spline along the X–direction and then

[7]For WIRC it is $\sim 60,000\, e^-$.

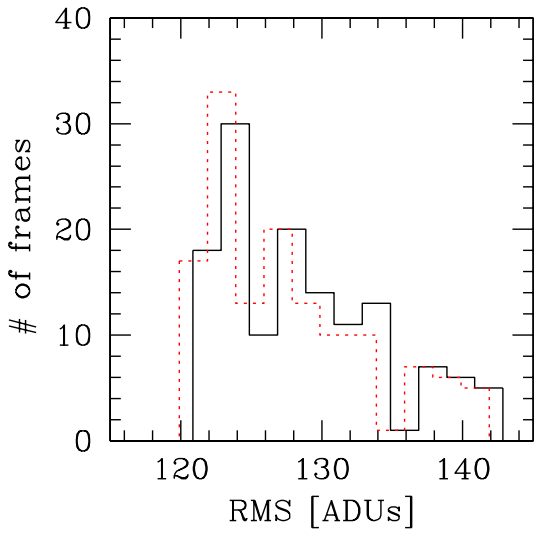

Figure 2.18: RMS histograms for 135 dithered, sky–cleaned and background residual–cleaned frames. The solid and dotted lines correspond to the frames before and after the destriping, respectively.

along the Y–direction of the object–masked frames. The impact of the destriping on the RMS of the science frames is demonstrated in Fig. 2.18. The results presented are based on the statistics over 135 frames, with the solid and dotted lines corresponding to the statistics before and after the destriping, respectively. The average increase in the RMS of the destriped frames was less than 1%, meaning that we have successfully removed the systematic features without increasing significantly the random noise.

2.3.8 Identification of cosmic rays and additional outliers

Once all sky residuals have been properly removed and the Bad Pixel (BPM) and Object masks (OM) have been created, the cosmic rays (CR) were removed from the individual dithered frames using PHIIRS's task crzap. The task detects the cosmic rays by (a) applying a median filter to the science image and (b) subtracting the filtered image from the original one. The pixels with values exceeding the background by a given factor are then characterized as possible candidates, and if they have not been flagged as objects they are then replaced by the local median.

Apart from producing CR–clean frames, it is also possible to produce a cosmic ray binary mask using crzap. This mask, combined with the binary masks described in previous steps of the data reduction, can help in characterizing the majority of the outliers of the pixel value distribution in each frame. However, the pipeline–reduced science dithered frames will still suffer from a few problematic pixels that the algorithms have failed to

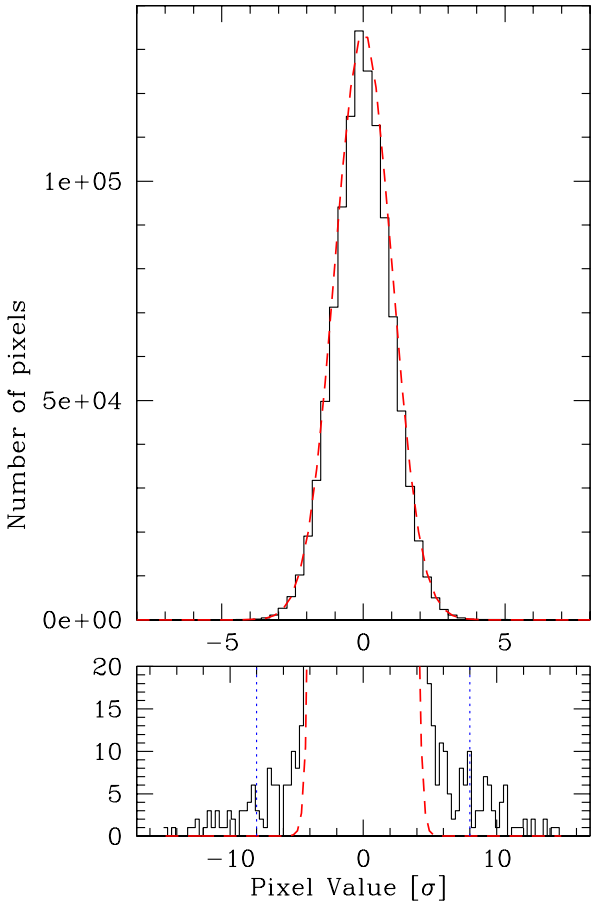

Figure 2.19: Pixel value distribution of a pipeline–reduced coadded frame, with the bad pixels, pixels with cosmic rays and object–pixels masked. *Upper panel:* The pixel value distribution can be well fitted by a Gaussian (dashed line). *Lower panel:* Zoom–in of the upper panel. All values outside the $\pm 8\sigma$ limits (vertical lines), are considered to be additional bad pixels and cosmic rays which were not previously detected.

detect. For identifying these outliers, the bad pixels, the pixels with cosmic rays and the object–pixels were masked and the RMS of the remaining pixels were computed. The intensity of all non–masked pixels was rescaled in terms of RMS and the pixels with values outside the $\pm 8\sigma$ range were considered to be additional pixels that had to be masked. The 8σ value was decided on the basis of number outliers as a function of threshold. Once the identification of these pixels was achieved, a combined pixel mask (CPM), with the position of the bad pixels, the object–pixels, the cosmic rays and the outliers was generated for each dithered frame.

The characterization of the outliers is illustrated in Fig. 2.19, where, for the sake of clarity, we have plotted a histogram of the pixel distribution of a sky–cleaned, corrected for sky–residuals and destriped coadded frame (instead of a dithered frame). The upper

panel shows the pixel value distribution, having the bad pixels, object–pixels and cosmic ray pixels masked. The histogram of the unmasked pixel values follows a Gaussian distribution, marked as a dashed line. The lower panel is a zoom–in of the upper panel. The dotted vertical lines indicate the 8σ cutting limit.

2.3.9 The end of the pipeline

The detection of the outliers marks the end of the main part of the pipeline. Thanks to the tests performed during the individual steps of the data reduction, we managed to build an efficient pipeline, which can correct for systematic effects, while maintaining random the noise properties, and the uncertainties introduced by our corrections to a minimum.

A flow chart of the second part of the pipeline line is presented in Fig. 2.20. The individual dithered exposures are masked for objects, bad pixels and cosmic rays, the sky background and its residuals are removed, the frames are destriped, the outliers are detected and the new frames, together with the updated bad pixels and object masks are re–introduced in the pipeline, for producing a new version of the dithered frames.

2.3.10 Construction of a combined science frame

The construction of a single coadded frame from the dithered frames has already been mentioned in paragraph 2.3.6.2. The input to the task ircoadd is a list of frames with the following characteristics: (a) the sky and its residual pattern have been removed (b) the name of the associated combined bad pixel mask has been written in their headers and (c) the relative shift (Δx_{ij},Δy_{ij}) of any i-th frame with respect to a reference frame j has been computed using the task irshift. ircoadd will produce a final image based on the relative shifts, whose size will be larger than the individual frames because of the nine–step dithering pattern (see Fig. 2.4).

Because of the different noise properties of the dithered images it is essential to minimize the effect of the noisier frames. By using weights, the coadded frame will have an improved signal–to–noise ratio compared to a coadded frame constructed from non–weighted images. The weights used are of the form:

$$w_i = \frac{1}{rms_i^2} \quad \text{with} \quad \sum_{i=1}^{9} w_i = 1 \tag{2.5}$$

where rms_i^2 represents the variance of the i-th frame of each cube ($i = 1...9$). ircoadd internally calls imcombine (see section 2.3.1) for producing the coadded frame. imcombine's main parameters used were scale=median, rejection=2 highest/2 lowest and combine=average.

An example of the initial, intermediate and final steps of the data reduction is displayed in Fig. 2.21. The upper left panel shows an image at the very early stage of the pipeline:

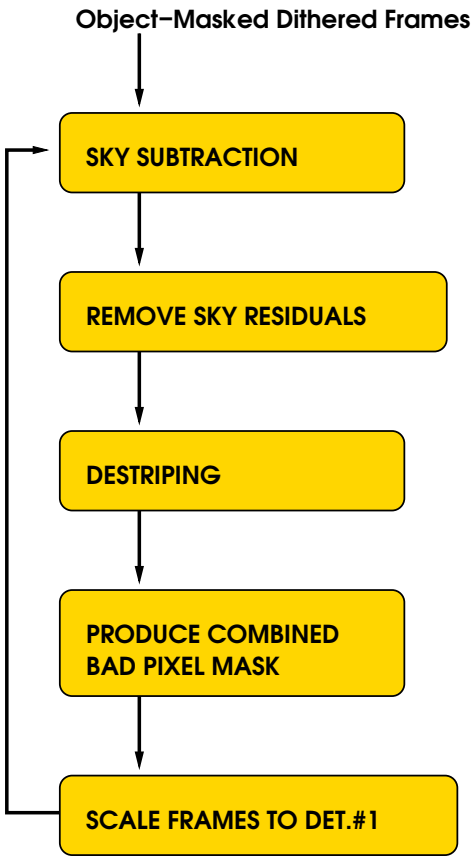

Figure 2.20: Flow chart of the second –and final– part of the pipeline. The dithered frames are masked for objects, bad pixels and cosmic rays, before being introduced in the pipeline. Once it has arrived to its end, the dithered frames together with the updated pixel masks are re–introduced in the pipeline, to produce a new version of cleaned dithered frames.

four 28–sec images have been flat-fielded and coadded, to produce a single dithered frame of 112 seconds total integration time. The upper right panel shows the same frame after it has been fully processed by the pipeline. Nine cleaned dithered frames are weighted, shifted and put together, in order to produce a single coadded frame, shown at the lower right panel. The black box indicates the region on the sky covered by the dithered frame of the upper right panel. The exposure map associated with the specific coadded frame is shown on the lower left panel. As the telescope points on the sky according to a specific dithering pattern, the region covered by the coadded frames will have been covered at different depths, indicated by the different numbers (in seconds).

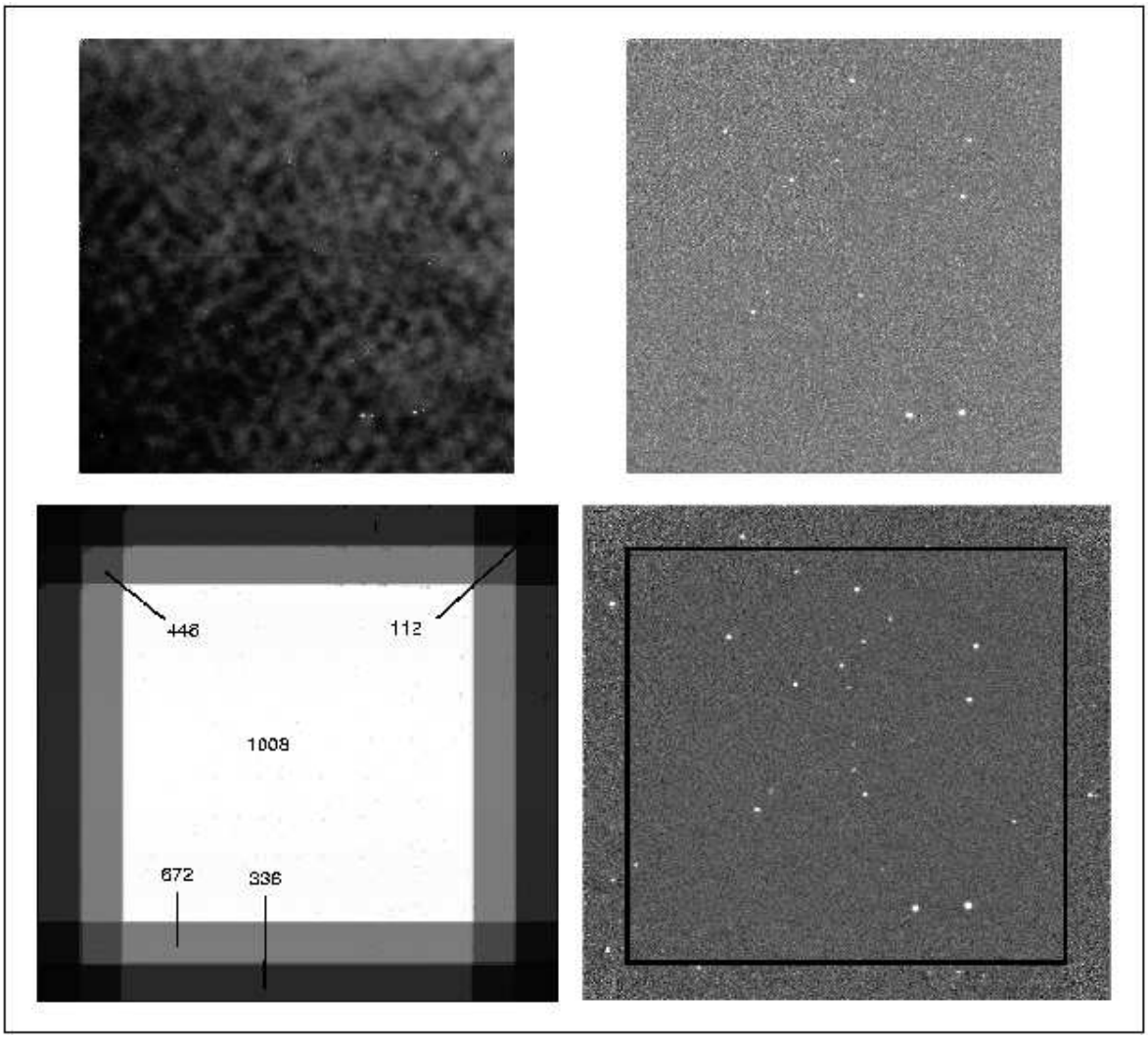

Figure 2.21: An example of the initial, intermediate and final steps of the data reduction. *Upper left panel:* A 112–sec dithered frame, just before being introduced in the pipeline. *Upper right panel:* The dithered frame has been fully processed by the pipeline. *Lower right panel:* Coadded frame, produced by combining the nine dithered frames belonging to the same cube. The black box indicates the region on the sky covered by the dithered frames of the upper right panel. *Lower left panel:* Exposure map associated with the coadded frame. The exposure time, in seconds, corresponding to each region is also shown.

2.3.11 Construction of a combined exposure map

For every science frame it is necessary to produce an exposure map that will allow, in the final mosaic image, to associate every pixel with its total integration time. Using the image headers, ircoadd produced FITS images whose pixel values were set to 112

ADUs, the total integration time of any dithered frame. The individual exposure maps were appropriately weighted and, together with the ASCII files describing the dithering pattern of their associated science frames, were introduced in ircoadd, in order to produce for each combined–weighted frame its corresponding exposure map.

2.3.12 Photometric calibration

The purpose of the subsection is to explain the steps followed for photometrically calibrating the Las Campanas data. This needs to be done in order to correct for three different effects, that make the ensemble of the ensemble of the reduced images quite inhomogeneous. The origin of this inhomogeneity is related to the instrumentation used for performing the observations, to the strategy followed for covering the LCO area and for the variation in the observing conditions: More analytically,

- The dithered frames were taken using four detectors having similar, but not identical, properties.
- The "building blocks" of this work, the coadded frames, were generated using 9 dithered frames taken under different airmasses.
- The observations were conducted during four nights with quite varying photometric conditions.

In the following paragraphs we will explain the operations performed for correcting the data for the above effects.

2.3.12.1 Rescaling the WIRC detectors to the same sensitivity level

Even if the four WIRC detectors share the same technical characteristics, they do not necessarily have the same sensitivity. In a mosaic image, however, frames obtained with different detectors will have to be processed so that uniform photometry can be extracted from the final mosaic. Bringing all detectors to the same sensitivity level is essential to produce a uniform data set. Six standard stars (Persson et al. 1998), observed on October 26, were used for this purpose. Each standard star was first put in the field of view of chip #1 and five exposures of a few seconds each were obtained. Then the telescope was slightly moved in order to sequentially position the star in chips #2, #3 and #4, each time executing the same observing loop. For each standard star, a mean flux F_i, averaged over the five observations obtained with chip i ($i = 1, 2, 3, 4$) was obtained. The conversion factor f_{ij} relating the mean flux F_i of any standard star, as it has been measured from the observations obtained with chip i, to its mean flux F_j measured on the frames obtained with chip j, is defined as:

$$f_{ij} = F_i/F_j \tag{2.6}$$

By taking the mean of the conversion factors obtained with the six standard stars for every chip, chips #2, #3 and #4 were scaled with respect to chip #1 according to the following formula:

$$S'_i = S_i / f_{i1} \tag{2.7}$$

where S_i and S'_i are the initial and the rescaled sensitivity of chip i ($i = 2, 3, 4$), respectively and the average (over the six standard stars) conversion factors computed are:

$$f_{21} = 1.111 \pm 0.004$$
$$f_{31} = 1.187 \pm 0.020$$
$$f_{41} = 0.959 \pm 0.003$$

2.3.12.2 Relative calibration of the dithered frames (task photcheck)

The time elapsed between the first and last frames belonging to a same cube is of the order of 25 minutes (net observation time plus overheads, mainly read–out time and telescope dithering). Over the time period when the observations are being taken, both the airmass and the sky transparency may vary. To correct the differential extinction among the dithered frames of each cube we used the task photcheck. The input to this task is the name of the coadded frame produced using ircoadd (subsection 2.3.10) and its corresponding exposure map (subsection 2.3.11), the list of the dithered frames used for producing the coadded frame and the ASCII file containing the relative shift between the dithered frames (see 2.3.6.1). The user has to indicate, in the coadded frame, the position of at least one object common in all frames. Using the relative shifts the task computes the object's (x,y) coordinates at each dithered frame and calls IRAF's task phot to measure the flux (in ADUs/sec) of the object at each time–divided frame. The user can modify, before executing photcheck, the parameters used by phot for performing the aperture photometry (e.g. the size of the aperture radius, the position at which the sky background will be measured, etc). The aperture radius chosen was $5''$, about three times the typical seeing conditions of the observations (FWHM $\sim 1.7''$). Based on the fluxes measured on the dithered frames, photcheck calculates the factor f_i by which the i-th frame ($i = 1...9$) has to be scaled such that the flux F_i of the star used for the relative calibration comes to the same level as its flux F measured from the coadded frame. Since for any telescope pointing i, belonging to the same tile ($i = 1...9$), four individual exposures I_j (one per detector) were simultaneously produced, it was sufficient to measure the relative scaling factors in one cube in order to correct for the effect of differential extinction for the other three sequences of dithered frames.

By the end of this step of the reduction a new version of coadded frames was created. The input to ircoadd was the sequences of nine clean (sky–subtracted, cleaned for the sky residuals and having properly masked the bad pixels) dithered frames, calibrated to the sensitivity of chip #1 and properly rescaled to correct for the differential atmospheric extinction.

2.3.12.3 Calibration to the K_s (Vega) system

The data of our run were obtained under very different conditions, since (a) during the same night the observations were being carried out at various airmasses and (b) the atmospheric transparency could experience variations from one night to another, or even during the same night. The final calibration that rescaled the data to the level of the K_s–band (Vega) photometry was performed using the observations of the standard stars observed on the night of October 26. Standard stars were also observed on October 27, but the night was not photometric, thus those observations were of no use.

The formula relating the instrumental magnitude of a star, m_i, with its flux F (measured in ADUs/sec) is:

$$m_i = -2.5 \log F \tag{2.8}$$

The instrumental magnitude m_{io}, corrected for any atmospheric effects, is related to m_i according to the formula:

$$m_{io} = m_i - k' X + k'' C X \tag{2.9}$$

where k' is called the first order (or primary) extinction coefficient and X is the airmass of the observations. The last term, $k'' C X$, is the product of the second order extinction coefficient k'' with the color index of the star C and the airmass X, and describes the way by which stars of different colors experience the atmosphere's absorption.

To measure the primary extinction coefficient one needs to observe at least one standard star during each photometric night, at various airmasses. Implementing the first order extinction correction on the objects present on any frame is straightforward, since one needs to know k' and the airmass at which the observations took place.

The measurement of k'' demands a rather time-consuming procedure, with observations of a large number (~ 10) of photometric standards at various airmasses, covering a wide color range. Given the requirements of our primary observing program, such a large amount of time could not be dedicated to measure k''. On the other hand, the correction of the instrumental magnitudes of the objects present on a science frame for second order extinction effects is not possible unless we know their color index C, which, a priori, is not available. Thus, the second order extinction correction ends up to be a rather complicated procedure.

A very good analysis regarding the second order atmospheric effects, when observing in the infrared, and the complexity of modeling its impact on the photometric measurements, is given by Bailer–Jones & Lamm (2003) and references therein. For a model of the Earth's atmosphere with a relatively low water column density, the differential extinction between a cool giant ($T_{eff} = 4000$K) and Vega ($T_{eff} = 9650$K) is 0.002 magnitude for the K–band. Increasing the water density by a factor of 35 changes the differential extinction to 0.001. From the above we conclude that the correction for the second order extinction is

negligible with respect to the uncertainty due to the imperfections in the sky-subtraction, the detector instabilities and the zero–point correction. Hence, we finally proceeded with the photometric calibration by simplifying Eq. (2.9) to:

$$m_{io} \approx m_i - k' X \tag{2.10}$$

The formula relating the instrumental magnitude of any star, corrected for any atmospheric effect, m_{io}, to its literature value m_l, for a given filter, is:

$$m_l = m_{io} + ZP \approx m_i - k' X + ZP \tag{2.11}$$

where ZP is called zero point and is an instrument–linked (i.e. telescope + detector + filters) additive term. In contrast to the extinction coefficient, which has to be computed every night, the zero point is considered to be constant for a period of several months, under the assumption that the observing configuration has *not* experienced any modification (e.g re–aluminization of the mirror, change of detector, etc).

The extinction coefficient and the zero point can be computed through a system of linear regressions of the form:

$$y = a\,x + b \tag{2.12}$$

with

$$y = m_l - m_i$$

and x being the airmass. The above system can either be represented by multiple observations of the same standard star at different airmasses, or, equivalently, by observations of various standard stars at a certain airmass. As explained in paragraph 2.3.12.1, each of the six standard stars was sequentially observed by all four detectors. For computing the extinction coefficient and the zero point, four independent systems of six equations each were solved. As expected, since the data obtained with chips #2, #3 and #4 were rescaled to the sensitivity of chip #1 (see paragraph 2.3.12.1), the four systems gave, within the 1σ error, the same zero point. The extinction coefficient solutions were also statistically equal and in agreement with the typical Las Campanas coefficient for the K_s–band (0.08 magnitude/airmass, Persson et al. 1998). Therefore, the extinction coefficient was fixed to this value and a system of 24 linear regressions was solved over ZP. The mean zero point $\overline{ZP}$ calculated for the night of October 26 was:

$$\overline{ZP} = -2.392 \pm 0.003$$

Fig 2.22 shows the magnitude difference $\Delta m = m_l - m_i$ as a function of the airmass, for the measurements of the standard stars obtained by all four chips on the night of October 26. The slope of the linear regression fit is equal to $a = 0.08$ and the constant term $b = -2.392$.

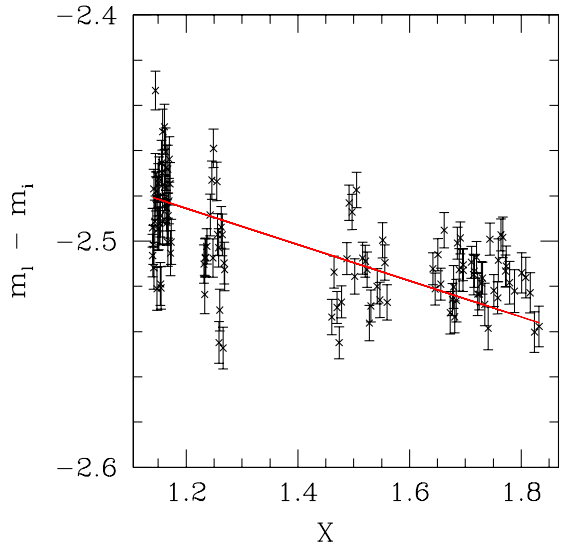

Figure 2.22: Extinction curve for the night of October 26. Six standard stars were observed at various airmasses, by all four detectors. The solid line corresponds to a linear regression with a slope equal to 0.08 and a constant term of −2.392.

Due to the amount of time lost because of the mediocre weather conditions, the time scheduled for calibrating purposes was finally allocated to the main observing program. As a consequence, no standard stars were observed on the photometric night of October 28. Due to the lack of observations of standard stars, the calibration of the data obtained during that night was achieved using the Two Micron All Sky Survey archive (Skrutskie et al. 2006). The 2MASS project has uniformly scanned the entire sky in three near–infrared bands, namely in J (1.25 μm), H (1.65 μm) and K_s (2.17 μm), using two highly-automated 1.3 m telescopes. Each telescope was equipped with a three-channel camera, each channel consisting of a 256 × 256 array of HgCdTe detectors, capable of observing the sky simultaneously. The spatial resolution obtained with the specific configuration was 0.2 ″/pixel. For the K_s–band, the 10σ sensitivity limit for the point–like sources is 14.8 (Skrutskie et al. 2006). The survey operations were completed in 2001 and the 2MASS data products were available at the Infrared Processing and Analysis Center (IPAC, http://www.ipac.caltech.edu/2mass/). Fig. 2.23 shows a K_s–band image downloaded from the 2MASS archive, whose central (RA,Dec) coordinates coincide with those of the coadded frame shown in the lower right panel of Fig. 2.21. The region of the sky covered by the latter is indicated by the gray box. North is up, East is to the left.

Because of the science objectives of the 2MASS survey, its limiting magnitude is much brighter than that of our observations. For calibrating the observations obtained on the night of October 28 we only used stars with $K_s \leq 14.5$, for which the error is lower than 0.1 magnitude. Fig. 2.24 shows the error in the 2MASS photometry as a function of K_s–band magnitude.

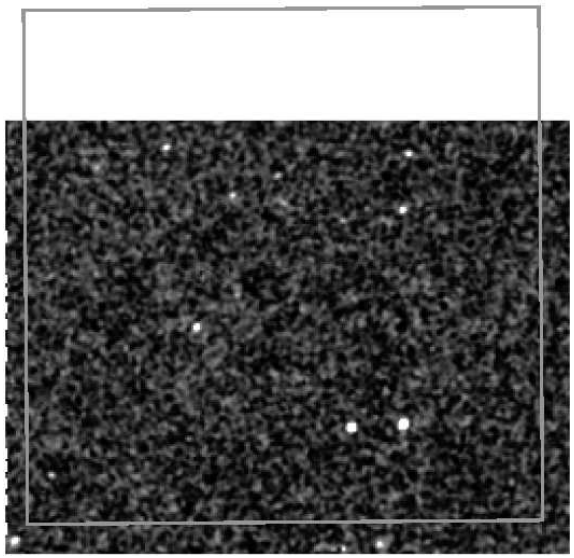

Figure 2.23: K_s–band image downloaded from the 2MASS archive, whose central (RA,Dec) coordinates coincide with those of the coadded frame presented in Fig. 2.21. The region of the sky covered by the latter is indicated by the gray box. North is up, East is to the left.

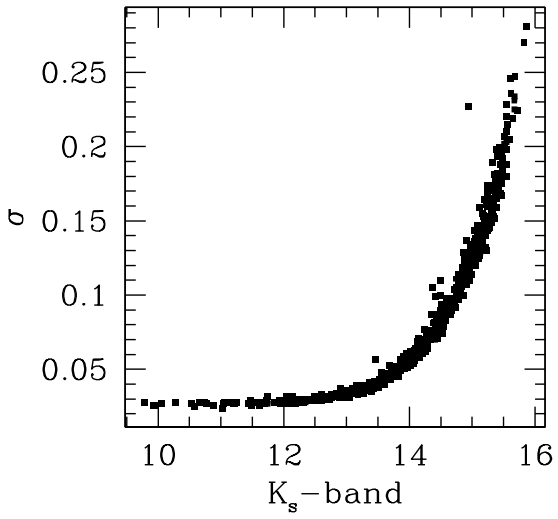

Figure 2.24: Standard deviation as a function of the K_s–band magnitude for the 2MASS survey (Skrutskie et al. 2006).

The celestial coordinates (RA,Dec) for every new pointing of the telescope were registered in a log–file. These coordinates refer to the center of WIRC's field and correspond to a point found in the gap between the four WIRC detectors (see Fig. 2.3). From the WIRC's configuration it is possible to *approximately* compute the (RA,Dec) coordinates of the field covered by each detector. In this way, for each tile, we succeeded in orienting its associated (four) coadded frames on the sky. Via the available 2MASS interface, we downloaded FITS files and ASCII catalogs with the position and the K_s magnitude for the regions of the sky covered by our observations. By visually comparing our images with the 2MASS data we succeeded in matching ~ 100 common point–like objects. By assigning the slope of the linear regression described by Eq. (2.12) to 0.08, the typical value for the Las Campanas extinction coefficient, we solved equation (2.11) over the zero point. The obtained solution was found to be statistically equal to the zero point computed for October 26, confirming that the night of October 28 was also photometric.

Since the nights of 25/10 and 27/10 were not photometric, it was impossible to use the methods described in the previous paragraphs for calibrating photometrically the data of each night as an ensemble. The only way to perform a calibration would be to work on each tile (i.e. four coadded frames) independently (tiles 05–19 and 37–46 for the nights of October 25 and 27 respectively). The 2MASS archive was once more of great help. For each tile we visually identified the point-like sources common between the 2MASS and our coadded frames. From the magnitude differences between the instrumental magnitudes (coming from our observations) and the values in the 2MASS catalog, it was possible to compute an average magnitude difference for each tile. This mean value represented the zero point by which the observations of each tile would have to be shifted for matching the 2MASS measurements. Fig. 2.25 shows the zero point variations, as a function of the tile identification, for the two non–photometric nights.

The magnitude difference between the 2MASS minus the Las Campanas calibrated magnitudes, as a function of the latter ones, is shown in Fig. 2.26. The graph is based on some hundreds of objects matched between the two catalogs. The higher Δmag uncertainties and scatter, for the fainter magnitudes, is dominated by the photometric uncertainty of the 2MASS measurements, since for magnitudes fainter than 14.5 we are very close to the survey's detection limit.

Fig. 2.27 shows the magnitude difference between the 2MASS and LCO magnitudes as a function of the (J-K) 2MASS color index. No color trend is visible, reassuring that the fields that have been photometrically calibrated using the 2MASS catalog are free from any color–related effects.

2.3.13 Astrometric calibration

A general description on how a two dimensional array can be represented by a pixel coordinate system has been given in section 2.3.6.1. All the information relating the physical elements of the detector, its pixels, with the Cartesian coordinate system is

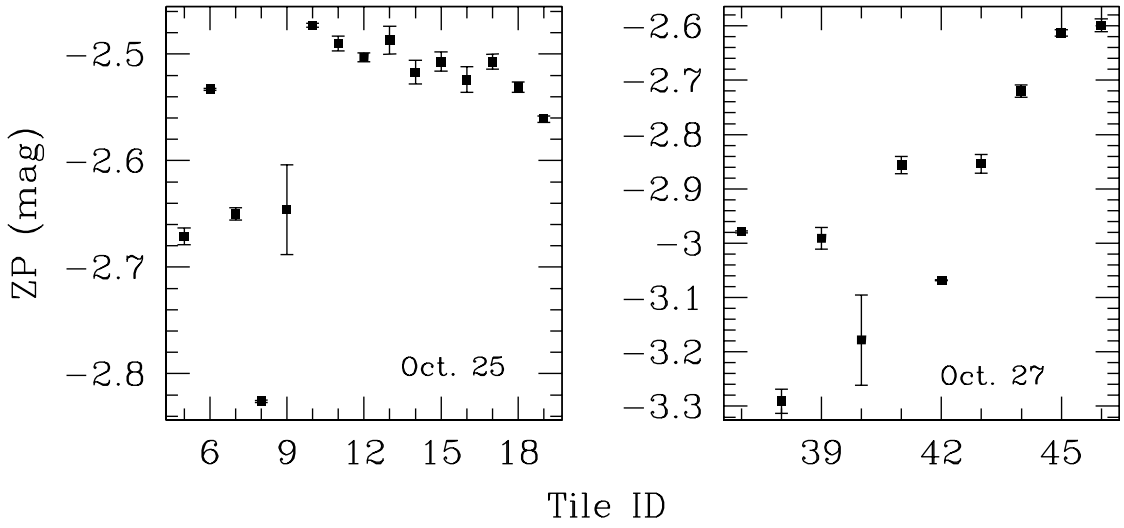

Figure 2.25: Zero Points for the tiles $05-19$ and $37-46$, corresponding to the nights of October 25 & 27 respectively.

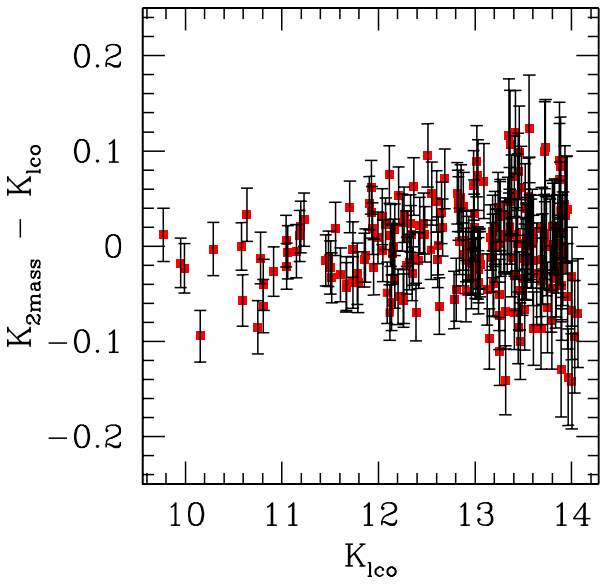

Figure 2.26: Magnitude difference as a function of the Las Campanas K_s–band calibrated magnitudes. The larger Δmag uncertainties and scatter, for the fainter magnitudes, are explained by the fact that we are very close to the 2MASS detection limit.

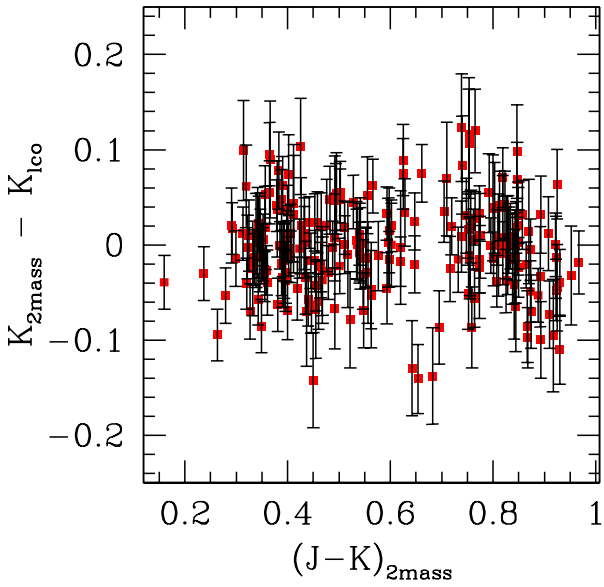

Figure 2.27: ΔMag versus the $(J-K)_{2mass}$ color index. No color trend is visible, ensuring that the Las Campanas fields calibrated through the use of 2MASS objects are free from any color–associated effects.

stored in the image headers. By manipulating the image headers it is possible to re-define this relation, e.g. re-define the pixel that will represent the (0,0) reference point, or the quantity by which the Cartesian coordinates will be modified for a given translation in the pixel array. The astrometric calibration is the process through which any pair of (x,y) Cartesian coordinates will be associated with a pair of celestial coordinates (RA,Dec). By including, in an encoded way, this transformation in the image headers, all science frames can therefore be astrometrically calibrated.

As part of the XMM–LSS project, the region on the sky covered by the Las Campanas observations had also been observed with the Mosaic II CCD imager mounted on the 4 m Blanco telescope at CTIO observatory (Chile). The observations were obtained in November 2001, in the R and z'–band [8]. The final throughput for the R and z' bands, where the transmission of the filters is convolved with the quantum efficiency of the CCD camera and the typical atmospheric transmission, is shown in Fig. 2.28.

The Mosaic II camera consists of eight 2048×4096 CCD arrays. The scale at the center of the camera is $0.27''$/pixel, covering in total $36' \times 36'$ on the sky. Due to the large FoV of the detector, the pixel scale decreases quadratically by 6.5% out to the corners, an effect that has to be taken under consideration during the astrometric calibration process. The eight arrays are read by two amplifiers, generating 16 individual images at a single exposure. The Las Campanas field was covered with four pointings, with an average of

[8] The transmission curves of the filters can be found at the following URL: http://www.ctio.noao.edu/instruments/filters/filters_66.html

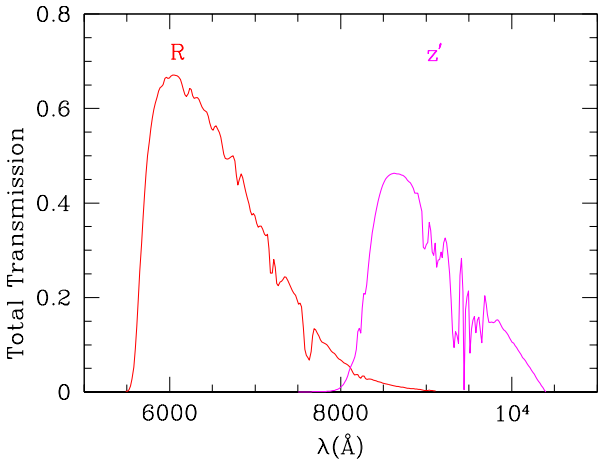

Figure 2.28: Final transmission curves for the CTIO R and z' filters.

three exposures per pointing. Therefore the total number of R and z'–band individual frames acquired was $\sim$ 190 ($2 \times 8 \times 4 \times 3$). Detailed information regarding the data reduction procedure can be found in Andreon et al. (2004).

The images were astrometrically calibrated using the USNO–2 catalog (Monet et al. 2003), with final goal to generate mosaic images of the observed area. Four z'–band frames, covering the LCO field, were kindly provided by Dr. S. Andreon (Osservatorio Astronomico di Brera, Milano, Italy) in order to astrometrically calibrate the K_s–band data. By visually comparing the Las Campanas and the CTIO images we created, when possible, an ASCII file for each coadded frame with the (x,y) coordinates of its objects, in the Las Campanas frames, and their corresponding (RA,Dec), coming from the CTIO data. The photocenters of the objects were measured using a Gaussian fitting. All possible candidates were visually inspected and we finally selected the most reliable ones, i.e. point–like, high signal–to–noise detections, whose counterpart in the CTIO data was not saturated.

The astrometric calibration of the coadded frames was performed using IRAF's task ccmap. The input to the program is the list of (x,y) pixel coordinates and their matched celestial coordinates, corresponding to the specific coadded frame. The program projects the celestial coordinates to a plane, according to a user–defined model. An extensive review on the basic concepts of projections and their representation in the FITS headers can be found in Calabretta & Greisen (2002). For astronomical applications, the *tan* projection works well for fields with size less than a degree. The result of the projection of the celestial coordinates is a new set of coordinates, (χ, η), called "standard". Once projected onto a plane, the program tries to compute the linear transformation between the pixel and the standard coordinates, according to the chosen model. The output of ccmap are the scaling factors (x_{scale},y_{scale}) and the rotation angles (x_θ,y_θ) of the x and y axis respectively, and the celestial coordinates (RA_{ref},Dec_{ref}) corresponding to the

reference pixel coordinates (x_{ref},y_{ref}) for the specific coadded frame. The user can interact during the fitting process with the program and reject values giving high residuals. Once the good transformation coefficients have been obtained, the program can store the above information (i.e. reference points, scales, rotations) in the image headers, following the current FITS conventions (Greisen & Calabretta 2002).

The model used by ccmap can be described by the following equations:

$$\chi = a + b\,x + c\,y \tag{2.13}$$

$$\eta = d + e\,x + f\,y \tag{2.14}$$

$$a = RA_o - b\,x_o - c\,y_o \tag{2.15}$$

$$b = +x_{scale}\cos\theta_x \tag{2.16}$$

$$c = -y_{scale}\sin\theta_y \tag{2.17}$$

$$d = Dec_o - e\,x_o - f\,y_o \tag{2.18}$$

$$e = x_{scale}\sin\theta_x \tag{2.19}$$

$$f = y_{scale}\cos\theta_y \tag{2.20}$$

$$\theta_x = O - S/2 \tag{2.21}$$

$$\theta_y = O + S/2 \tag{2.22}$$

where

- (χ, η) are the "standard coordinates" in the *tan* projection.
- (x_{scale},y_{scale}) and (θ_x,θ_y) are the scaling factors (in arcseconds/pixel) and the rotation angles (in degrees, measured counter-clockwise), for the x and y axis, respectively.
- (RA_o,Dec_o) and (x_o,y_o) are the origins of the celestial and pixel coordinate systems, respectively.
- O and S are the orientation of the x and y axes and their deviation from perpendicularity, respectively.

The above fitting geometry does not correspond to a physically meaningful model. Nevertheless, the computed coefficients can be interpreted as a combination of linear terms associating rotations and translations. If necessary, it is also possible to incorporate higher order corrections in the model. By default this was not the case, and the study of the results we obtained showed that the astrometric calibration was quite satisfactory without introducing any additional terms.

To reach an acceptable astrometric solution, ccmap needs at least three pairs of coordinates. Once the fitting has converged, the task gives the scaling factors (x_{scale},y_{scale}),

the rotation angles (x_θ,y_θ) and the origins of the two coordinate systems (RA_o,Dec_o) and (x_o,y_o) corresponding to the best solution for the specific frame. This information is stored in the image header of each frame, in an encoded way. A typical FITS header of an astrometrically calibrated frame is shown below:

```
====> file I18.c2w.coadd.1.fits (main) <====
NAXIS1      =    1239              / Axis length
NAXIS2      =    1215              / Axis length

CRPIX1      =    673.2222727       / Reference pixel
CRPIX2      =    515.3682889       / Reference pixel

CRVAL1      =    35.9055193        / Coordinate at reference pixel
CRVAL2      =    -4.94749403333    / Coordinate at reference pixel

CTYPE1      =    'RA---TAN'        / Units of coordinate
CTYPE2      =    'DEC---TAN'       / Units of coordinate

WCSDIM      =    2
WAT1_001    =    'wtype = tan axtype = ra'
WAT2_001    =    'wtype = tan axtype = dec'
RADECSYS    =    FK5
CD1_1       =    -5.54210285792E-5
CD2_2       =    5.54054737735E-5
CD1_2       =    6.22839780888E-7
CD2_1       =    7.642879082144-7
```

The NAXIS1 and NAXIS2 descriptors define the length (in pixels) of the x and y axes respectively. According to the FITS language, any coordinate system different from the pixel coordinate system is called "World Coordinate System" (WCS). The descriptor WCSDIM informs that in the given frame the x and y axes also contain information on an additional coordinate system, the type of which is provided by the descriptors WAT1_001 and WAT2_001 (a projection of type *tan*, the x–axis contains information on the right ascension, the y–axis on the declination). What, in projection terms, was called origin of the pixel coordinate system is called reference pixel in the WCS terminology. The role of the reference pixel is to indicate the array position that relates the two coordinate systems. The (x_{ref},y_{ref})=(x_o,y_o) coordinates are stored in the (CRPIX1,CRPIX2) descriptors. The corresponding values in the World Coordinate System (RA_{ref},Dec_{ref})=(RA_o,Dec_o) are stored in the descriptors (CRVAL1,CRVAL2). The rotation angles (θ_x, θ_y) and the scaling factors (x_{scale},y_{scale}) are encoded in the descriptors CDi_j and are actually products of the form:

$$CDi\_j = \sin\theta_i \times scale\_j$$

Chip No	scale$_x$ ($''$/pixel)	scale$_y$ ($''$/pixel)	$\theta_x(^\circ)$	$\theta_y(^\circ)$
#1	0.1991 ± 0.0001	0.1992 ± 0.0001	180.6173 ± 0.0214	0.5089 ± 0.0327
#2	0.1991 ± 0.0001	0.1991 ± 0.0001	180.7637 ± 0.0144	0.6571 ± 0.0228
#3	0.1987 ± 0.0003	0.1991 ± 0.0002	179.4886 ± 0.0959	-0.5988 ± 0.0472
#4	0.1986 ± 0.0003	0.1993 ± 0.0001	180.2802 ± 0.0283	0.1245 ± 0.0392

Table 2.3: Average values for the scaling factors and rotation angles for each of the WIRC detectors, computed using a high signal–to–noise data set.

The descriptor RADECSYS informs that the Equinox used for the astrometric calibration was J2000.

The quality of the astrometric solution strongly depends on the number of (x,y) and (RA,Dec) pairs in the ccmap's input file. This number varied between three to six objects per frame. However, since during the 2002 run neither WIRC's orientation on the sky was modified, nor any significant temperature variations (that could strongly alter the focus of the telescope) took place, a single solution, based on a high signal–to–noise sample, would better describe the scaling factors and rotation angles for each detector than the individual solutions. For obtaining such a solution we converted the Cartesian and celestial coordinates of the objects used for calibrating each coadded frame into relative values (x$_{rel}$,y$_{rel}$) and (RA$_{rel}$,Dec$_{rel}$), by subtracting the coordinates of the reference pixel (x$_{ref}$,y$_{ref}$) and (RA$_{ref}$,Dec$_{ref}$) of the specific frame, provided by ccmap. By merging to a single file all the relative positions obtained from the individual frames for each detector, we obtained a higher quality solution, based on some hundreds of objects. The CDi_j descriptors were updated accordingly and in this way a single solution for the scaling factors and rotation angles was imposed to all frames obtained with the same chip. Table 2.3 provides the average scaling factors and rotation angles, together with their standard deviation, for each of the WIRC detectors. The error on the scaling factors is of the order of 10^{-4}, i.e. for a typical coadded–image size of 1200 × 1200 pixels the error in Right Ascension and Declination, due to the scale uncertainty, is of the order of $\sim 0.12''$.

ccmap was successfully run for about 45% of the frames. For the cases where some of the reference objects in the LCO data were saturated in the CTIO images, or the Gaussian fitting could not converge for three point–like sources because of the low signal–to–noise ratio of the detections, we tried a different approach for calibrating the fields. As it has been already explained, the representation of any World Coordinate System on a pixel array is defined by an ensemble of descriptors. By updating the CD_ij descriptors of the non–calibrated frames we can impose the same scaling factors and rotation angles, but this is not enough for having the frames astrometrically calibrated. Without a reference point, the frames can not be associated with a position on the sky. Starting from the (x,y) and (RA,Dec) coordinates of at least one point–like source matched to the CTIO data, we computed the Right Ascension and Declination corresponding to the (x$_c$,y$_c$) pixel

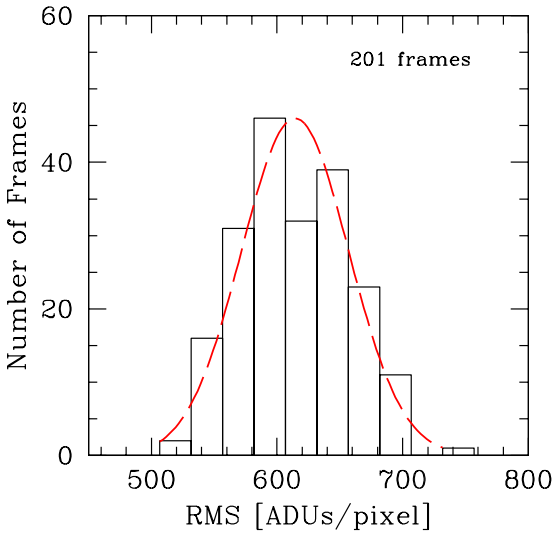

Figure 2.29: RMS histogram of the 201 astrometrically calibrated frames that were used for making the Las Campanas mosaic. A Gaussian fit, plotted as a dashed line, is also shown.

coordinates at the center of each frame. Using the center instead of the first pixel (found at one of the edges of the frame) as the reference point has the advantage of minimizing the average position errors due to the scaling factor uncertainties. When more than one Las Campanas point–like source could be matched to a CTIO detection, a mean central position was calculated. The percentage of frames that was astrometrically calibrated using this approach was $\sim 48\%$.

7% of the frames were not calibrated, either because the data were very noisy or because it was not possible to identify (or match) any point–like sources on the LCO data. Out of the 210 calibrated frames, 17 had to be partially masked, because at least one quadrant was much noisier than the others. Another 10 frames, belonging to tiles 41-49 and 61–63 (see Table 2.2), had to be left out from any further steps of the data reduction, because they were very noisy. Fig. 2.29 displays the RMS distribution of the 201 frames that were finally selected for the mosaicking, which can be well fitted by a Gaussian distribution, as demonstrated by the dashed red line.

Fig. 2.30 shows the difference in Right Ascension and Declination for some thousands of objects detected in the Las Campanas fields with respect to their matched coordinates in the z′–band CTIO data. The inner and outer circle correspond to the 1σ and 3σ errors, respectively.

As a final step, the image headers of the exposure maps were appropriately updated in order to have exactly the same astrometric solution stored in each science frame and its associated exposure map.

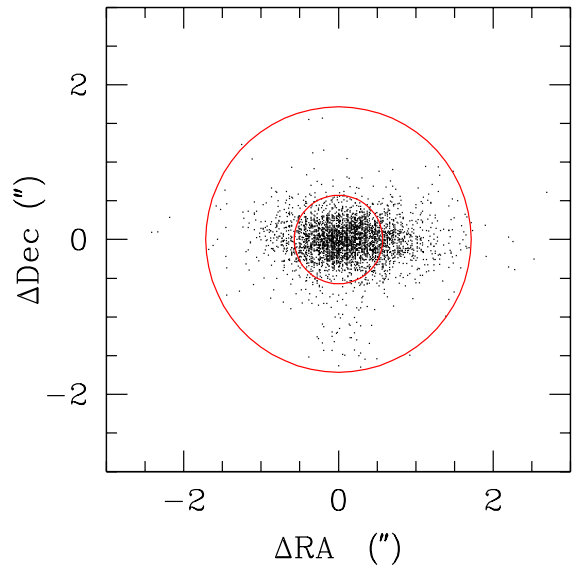

Figure 2.30: Right Ascension versus Declination diagram for some thousands of Las Campanas objects, with respect to their matched coordinates in the z′–band CTIO data. The inner and outer circle correspond to the 1σ and 3σ errors, respectively.

2.3.14 Science mosaics

The construction of the mosaics was performed using a sequence of four IRAF tasks, taken from the noao–artdata package, according to the following steps:

- Using the task mkpattern we created an empty image of $N \times M$ pixels. The N and M values should approximately correspond to the size of the final mosaic.

- ccsetwcs takes mkpattern's empty image and creates a World Coordinate System grid. The user has to define the pixel scale in the output image, the projection type, the Equinox of the coordinates and approximately the Right Ascension and Declination in the middle of the frame.

- wregister takes a list of the astrometrically calibrated frames and for every frame i, of $N_i \times M_i$ pixels, a new frame of $N \times M$ pixels will be created. The task computes the spatial transformation function required to register the input image to the $N \times M$ reference image. This is done by matching the (x,y) coordinates in the input image with the same grid of points in the reference image, using the world coordinate information stored in the two image headers. wregister computes the output image by evaluating the fitted coordinate surfaces and interpolating in the input image at the position of the transformed coordinates. The interpolation type chosen was the "drizzle" ,with a pixel fraction of 0.25 (i.e a bilinear interpolation with an oversampling of 4).

- imcombine takes all the wregister-ed images, and combines them to a single image. The user can define the type of combination (average, median) and the type of the pixel rejection.

The above four steps are schematically described in Fig. 2.31. The horizontal lines separate the individual steps. The empty box shown in the upper panel represents the empty $N \times M$ pixel frame, produced by mkpattern. ccsetwcs's output is shown in the second panel: the dashed lines represent the WCS grid and the cross the reference point, placed in the center of the new frame, having coordinates (x_{ref},y_{ref}) and (RA_{ref},Dec_{ref}). The third panel describes how wregister works: for each of the three astrometrically calibrated frames, represented by the small shadowed boxes, the task will produce a frame of $N \times M$ pixels into which the frames will be inserted via an interpolation algorithm. The white region in the large boxes corresponds to pixel values set to zero, the black region to the (interpolated) values of the individual frames. The lower panel represents imcombine's output, the mosaic–frames produced in step 3 are combined into a single mosaic.

Due to the large area covered by the Las Campanas observations (about 0.9×0.9 square degrees, or $16,000 \times 16,000$ pixels) working on a single mosaic would be quite cumbersome. It was therefore decided to construct four sub–mosaics, hereafter named as MosaicA, MosaicB, MosaicC and MosaicD. This splitting was based on logical criteria: the gap produced by the noisy tiles 41–49 served as the east-west separating line and the northern and southern sub–mosaics to be about the same size. Fig. 2.32 shows the final area covered by the Las Campanas Observations. Due to (a) imperfections in the telescope pointings and (b) the different rotation angle of each detector, it would be more realistic to consider that the ~ 0.8 square degrees on the sky were covered by a slightly irregular patch of cells, rather than a grid of observations perfectly aligned in the North–South and East–West directions. The gaps at the lower right and upper left part of the mosaic correspond to the area that would have been covered if a full night were not lost due to bad weather conditions. The empty regions in the central part of the mosaic are due to the noisy tiles 41–49, that were excluded from the mosaicking process. The dashed lines represent the virtual borders based on which the frames were distributed. The red, green, blue and magenta colors indicate the area covered by mosaics A, B, C and D, respectively. The overlapping regions formed by adjacent pointings are also visible.

2.3.15 Exposure Map mosaics

From what has been mentioned until now about the observational strategy followed (section 2.2) and the observing conditions during the observations (section 2.3.12.3), it is obvious that the four sub–mosaics will consist of frames taken under very different observing conditions (i.e. airmass, photometric or non–photometric nights). Additionally there will be overlapping and non–overlapping regions, the latter ones produced by telescope pointings taken with an elapse of time that spans from a few minutes to a day. In order to build a K_s–band catalog using the mosaics, and for studying its properties,

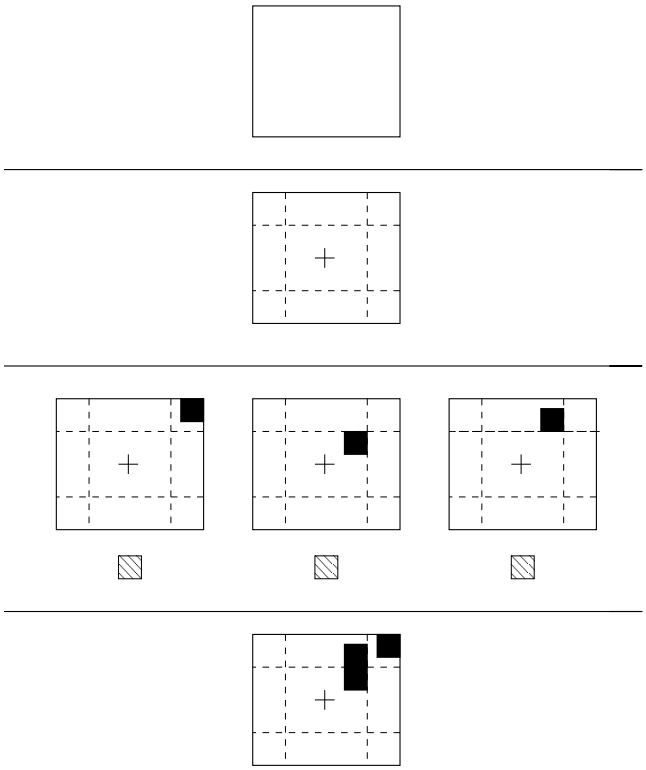

Figure 2.31: Schematic representation of the four–step procedure followed for producing a final mosaic image starting from the individual astrometrically calibrated coadded frames. Upper panel: the task mkpattern produces an empty image. Second panel: ccsetwcs defines a World Coordinate System grid (dashed horizontal & vertical lines) with a given scaling, reference pixel (marked with a cross) and projection type. Third Panel: wregister takes every coadded frame, interpolates its and registers the output image to a frame sharing the same properties (dimensions, coordinate grid) as the frame generated by ccsetwcs. Lower panel: The mosaics produced by wregister are combined into a single mosaic, by the task imcombine.

it is essential to associate every "science pixel" with a pixel–map that will contain the information on the total integration time, the zero point correction, etc. This can be done through the use of the exposure maps. The calibrated magnitude mag of an object observed at a given airmass X, on a night with extinction coefficient k and a zero point ZP is:

$$mag = -2.5 \log_{10}(\frac{F}{T}) - k\,X + ZP \tag{2.23}$$

where F is the flux of the object in ADUs, and T the exposure time of the observations. Eq. (2.23) can be re–written as:

$$mag = -2.5\ \log_{10}(\frac{F}{A}) \tag{2.24}$$

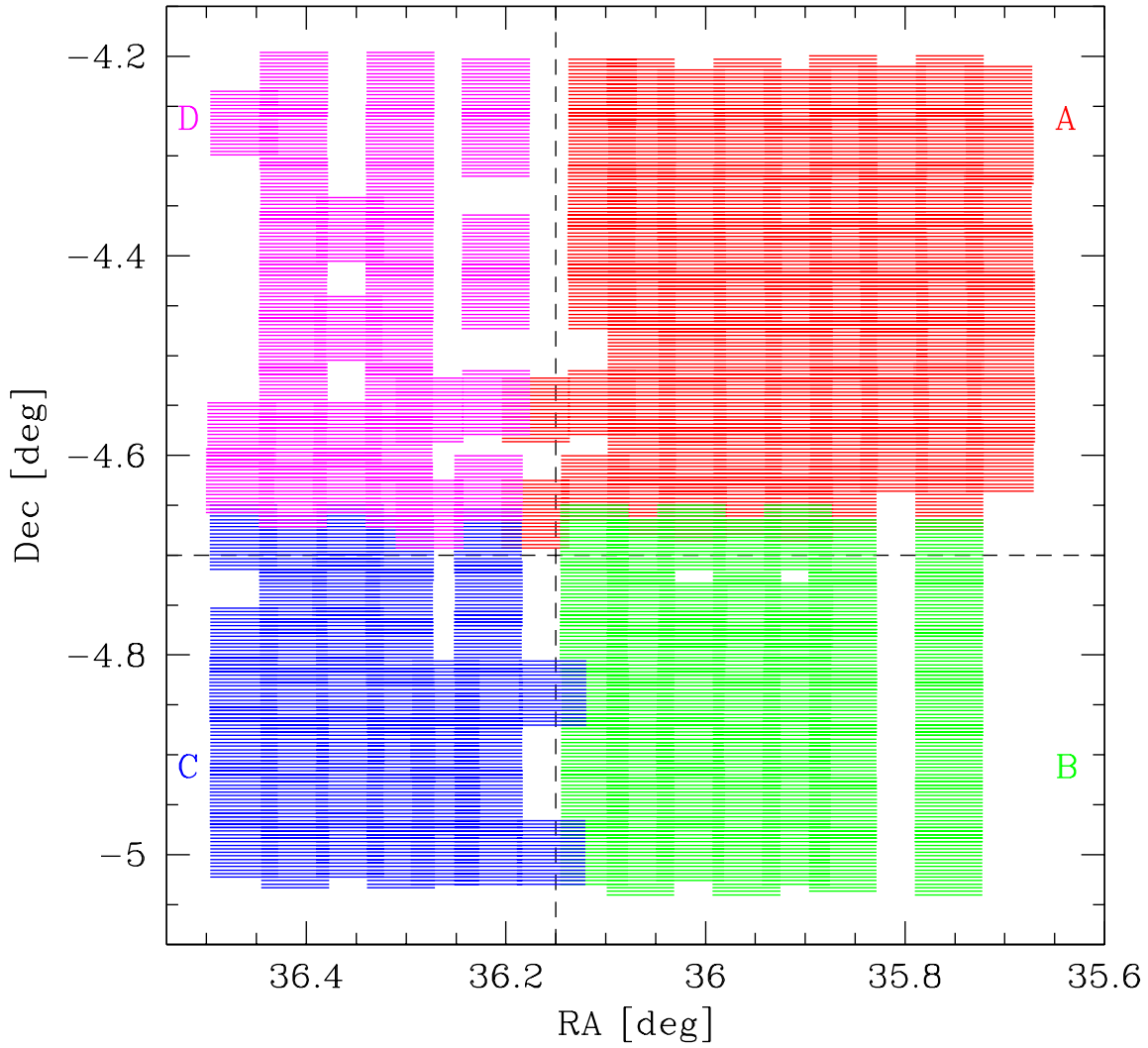

Figure 2.32: Mosaic image of ∼200 fields, centered at RA=$2^h24^m24^s$, DEC=4^o $36'$. The gaps in the lower right and upper left part of the mosaic correspond to the area on the sky that would have been covered if a full night were not lost due to the bad weather conditions. The empty regions in the central part of the mosaic are due to some very noisy frames, that were excluded. The dashed lines represent the virtual borders used for re–distributing the coadded frames, in order to form mosaicA (red area), B (green,), C (blue) and D (magenta). The overlapping regions formed by adjacent pointings are also visible.

where

$$A = T\,10^{0.4(ZP - k\,X)} \tag{2.25}$$

For the photometric nights of October 26 and October 28 the above equation becomes:

$$A_i = T_i\,10^{0.4(-2.392 - 0.008\,X_i)} \tag{2.26}$$

where T_i is the initial exposure map associated with the i-th coadded frame, A_i the new exposure map, and X_i the *mean* airmass at which the observations of the dithered frames used for constructing the coadded frame were obtained. For the non–photometric nights of Oct. 25 and Oct. 27, equation (2.25) is simplified to:

$$A_i = T_i\,10^{0.4\,ZP_i} \tag{2.27}$$

where ZP_i is the zero point corresponding to the i-th tile, as explained in paragraph 2.3.12.3. Based on Eq. (2.26) and (2.27), the exposure maps of the coadded frames were updated and were then distributed in order to form mosaic exposure maps. In this way, any magnitude measured on a time–divided mosaic (the result of the division of a science mosaic by its corresponding exposure map), will be directly expressed in terms of K_s (Vega) magnitudes.

Fig. 2.33 shows the final exposure map for mosaic A. The individual exposure maps have been corrected for extinction and zero point before producing the mosaic frame. Because of the technique used for the mosaicking (see paragraph 2.3.14), the mosaic image will contain areas where no observations were performed. These empty regions are marked with black color. The ratio between the deepest exposure maps (white regions) over the shallowest (dark blue) is of the order of three.

2.4 Detection Probability (Completeness Tests)

Making the four time–divided mosaics is the final step prior to the construction of the K_s–band catalog. The extraction of the photometry and positions of the objects was performed with SExtractor (Source Extractor, Bertin & Arnouts (1996)), a package whose main task is the building of catalogs from astronomical images. Via an input file the user control the object–detection threshold, the way the photometry will be performed, etc. Since SExtractor will be extensively mentioned in the following sections, we advice the reader to consult Appendix A for a more detailed description.

To appreciate the properties of the K_s–band catalog it is essential to understand the relation between the true properties of the area on the sky and the ones of the area registered on our data, since the sensitivity in the mosaics is spatially varying. The limiting magnitude at any pixel is a convolution of the detector's properties with the number of frames superposed, their noise properties and the total integration time. This

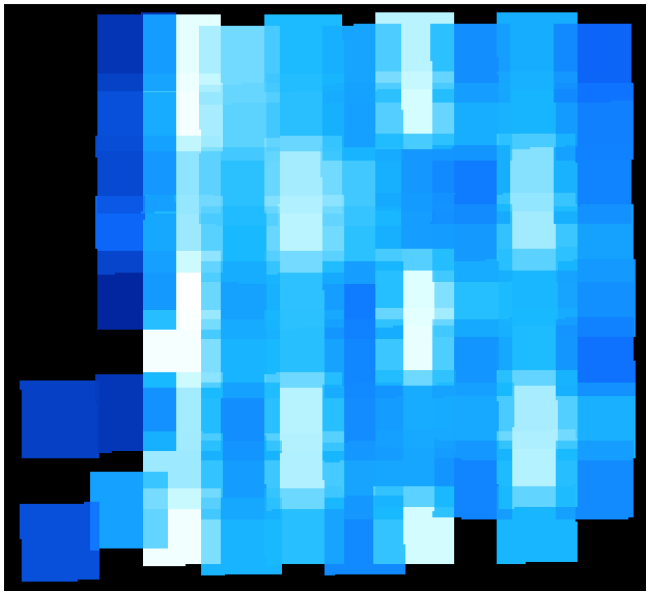

Figure 2.33: Final exposure map of MosaicA. The zero point correction and the extinction effects, are incorporated in the individual frames used for constructing the mosaic image. The overlapping regions formed by adjacent pointings are also visible.

study was performed using a code that was kindly provided to us by E. Labbe (PUC) and Dr. F. Courbin (Ecole Polytechnique Fédérale de Lausanne, Switzerland) that simulates the presence of (artificial) objects in our data. The basic concept of the simulations is as follows:

- For each mosaic the program creates a grid of $N \times M$ cells. The values for N, M depend on the size of the (rectangular) cell, defined by the user.
- An object of a certain profile and magnitude is inserted in each cell. For our study we chose the Gaussian distribution for describing the profile of the point–like sources. The positions of the objects are stored in a temporary file.
- The program calls SExtractor for checking whether there is a detection, within a certain radius, from the positions stored in the temporary file. If there is a detection and the measured magnitude lies within the $[m - 0.5, m + 0.5]$ region (where m the inserted magnitude), then the program flags the cell with one, denoting a successful detection. Otherwise, the flag gets the value zero, denoting a failure to detect.

The total number of objects that will be introduced in each cell is one of the user–defined parameters in the program's input file. For a total number of K objects, the steps described above will be repeated an equal amount of times, having the objects shifted in such a way in order to fill–in each cell with a uniform grid of $\sqrt{K} \times \sqrt{K}$ objects. The ratio of detections over the number of inserted objects is called *detection probability* and represents the probability an object of a given magnitude to be detected in the given cell.

The presence of real astronomical objects in the field produces a certain bias. Since the program produces a regular grid of cells into which it inserts objects, there will be some cases where the inserted objects will fall very close to, or even exactly on, pixels where real astronomical objects are registered. As a result, SExtractor might wrongly identify a real object as the inserted object, biasing the number of detections and the average magnitude of the detected objects. Additionally, the presence of saturated stars in the images might also affect the detection and photometry of the recovered objects. In order to avoid these effects, it is possible to associate each science frame with an object mask, i.e. a binary file where the pixel value will be set to zero in the "forbidden" regions (i.e. pixels occupied by objects) and one in the "permitted" ones.

Fig. 2.34 illustrates how the simulation code works for two cells with different characteristics. The program sequentially inserts 16 objects of magnitude m, indicated by the open circles. The crosses indicate SExtractor's detections. The shadowed region denotes a noisy region, where SExtractor fails to detect objects for a given detection threshold, the black region denotes a masked region (pixel values = 0) that should not be taken into account in the final calculations. In the left panel, SExtractor detects 14 objects out of 16, thus the probability of detecting an object of magnitude m in the given cell is $DP = 87.5\%$. In the right panel, three objects fall in masked regions, and SExtractor detects 13 objects out of 13, i.e. the detection probability is 100%.

As explained in Appendix A, there are two parameters controlling SExtractor's detections: the threshold above which the program searches for pixels that will be considered as objects (DETECT_THRESH) and the number of contiguous pixels above DETECT_THRESH in order to be considered a valid object (DETECT_MINAREA). Prior to detection SExtractor computes a mean sky value for each science frame. It then filters the sky–subtracted frame, in order to detect the objects. The type of filter is controlled by the parameter FILTER_NAME and its choice strongly depends on the type of sources we want to detect. A "Mexican Hat" is a very sharp filter, that makes the intensity at the wings of an object's profile drop very fast. It favors crowded fields with bright objects, but it is unrecommended for fields with faint objects, since they will be smoothed at the level of the background. On the other hand, a Gaussian filter, which favors the detection of faint objects, comprises the problem of blending in regions with high object density or regions where an inserted object falls very close to a real object.

Although the spatial resolution of the simulations is controlled by the size of the cells, it is the number of iterations (i.e. the incremental step) that will allow to know the detection probability within any cell with a given precision. Choosing a large cell size decreases the resolution but at the same time decreases the time needed for the simulations, allowing to experiment with different sets of SExtractor parameters (e.g. detection threshold, filter type). On the other hand, reducing the size of the cells not only increases the time needed for the simulations but also increases the risk of having stars covering a large part (or even the full area) of a cell (see Fig. 2.34), biasing the detection probability maps.

Before launching the simulations we experimented with various parameter configu-

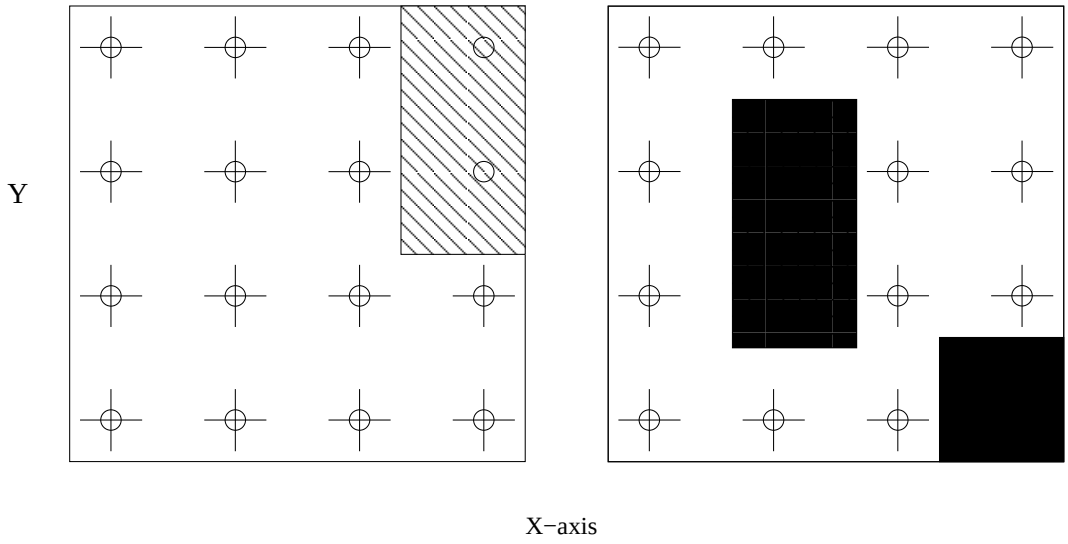

Figure 2.34: Schematic representation on how the detection probability is computed for two cells with different properties. The program sequentially inserts 16 objects of magnitude m, indicated by the open circles. The crosses indicate SExtractor's detections. The shadowed region denotes a noisy region, where SExtractor fails to detect objects for a given detection threshold; the black region denotes a masked region (pixel values = 0) that should not be taken into account in the final calculations. *Left panel:* SExtractor detects 14 objects out of 16, thus the probability of detecting an object of magnitude m in the given cell is $DP = 87.5\%$. *Right panel:* three objects fall in masked regions, and SExtractor detects 13 objects out of 13, i.e. the detection probability is 100%.

rations on small science sub–frames of $\sim 3000 \times 3000$ pixels. Based on the experience we obtained, we finally decided to use a size of 10 arcseconds (50 pixels) and introduce 36 objects per cell (which corresponds to an incremental step of 2.778%, the detection probability resolution). The range of magnitudes of the artificial objects inserted in the frames was between 16.5–18.5. The σ used in the Gaussian profiles corresponds to a FWHM $= 1.4''$ ($\sigma = \mathrm{FWHM}/2\sqrt{2\ln 2}$). Concerning the SExtractor detection thresholds, we finally used DETECT_MINAREA $= 10$ and DETECT_THRESH $= 0.5\sigma$. We also used weighting maps[9] of the form:

$$W = 1/RMS^2$$

where RMS is the root mean square mosaic–image corresponding to the science mosaic on which the simulations are being performed, produced by SExtractor. For practical reasons, each of the mosaics was split in four parts before performing the detection probability simulations. The necessary time for having a full simulation run on the 16 sub–mosaics, using a PC with a Pentium IV processor, running at 2 GHz and with 1 GByte of RAM memory was about 12 days.

[9]See Appendix A for more details.

Four detection probability maps, covering the same $\sim 17' \times 15'$ region, are shown in Fig. 2.35. The maps were produced using the same object–detection parameters, have the same resolution and are scaled in exactly the same way. The upper left, upper right, lower left and lower right panel correspond to simulations using artificial point–like sources of magnitude $K_s = 16.5$, 17.0, 17.5 and 18.0, respectively. Each pixel in the probability maps corresponds to a 50×50 pixel region in the science frames. The white color corresponds to a detection probability of 100%, the black color to 0%. The color scaling has a step of $\sim 3\%$. The regions with systematically lower detection probability in the right part of the maps are due to the bad sky subtraction of the coadded frames of tile #22 (see paragraph 2.3.5.1).

As expected, the simulations showed that, in the mosaic images, the detection probability for the overlapping regions has increased with respect to the probability in the individual coadded frames. This effect is clearly demonstrated in Fig. 2.36: the upper panel shows the probability map for a science "mini–mosaic" image which consists of four frames. The area covered by each coadded frame is marked with a red box. The identifier in the middle of each box is a combination of the tile sequential number and the chip's identification (e.g. "I05 c2" means pointing 05, chip #2). The magnitude of the inserted objects was $K_s = 18$. The four smaller panels display the probability maps corresponding to the coadded science frames used to make the mosaic image. The drop in the detection probability at the edges of the individual frames, due to the lower exposure time, is clearly visible. On the contrary, in the mosaic image, the probability remains high, allowing to detect faint objects that would be difficult to detect in the individual frames.

Fig. 2.37 shows the average detection probability, computed over all four mosaics A–D, as a function of the magnitude of the artificial objects introduced in the frames. The fact that even for the brighter objects the probability does not reach 100% is a bias due to the type of filter used for the object detection, since the SExtractor parameters were optimized for the detection of the fainter sources.

2.5 False Detections

The detection probability is the first of the two simulations that will allow to define the properties of the K_s–band catalog. Knowing the completeness limit permits to calculate the number of sources we miss, since the depth covered as a function of area depends on the observing configuration (telescope, detector), the observing strategy followed (which reflects to the minimum/maximum total integration time) and the atmospheric conditions. However, it is also important to estimate the number of false detections contaminating our sample, because of the different spatial noise properties of the mosaics.

In contrast to the detection probability simulations, the study of false detections has to be performed on object-free frames. Masking the science frames is not sufficient, because the number of objects to be identified and masked, for a given detection threshold, depends

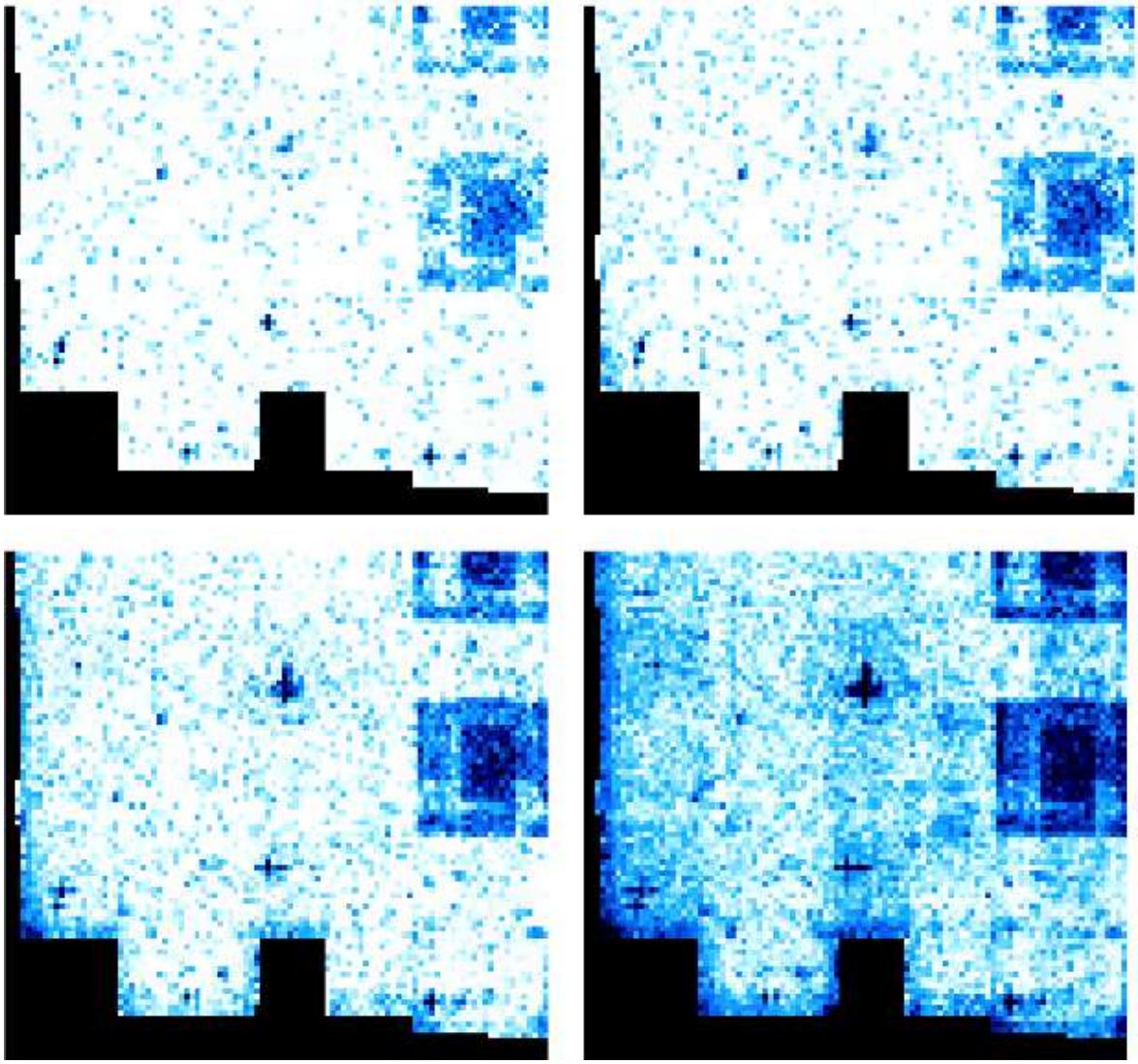

Figure 2.35: Detection probability maps covering a region of $\sim 17' \times 15'$. The maps were produced using the same object–detection parameters, have the same resolution and are scaled in exactly the same way. The upper left, upper right, lower left and lower right panels correspond to simulations using artificial point–like sources of magnitude $K_s = 16.5$, 17.0, 17.5 and 18.0, respectively. Each pixel in the probability maps corresponds to a 50×50 pixel region in the science frames. The white color corresponds to a detection probability of 100%, the black color to 0%. The step is $\sim 3\%$. The regions with systematically lower detection probability in the right part of the maps are due to the bad sky subtraction of the frames of Tile #22.

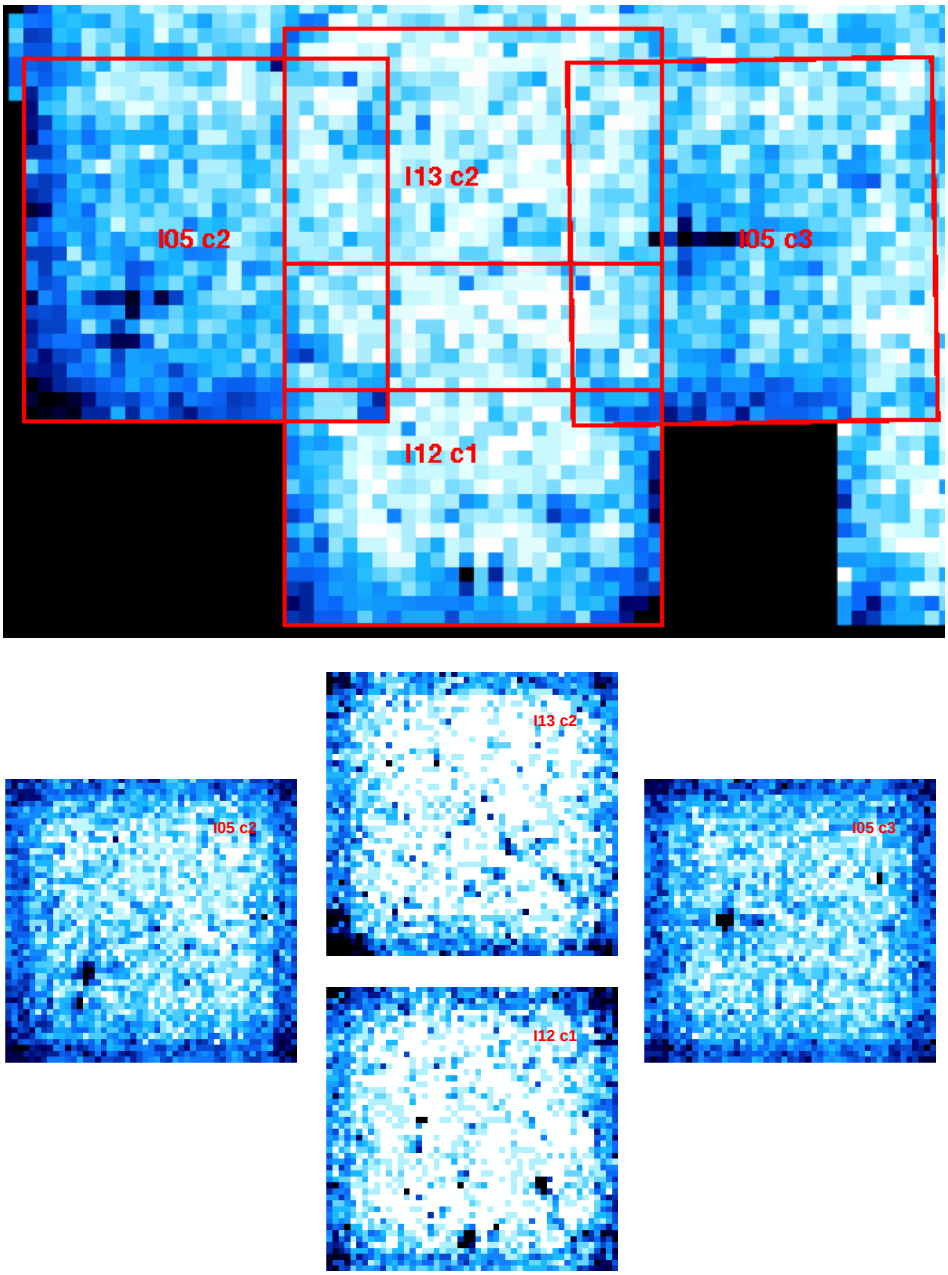

Figure 2.36: Comparison between the probability maps of a mosaic science frame, and the probability maps of the science frames used to create the mosaic. The (red) boxes in the mosaic image indicate the area covered by the individual frames. The drop in the detection probability at the edges of the individual probability maps is visible. In the mosaic image the edges form overlapping regions with a higher signal–to–noise and, therefore, higher detection probability.

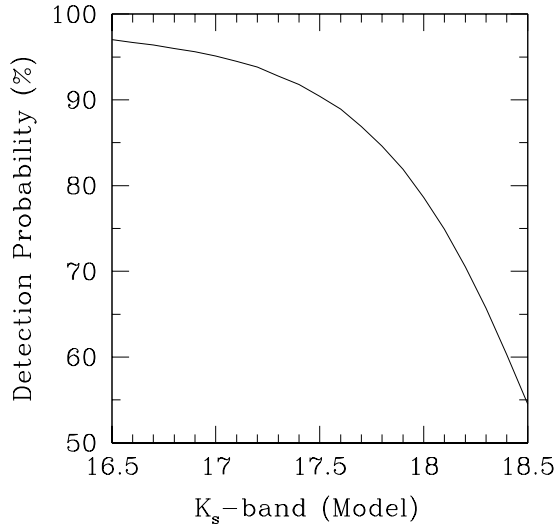

Figure 2.37: Detection probability as a function of magnitude, averaged on all simulations performed on mosaics A, B, C and D.

on the position on the mosaic, which is exactly what we want to study.

For estimating the false detection rate one has to perform the following steps:

- Compute the RMS background as a function of position on the mosaics and produce a two–dimensional (smooth) RMS map, by
  - discarding pixels associated with objects (using object masks),
  - computing the RMS value for a cell of $N \times N$ pixels (we chose $N = 5$) around every pixel,
  - replacing every pixel with the RMS computed in the previous step.
- Simulate an object–free image, with noise properties similar to those of the real image, by throwing on the RMS map random numbers from a Gaussian distribution of mean $\mu = 0$ and $\sigma = rsm_i$, where $\sigma = rsm_i$ the root mean square of the real image we want to study. The new RMS map will have the same noise level as the initial frame, but for successive pixels the noise variation will be random instead of smooth.
- Run SExtractor, using the same parameters as for the detection probability simulations and store its output in an ASCII file.
- Repeat the simulations N times ($N = 500$) and compute the average magnitude distribution of fake objects.

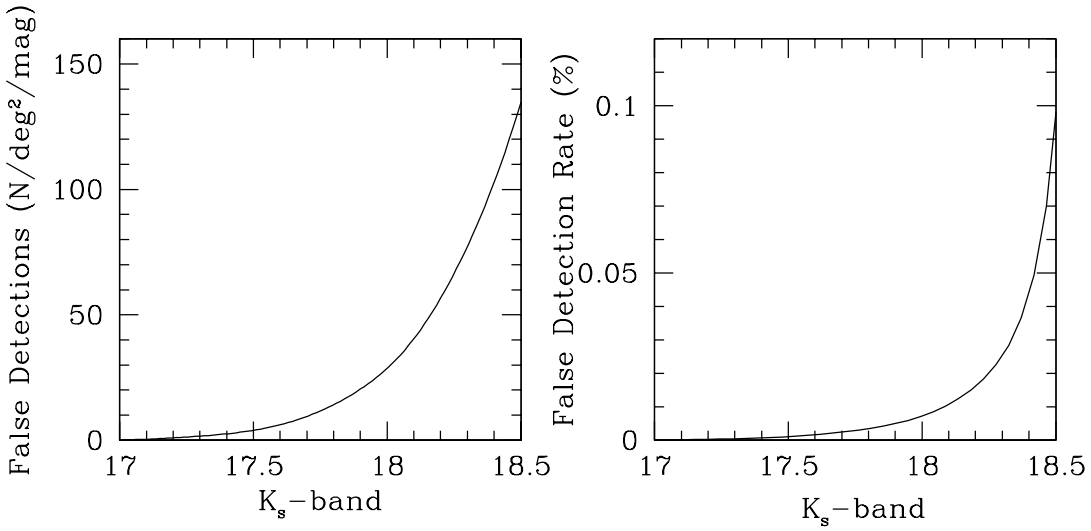

Figure 2.38: Results for the simulations on the False Detections. *Left Panel:* Number of false detections per magnitude bin, per square degree, as a function of magnitude. *Right Panel:* Average false detection rate, as a function of magnitude.

As for the detection probability, the simulations were performed in the 16 sub–mosaics A1–D4. The results from all simulations were merged and an average false detection rate was calculated. The results of the statistical analysis performed are shown in Fig. 2.38. The number of false detections per magnitude bin and per unit surface is shown in the left panel. The right panel shows the mean false detection rate as a function of magnitude. From the plots it is clear that the percentage of false detections until $K_s \sim 18.0$ still remains low.

2.6 Building the K_s–band catalog – Selection criteria

The construction of the final K_s catalog was done using SExtractor. Based on our experience from the detection probability simulations, the parameters used for the object detection were DETECT_THRESH= 0.5σ and DETECT_MINAREA= 10 pixels. The use of weight maps, as explained in section 2.4, helped in avoiding the majority of spurious detections in the noisy regions.

As explained in Appendix A, SExtractor offers different possibilities for measuring the magnitude of the detected objects. The final catalog contains the photometry measured within a fixed aperture of $1''$ and $3''$ diameter, together with the Kron–magnitudes and what is called "BEST" photometry. The (RA,Dec) and (x,y) coordinates of each detection are also included. The initial catalog contained ∼25 000 sources. However, a large fraction of the detections are false detections, mainly produced in the noisy regions or regions

with lower signal–to–noise, usually located at the borders of the mosaics. Using the (x,y) coordinates and the Kron magnitude, a detection probability was assigned to every object, based on the probability maps produced by the simulations. Once the catalog has been updated with the new information, entries with a detection probability lower than 70%, or Kron magnitude fainter than 18, were rejected.

Using the above criteria we generated a K_s–band catalog with ∼4500 reliable entries, ready to be correlated with the R and z'–band CTIO catalogs. The catalog matching and the quasar identification, through the use of the K–excess technique, is the scope of the following chapter.

Chapter 3

Generating the KX–selected quasar sample

The K–excess technique will be implemented in the $Rz'K_s$ color space, where the identification of quasars will be based on a combination of morphological and color criteria. Existing information on spectroscopically identified quasars and galaxies, will help in setting better constraints on the selection criteria. This chapter describes the properties of the R and z'–band photometric catalogs, together with the steps followed in order to generate the sample of the KX–selected quasar candidates.

3.1 The R, z'–band CTIO data

The basic characteristics of the R and z'–band data have been described in section 2.3.13. Fig. 3.1 shows a region of $\sim 4' \times 3.4'$ on the sky, as it has been observed in the R–band (upper left panel), z'–band (upper right panel) and K_s–band in the Las Campanas data (lower left panel). A K_s–band 2MASS (Skrutskie et al. 2006) image is also displayed in the lower right panel. North is up, East is to the left. A 30–arcsecond ruler is shown at the lower right side of each panel. The displayed region has been chosen on purpose, since it indicates all problems related to the matching of the R, z' and K_s–band catalogs:

- LCO detections with $K_s < 13.5$ will correspond to saturated objects in the R, z'–band data. Regions around saturated objects have been masked appropriately, to avoid matching K_s detections with R, z' detections having unreliable photometry.
- Because of saturation effects, the spikes extending up to one arcminute in the East–West direction will affect the photometry of a small percentage of R, z' detections. Regions with spikes were masked in order to have the affected detections excluded from the corresponding catalogs.

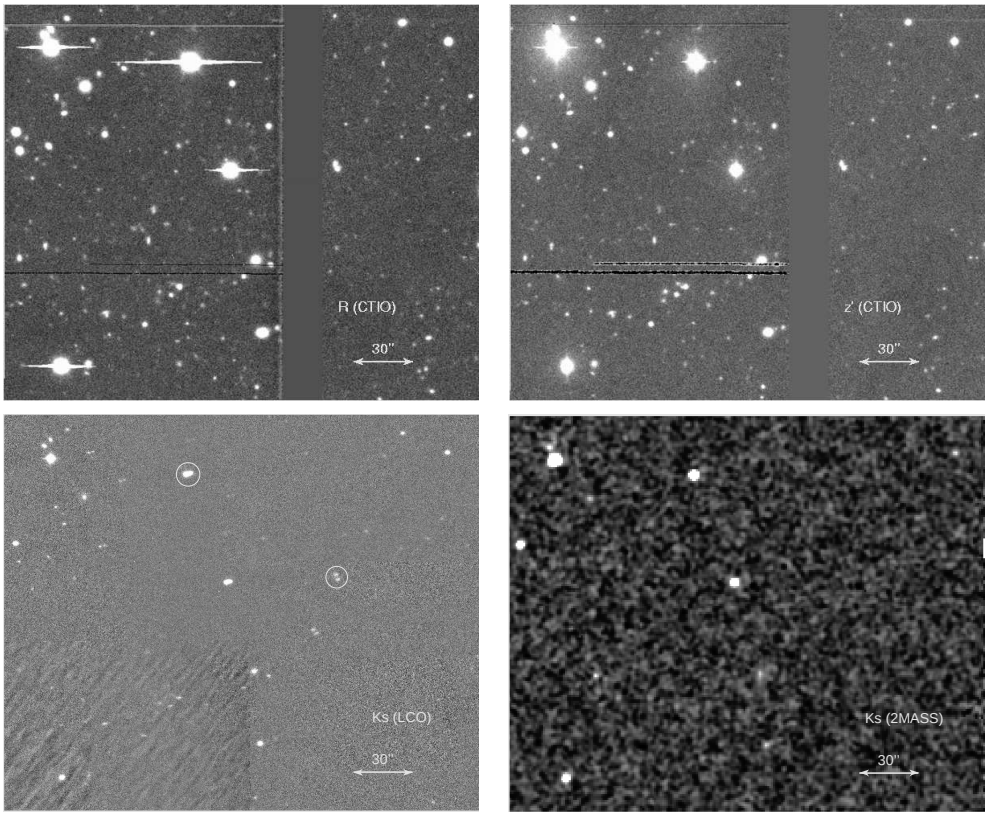

Figure 3.1: R (upper left panel), z' (upper right), Las Campanas K_s (lower left) and 2MASS K_s–band (lower right) data. The field of view is $\sim 4' \times 3.4'$. At the lower right part of the panels, a 30–arcsecond ruler is shown. North is up, East is to the left.

- Since the object density is much lower in the LCO data, the blending because of sources being very close to each other is much stronger in the R, z'–bands. As a consequence, K_s detections matched with R, z' blended objects will have unreliable SExtractor photometry. These objects have been removed, for minimizing the presence of objects with weird colors in the color–color space.

- The CTIO MOSAIC II imager consists of eight CCDs. Because of the configuration of the camera, there are gaps between the CCD arrays, both in the East–West and North–South direction, of about $\sim 20''$. As a consequence, there will be K_s detections with no counter–part in the R, z'–band data.

- Because of the gaps in the CTIO images, some LCO detections will be matched to CTIO detections falling on the edges between the CCD arrays and the gaps. The data were visually inspected and CTIO detections falling on the edges were removed from the corresponding catalogs.

- The most eastward of the two circles in the LCO image of Fig. 3.1 indicates an object which appears to be extended, but with a slightly "peculiar" morphology. The R

and z'–band images are not of much help regarding the morphology of the object, since they are saturated. The answer comes from the 2MASS image, despite its lower resolution ($\sim 1''$ pixel^{-1}): the extended morphology of the marked object is an artifact: the object falls in an overlapping region[1] and the astrometric uncertainties of the individual coadded frames are such, that their superposition has produced an elongated object in the mosaic image. The photometry obtained with SExtractor will not be reliable and the regions populated by the different classes of objects in the color–color space will have a relative scatter.

- The second circle, southwestwards of the first one, marks two objects at a very small angular separation. Based on the previous experience, it is necessary to inquire the possibility of another artifact. The objects are too faint to be detected in the 2MASS image, but in this case the R, z' images are of great help: there are indeed two objects, thus the artifact scenario is rejected. All K_s–band detections having a second object at an angular separation less than $2''$ were visually inspected in order to remove only the artifacts and not real detections at a close angular separation.

- As explained in section 2.6, the final K_s catalog consisted of objects having a Kron magnitude brighter than 18 and a detection probability $DP \geq 70\%$. Since these selection criteria are average values, the rejection of the false detections will be less effective in the noisier regions. Such an area is shown in the southeastern part of the Las Campanas image in Fig. 3.1 (lower left panel). Because of the presence of noisier regions in the LCO mosaics, the K_s–band catalog will contain a number of false detections that will not be matched with any object in the R, z' data.

The matching of the K_s–band catalog with the Rz' detections, which were pre–correlated by S. Andreon, was performed by means of IRAF's task xyxymatch, using the "tolerance" algorithm. The method matches the entries contained in two lists, the "input" and "reference", and for each object in the reference file finds the objects in the input list within a given tolerance. From our experience on the astrometric calibration of the Las Campanas fields, we estimated that a $2''$ tolerance was sufficient in order to have the majority of the real K_s–band detections matched to a Rz' detection, keeping the percentage of wrongly matched detections low. The final $Rz'K_s$ matched catalog contained $\sim$ 3400 detections. The missing $\sim$1100 entries, between the K_s–band and the matched $Rz'K_s$ catalog, are detections corresponding to one of the cases described above.

3.1.1 Bringing the $Rz'K_s$ data to the same photometric system

Before proceeding with the analysis of the photometric catalogs it is necessary to bring the $Rz'K_s$ magnitudes to the same photometric system. As mentioned in section 2.3.12, the

[1]The specific region in the mosaic image originates from the superposition of seven frames.

photometric calibration of the K_s–band data was performed through the use of photometric standard stars taken from Persson et al. (1998) and 2MASS. The R–band data were calibrated in the Kron–Cousins photometric system using standard stars from the Landolt catalog (Landolt 1992). These two catalogs have been established using as reference A0–type stars, mainly α Lyrae (Vega). Regarding the z' observations, however, things are slightly different. The z' filter is not part of the $UBVRIJHK$ Johnson–Kron–Cousins filter sequence, that has been calibrated on the basis of Vega type standard stars. Instead it shares the same properties with the filters of the Sloan Digital Sky Survey (SDSS). Therefore, the calibration of the z'–band observations has been based on a different set of photometric standard stars, which is closer to the so–called AB photometric system. Detailed information on the calibration of the SDSS photometric system and the transformation equations for converting SDSS to Johnson–Kron–Cousins magnitudes can be found in Smith et al. (2002).

The relation between the Vega and AB photometric system can be described as follows: the apparent magnitude m of any object in the Vega photometric system is given by the equation:

$$m_{\rm vega}({\rm obj}) = -2.5\,\log\frac{\int F_\nu({\rm obj})\,S_\nu\,d\nu}{\int F_\nu({\rm vega})\,S_\nu\,d\nu} \tag{3.1}$$

where F_ν(obj) is the object's monochromatic flux in $\mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}}$, S_ν is the instrumental (filter, detector, telescope) response and F_ν(vega) is the Vega flux, in $\mathrm{erg\,s^{-1}\,cm^{-2}\,Hz^{-1}}$. By definition, Vega's magnitude is 0 in all filters. The magnitude of any object in the AB system is given by the equation

$$m_{AB}({\rm obj}) = -2.5\,\log\frac{\int F_\nu({\rm obj})\,S_\nu\,d\nu}{\int S_\nu\,d\nu} - 48.60 \tag{3.2}$$

so that an object with constant F_ν (flat energy distribution) has the same magnitude in all bands and colors equal to zero. By multiplying the numerator and denominator of Eq. (3.2) by

$$\int F_\nu({\rm vega})\,S_\nu\,d\nu$$

we obtain:

$$\begin{aligned} m_{AB}({\rm obj}) &= -2.5\,\log\frac{\int F_\nu({\rm obj})\,S_\nu\,d\nu \times \int F_\nu({\rm vega})\,S_\nu\,d\nu}{\int S_\nu\,d\nu \times \int F_\nu({\rm vega})\,S_\nu\,d\nu} - 48.60 \\ &= -2.5\,\log\frac{\int F_\nu({\rm obj})\,S_\nu\,d\nu}{\int F_\nu({\rm vega})\,S_\nu\,d\nu} - 2.5\,log\frac{\int F_\nu({\rm vega})\,S_\nu\,d\nu}{\int S_\nu\,d\nu\times} - 48.60 \\ &= m_{\rm vega}({\rm obj}) + mag_{AB}({\rm vega}) \end{aligned} \tag{3.3}$$

and the conversion between the two systems is simply given by the AB magnitude of Vega:

$$\text{conv} = mag_{AB}(\text{vega}) = -2.5 \log \frac{\int F_\nu(\text{vega})\, S_\nu \, d\nu}{\int S_\nu \, d\nu} - 48.60 \tag{3.4}$$

The AB z'–band magnitudes of the $Rz'K_s$ catalog were converted to Vega magnitudes system according to the formula:

$$z'_{vega} = z'_{AB} - 0.549$$

3.1.2 Choosing the SExtractor photometric aperture

The R, z' and K_s catalogs, constructed with SExtractor, contain different types of photometry (fixed aperture magnitudes, Isocorr, etc)[2]. The final choice on the type of magnitudes to use depends on the type of research we want to perform. In order to implement the K–excess method, the colors have to be measured through a uniform aperture. Although quasars are point–like sources, their morphology might vary as a function of wavelength. Using aperture photometry magnitudes is equivalent to measuring the light emitted from the same region of the quasar, in physical units. On the contrary, magnitudes based on more complex algorithms (e.g. Kron, Isocorr, etc) are difficult to control and can not be interpreted in terms of physical units.

The implementation of the K–excess method will be performed on the basis of the $(R - z')$ versus $(z' - K_s)$ color–color diagram. Before computing the color indices it is necessary to examine any possible effects related to the fact that the observations were performed under different atmospheric conditions, since the amount of light within a fixed aperture decreases as the seeing conditions degrade. The aperture correction, i.e. the amount (in magnitude units) corresponding to light losses due to the degradation of the seeing (or deformations of the PSF, in general) can be calculated analytically. Since we are mainly interested in point–like sources, we will calculate the aperture correction under the assumption that the spatial distribution of the light falling on an array can be described by a two–dimensional Gaussian of the form:

$$I(x, y) = H \exp\left(-\left\{\frac{(x - x_c)^2}{2\sigma^2} + \frac{(y - y_c)^2}{2\sigma^2}\right\}\right) \tag{3.5}$$

where (x_c,y_c) corresponds to the position of the photocenter of the point–like source, H is a scaling factor relating the peak intensity of the Gaussian to the magnitude of the point–like source, and the standard deviation σ is related to the FWHM according to the relation σ =FWHM/$2\sqrt{2 \ln 2}$. By integrating Eq. (3.5) in polar coordinates we finally derive that the total flux within a two–dimensional Gaussian is:

$$\text{Flux} = H \, \pi \, 2 \, \sigma^2 \tag{3.6}$$

[2]A more detailed description on the different possibilities is provided in Appendix A.

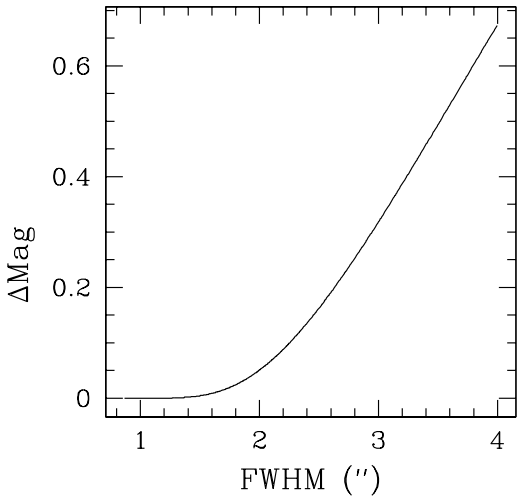

Figure 3.2: Aperture correction as a function of FWHM, considering a Gaussian profile for describing the light distribution emitted by a point–like source.

The ratio of the flux measured within a 3″ aperture, Flux_3, over the total flux of the Gaussian, will correspond to a magnitude difference:

$$\Delta\mathrm{Mag} = -2.5 \log \left(\frac{\mathrm{Flux}_3}{\mathrm{Flux}} \right) \tag{3.7}$$

Fig. 3.2 shows the aperture correction as a function of FWHM. For seeing conditions up to $\sim 1.5''$ the flux within the 3″ aperture is practically equal to the total flux contained in the Gaussian, thus no aperture correction is necessary. As the seeing degrades, the amount of flux missed by the fixed aperture starts becoming important and an aperture correction has to be applied in order to convert aperture photometry to total photometry.

When computing colors using a fixed aperture it is the *difference* in FWHM between observations performed at different energy bands that will indicate whether it is necessary or not to apply an aperture correction, since for the same seeing conditions the amount of flux missed in both bands will be the same.

Table 3.1 presents the typical seeing conditions corresponding to the four CTIO pointings covering the Las Campanas field. The first and second column provide the pointing identification and the central coordinates of each field, in degrees, respectively. The third and fourth column provide the FHWM (in arcseconds) for the R–band and z'–band observations, respectively. The maximum seeing difference between observations of the same field in the two bands is $0.2''$, meaning that no aperture correction is necessary for the $(R - z')$ color index.

To examine the possibility of implementing an aperture correction to the z' and K_s–band magnitudes, before computing the $(z' - K_s)$ color, we first have to study the

Field	(RA_c,Dec_c)	$R(''$)	$z'(''$)
B4	(36.45427,−4.32669)	1.20	1.00
B5	(36.45221,−4.92249)	1.35	1.40
C4	(35.86548,−4.32579)	1.35	1.20
C5	(35.86450,−4.91304)	1.35	1.48

Table 3.1: Mean seeing conditions during the CTIO observations, for the pointings covering the Las Campanas field. The second column gives the central coordinates of each pointing, in degrees. The third and fourth columns provide the average FWHM, in arcseconds, for the R and z'–band observations, respectively.

seeing variations in the Las Campanas data. The latter is not a trivial task, given that 200 images, obtained during very different seeing conditions, were put together for producing the K_s–band mosaics. To measure the seeing variations as a function of position in the LCO data we selected relatively bright sources ($K_s \leq 17$) with a SExtractor stellarity index higher than 0.95. The magnitude limit ensures that the sources have enough signal to trust the stellarity index provided by SExtractor, while the high stellarity cut–off ensures that only real point–like sources will be used.

The median FWHM measurements, for the K_s–band observations. are shown in Fig. 3.3. The grid size was based on the number of sources available for computing the median FWHM per cell. For the 36–cell grid, there are 6–25 sources per grid element. The three empty cells at Right Ascension ∼36.4–36.5 correspond to regions with no K_s–band sources complying with the selection criteria. However, because of (a) the observing strategy followed and (b) the information from the adjacent pointings, we can deduce that the differential seeing variations in those regions will be of the same order and therefore no aperture correction has to be applied for those fields either.

A histogram of the *differential* FWHM (in arcseconds), between the K_s–band and z'–band observations, as computed from the values shown in Table 3.1 and Fig. 3.3, is presented in Fig. 3.4. The binsize used is $0.1''$.

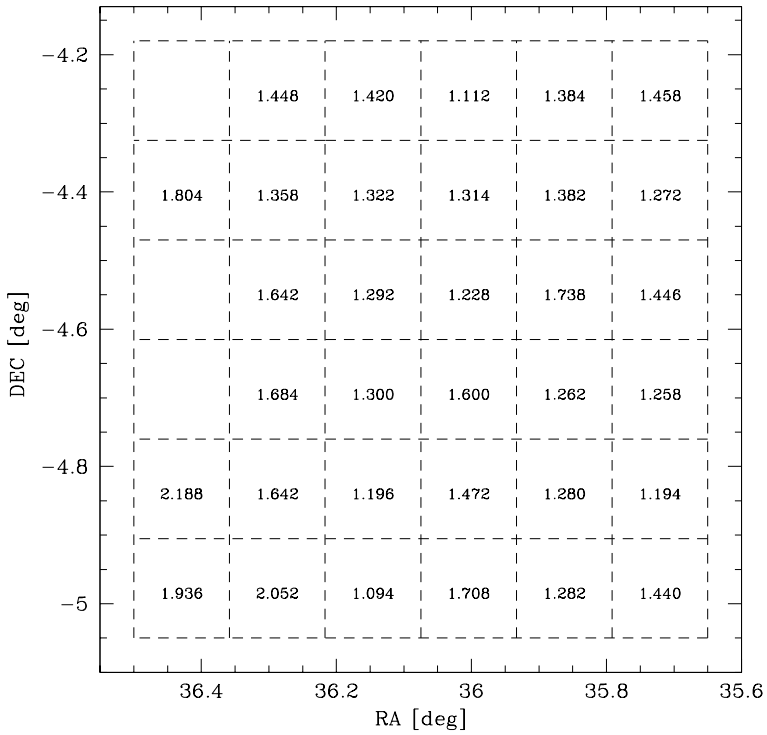

Figure 3.3: FWHM measurements in the Las Campanas field, for the K_s–band observations.

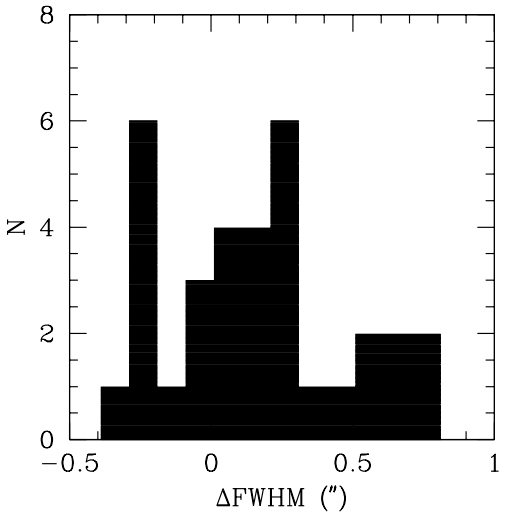

Figure 3.4: Histogram of the differential seeing conditions between the K_s and the z'–band observations (in arcseconds). The bin size is $0.2''$.

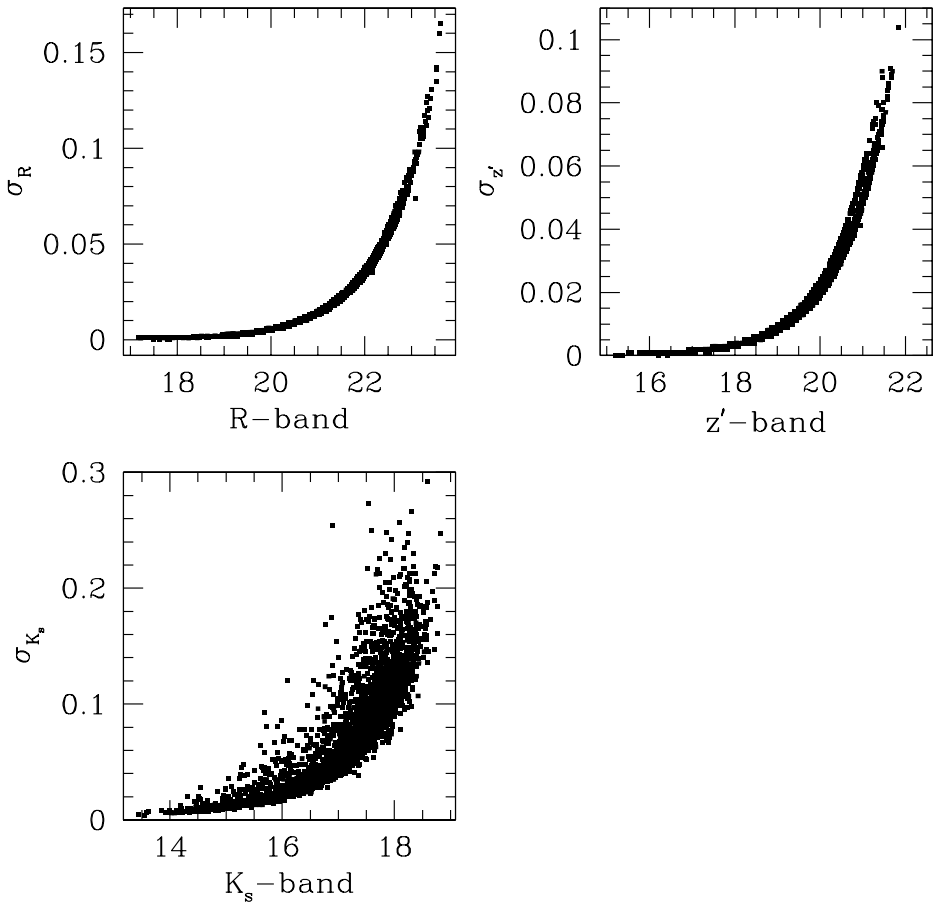

Figure 3.5: Standard deviation (σ) as a function of magnitude, for the R (upper left panel), z' (upper right panel) and K_s-band (lower panel) detections in the matched $Rz'K_s$ catalog.

Fig. 3.5 shows the SExtractor computed photometric error as a function of magnitude ($3''$ aperture photometry), for the R (upper left panel), z' (upper right panel) and K_s-band (lower panel) matched detections. Fig. 3.6 shows the number counts (in percentage) for the magnitudes (left column) and standard deviations (right column) for the R (upper raw), z' (middle raw) and K_s–band (bottom raw) for the $\sim$ 3400 matched sources. The bin size for the magnitudes and the standard deviation is 0.5 and 0.01, respectively. As expected, the R and z'–band detections share the same properties, since their data were obtained with the same telescope/detector and under very similar observing conditions. The large scatter in the standard deviation for the K_s–band photometry is mainly due to (a) the scatter of the errors related to the sky subtraction and (b) the instability of some of the 16 independent elements constituting the WIRC camera, which, together with the complex methodology used for covering the LCO area, resulted into an inhomogeneous mosaic with spatially varying properties.

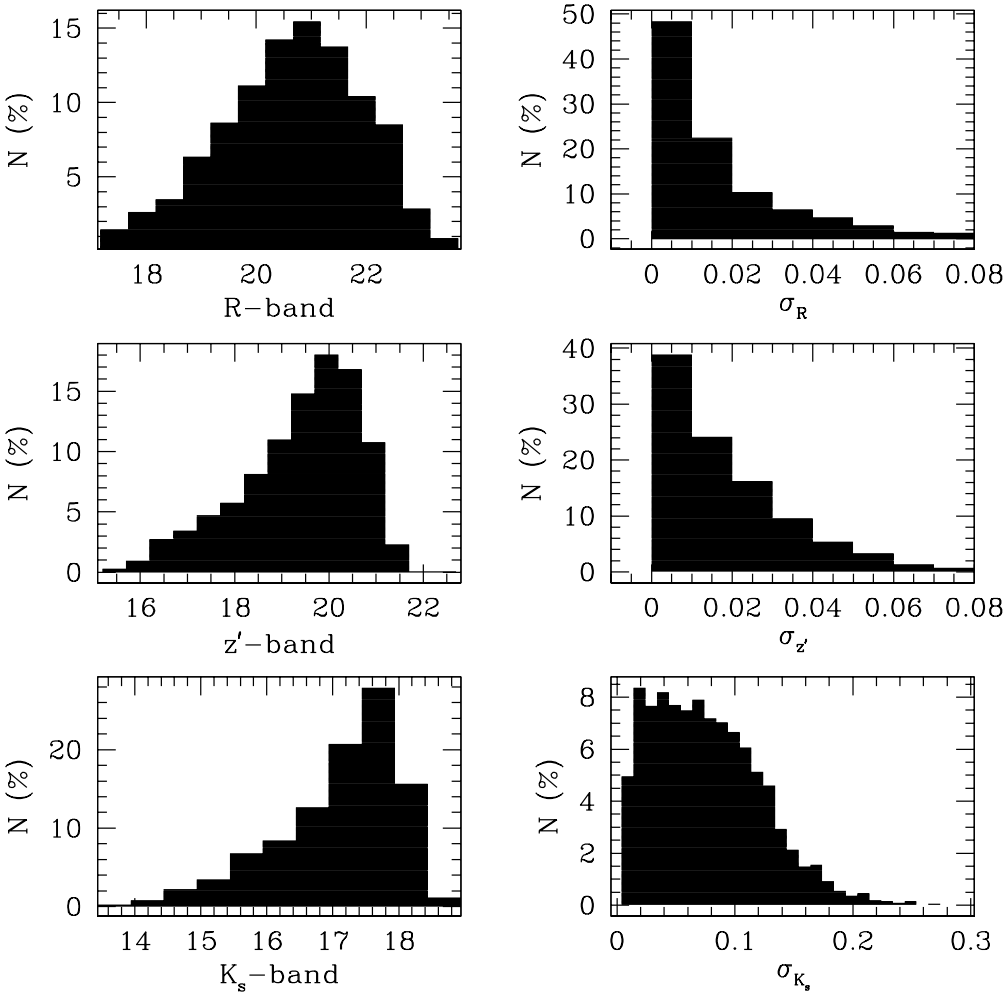

Figure 3.6: Number counts (in percentage) for the magnitudes (left column) and standard deviations (right column) for the R (upper raw), z' (middle raw) and K_s–band (bottom raw) for the ~ 3400 matched sources. The bin size for the magnitudes and the standard deviation is 0.5 and 0.01, respectively.

3.1.3 Separating the point–like from the extended sources

The color–color diagram, in which we want to implement the K–excess technique, will contain stars, galaxies and quasars. The galaxies have an extended morphology, while the stars and quasars look point–like. The first step for the identification of the QSOs among the thousands of sources in the $Rz'K_s$ plane is to use SExtractor's morphological classification, expressed by the so–called stellarity index. This index provides the probability of a source being point–like, given the quality and characteristics of the *specific* data. In other words, the stellarity index is not an absolute quantity: a point–like source might

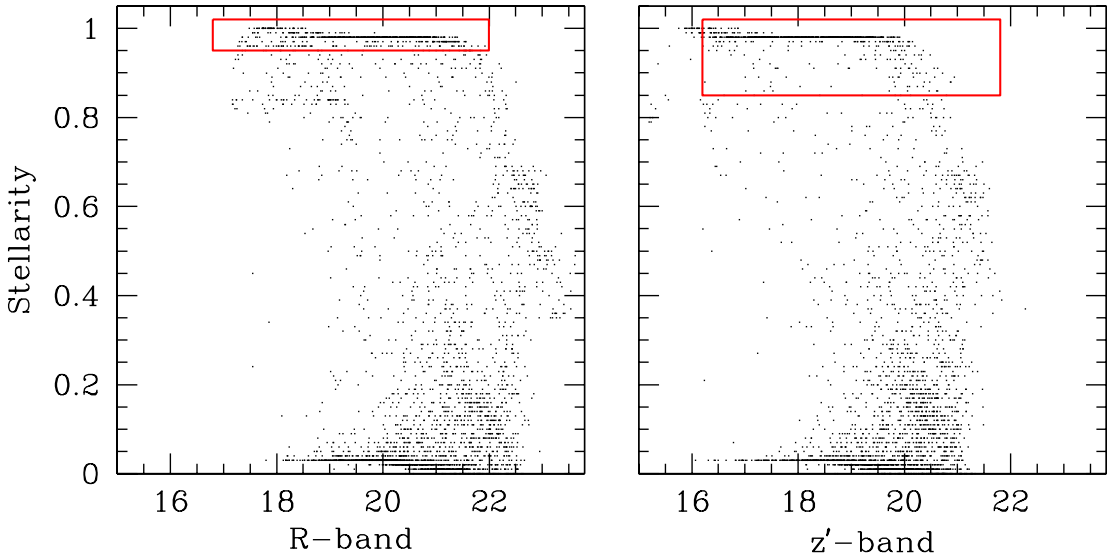

Figure 3.7: Stellarity index as a function of R (left panel) and z'–band (right panel) magnitudes. *Left panel: The rectangle indicates the region occupied by the sources classified as point–like (stellarity index higher than 0.95 and magnitude range* $16.8 \leq R \leq 22.0$*). Right panel:* The rectangle indicates the region occupied by the sources classified as point–like (stellarity index higher than 0.85 and magnitude range $16.2 \leq z' \leq 21.8$).

have a different stellarity index in different energy bands either because its morphology is different, or because the data do not share the same characteristics (e.g. different Point Spread Functions, seeing conditions, signal–to–noise ratios).

Due to the higher sensitivity and uniformity (i.e. few pointings under similar observing conditions) of the CTIO data, we preferred to perform the source characterization using a combination of R and z'–band stellarity indices and magnitudes. Such a combination allows to properly identify the sources, avoiding caveats because of the very bright/faint magnitudes, or a possibly different morphology in each band: at very bright magnitudes a saturated star will have a low stellarity index, while at faint magnitudes the morphological classification is rather unreliable.

Fig. 3.7 reveals how the stellarity index is related to the R (left panel) and z'–band (right panel) magnitudes. Given the quality of the data, the set the stellarity cutoff limit to 0.95 and 0.85 for the R and z'–band detections, respectively. The rapid drop of the stellarity index for magnitudes fainter than ~ 21.5 clearly indicates that after this limit we are unable to properly classify the sources as point–like or extended. The actual value chosen was 22 and 21.8 for the R and z'–band, respectively. For magnitudes brighter than $R = 16.8$ and $z' = 16.2$ the stellarity cutoff limit for saturated sources was set at 0.5. In each panel, the region occupied by the point–like sources s is marked with a rectangle.

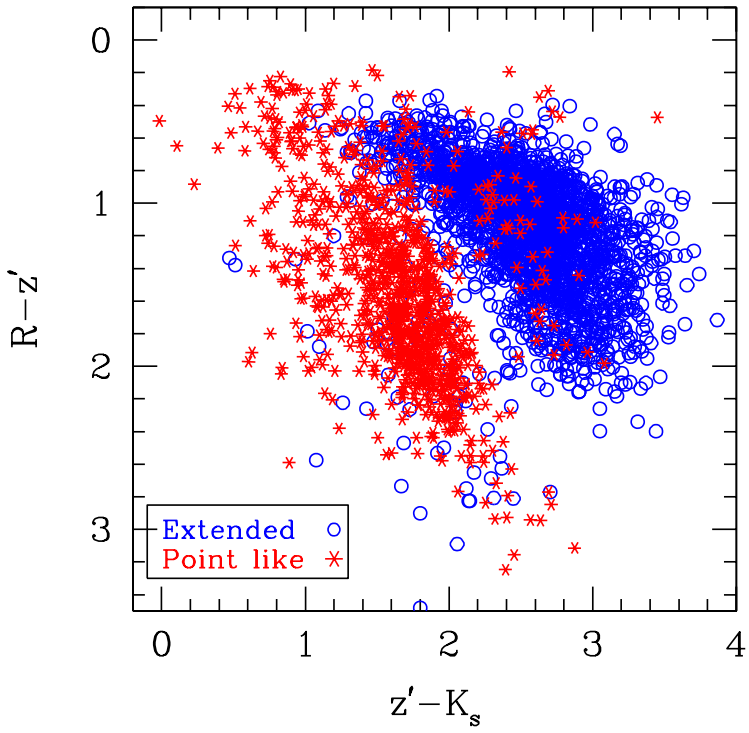

Figure 3.8: $(R - z')$ vs $(z' - K_s)$ diagram for the point–like (red asterisks) and extended sources (blue open circles). The computed colors are based on the 3″ photometry (Vega magnitudes). The point–like sources (mainly stars) form a locus well separated from the region occupied by the extended sources (galaxies). Small contamination effects are either due to false photometry, photometric uncertainties, or because of the presence of objects with colors different from the typical colors of the majority of the objects occupying each locus.

Any source in the $Rz'K_s$ catalog with:

$$\begin{array}{lllll} \mathrm{class}_R & > 0.95 & \wedge & 16.8 < R < 22.0 & \wedge \\ \mathrm{class}_z & > 0.85 & \wedge & 16.2 < z' < 21.8 & \end{array}$$

where the symbol $\wedge$ is the logical AND operator. was classified as a point–source, otherwise as a source with extended morphology.

3.1.4 The color–color diagram

Fig. 3.8 shows the $(R - z')$ vs $(z' - K_s)$ diagram for the $Rz'K_s$ matched sources. The computed colors are based on the 3″ aperture photometry. Morphologically extended (open circles) and point–like (asterisks) objects largely occupy two distinct regions of the

$Rz'K_s$ color space. There is some color overlap between the two samples and it is the purpose of the next sections to investigate their nature.

The characterization of the sources is done by analyzing the general properties of the whole sample, since it is not possible to work on every source individually. Therefore, as expected, a small percentage of detections will deviate from the global characteristics of the population. These detections will either (a) belong to problematic cases (e.g. wrong photometry, photometric uncertainties, wrong source matching), (b) were wrongly classified because they were not respecting at least one of the selection criteria (position in the color–color diagram, morphological classification), or (c) they indeed belong to a different category of objects. For every new plot all the outliers are visually inspected on the R, z', and K_s–band images in order to understand the nature of the deviating points. Problematic cases (i.e. cases with unreliable photometry) are removed from the catalogs and the plots are re–generated.

3.2 The 2dF catalog

The Two Degree Field (2dF) Survey is a project designed to spectroscopically cover a large area on the sky. It uses a multi–object fiber spectrograph, mounted on the 3.9 m Anglo–Australian Telescope, which allows the acquisition of up to 400 simultaneous spectra. The spectral resolution is between 500 and 2000, over the wavelength range from 4400Å up to 11 000Å. Extensive information regarding the technical specifications of the 2dF facility can be found in Lewis (2002).

As part of the XMDS, the 2dF spectrograph targeted some 800 sources, in November 2003. The observations were performed by M. Watson and M. Page. The spectroscopic survey was X–ray flux limited, with $F_{0.5-4.5} > 10^{-14}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. The expected optical magnitude limit for the survey's sensitivity is $r' \approx 22$, for point–like sources. The positioning of the fibers was based on the (RA,Dec) coordinates of the optical counterparts of the X–ray detections. The correlation of the X–ray and optical (CFHTLS) catalogs was performed using a radius of $3''$.

The data were reduced, using standard IRAF procedures, by astronomers at the Institute of Astrophysics of the University of Liège, namely C. Jean, O. Garcet and K. Libbrecht. 201 spectra, corresponding to 146 different targets (some objects were observed twice, due to overlapping between different 2dF pointings) were located in the Las Campanas area. For a large fraction of the targets, however, the optical counterpart of the X–ray detections exceeded the magnitude limit of the 2dF spectroscopic survey ($R \approx 21$). As a result, 136 of the obtained spectra were noisy (≈68%), making it impossible to extract any scientific information out of them. The remaining spectra targeted 57 different objects, out of which only 22 were finally matched with a single entry in the $Rz'K_s$ catalog. These spectra corresponded to one star, nine quasars and 12 galaxies. The positions of the 34 useful spectra, not matched to any object in the $Rz'K_s$ catalog,

were visually inspected in the R, z' and K_s images, in order to inquire on this. 29 of these objects were either:

- Falling in a region with no observations in at least one band.
- Were not included in the $Rz'K_s$ catalog because of unreliable photometry, due to blending effects in the R and z'–bands.
- Were not included in the $Rz'K_s$ catalog because of the cut–off limits (i.e. magnitude, detection probability) in the K_s–band catalog.

However, a query on the individual sub–catalogs (R, z' and K_s) allowed to recover the position and photometry of five objects, classified as QSOs. These objects were not included in the matched catalog because their detection probability in the K_s catalog was $\sim 67\%$, just below the threshold of 70%.

Table 3.2 presents the astrometry and photometry of the 27 spectroscopically identified sources. In order to keep unbiased the properties of the matched $Rz'K_s$ catalog, the five quasars previously mentioned will not be used in any statistical analysis involving the photometry and the general properties of the $Rz'K_s$ detections. Nevertheless, since they will be of great help at later stages of this work, we decided to also include them in Table 3.2.

Fig. 3.9 shows the distribution in the color–color space of the 22 sources spectroscopically identified as stars (asterisks), QSOs (squares) and galaxies (triangles), found in the $Rz'K_s$ catalog. The extended source area (see Fig. 3.8) is indeed occupied by the spectroscopically confirmed galaxies. As far as the quasars with 2dF spectra are concerned, all but one occupy the bluer part ($R - z' < 0.8$) of the "extended–source" region. A visual inspection of the R, z' and K_s–band data didn't reveal any information that could possibly explain the colors of the isolated quasar found at $(z' - K_s) \approx 1$.

ID	RA (J2000) h m s	Dec (J2000) ° ′ ″	Type	Rdsft	R	σ_R	z'	$\sigma_{z'}$	K_s	σ_{K_s}	$R-z'$	$z'-K_s$
528	2 22 55.08	−4 12 10.58	S	–	19.012	0.002	17.359	0.001	14.949	0.021	1.654	2.410
735	2 22 44.40	−4 33 47.02	Q	0.760	18.307	0.001	18.510	0.003	16.309	0.034	−0.203	2.201
736	2 22 47.90	−4 33 30.20	Q	1.629	20.151	0.006	20.122	0.013	17.900	0.128	0.029	2.222
97	2 23 26.47	−4 57 06.30	Q	0.826	20.994	0.013	20.968	0.029	17.843	0.135	0.026	3.125
100	2 23 54.82	−4 48 15.19	Q	2.463	18.163	0.001	18.371	0.003	16.092	0.023	−0.208	2.279
251	2 24 29.14	−4 58 08.11	Q	1.497	19.626	0.004	19.959	0.011	17.907	0.101	−0.333	2.052
23	2 25 14.40	−4 47 00.38	Q	1.924	19.225	0.003	18.456	0.004	16.918	0.045	0.769	1.538
41	2 25 40.61	−4 38 25.30	Q	2.483	19.860	0.005	19.898	0.013	17.625	0.096	−0.038	2.273
11	2 25 56.83	−4 58 53.29	Q	1.183	20.947	0.014	21.147	0.039	17.968	0.127	−0.200	3.179
14	2 25 57.62	−4 50 05.50	Q	2.263	19.454	0.004	19.722	0.011	17.826	0.084	−0.268	1.896
531	2 22 57.98	−4 18 40.50	G	0.237	19.047	0.003	18.112	0.002	14.674	0.010	0.935	3.438
544	2 23 02.04	−4 32 04.81	G	0.616	20.437	0.008	20.004	0.012	16.767	0.040	0.433	3.237
742	2 23 15.36	−4 25 58.69	G	0.190	19.695	0.004	19.423	0.007	17.099	0.064	0.272	2.324
743	2 23 19.66	−4 47 30.80	G	0.293	19.284	0.003	18.798	0.004	15.651	0.045	0.486	3.147
96	2 23 44.26	−4 57 25.42	G	0.157	19.803	0.005	19.576	0.008	16.635	0.034	0.227	2.941
65	2 23 51.29	−4 20 53.41	G	0.181	19.635	0.004	19.666	0.009	17.144	0.056	−0.031	2.522
125	2 24 18.79	−5 01 20.89	G	0.458	20.769	0.011	20.589	0.020	18.177	0.141	0.180	2.412
248	2 25 21.07	−4 39 49.90	G	0.265	19.416	0.003	18.941	0.005	16.074	0.048	0.475	2.867
243	2 25 24.70	−4 40 44.69	G	0.263	19.044	0.003	18.448	0.004	15.412	0.017	0.596	3.036
46	2 25 31.39	−4 42 20.30	G	0.209	20.015	0.006	19.929	0.013	17.806	0.074	0.086	2.123
42	2 25 32.02	−4 43 46.20	G	0.314	20.224	0.007	19.731	0.011	16.929	0.048	0.493	2.802
357	2 25 58.87	−5 00 54.50	G	0.148	18.324	0.001	17.921	0.002	15.106	0.029	0.403	2.815

Table 3.2: Sources in the Las Campanas field, with spectroscopic identification. Column 1 gives the identifier of the source in our spectroscopic data–base. Columns 2 and 3 give the Right Ascension and Declination of each source, for Equinox 2000. Columns 4, 5 provide the spectroscopic classification of the source (Star, Quasar or Galaxy) and its measured redshift (for the latter two cases), respectively. The error on the redshift estimation is smaller than 0.001 for all sources. Columns 6–7, 8–9 and 10–11 give the 3–arcsecond photometry and its error for the the R, z' and K_s–bands, respectively. Columns 11–12 provide the $(R-z')$ and $(z'-K_s)$ color index for each source, respectively. The five quasars not included in the matched $Rz'K_s$ catalog are presented at the bottom of Table 3.2 and are marked as "Q*".

ID	RA (J2000) h m s	Dec (J2000) ° ′ ″	Type	Rdsft	R	σ_R	z'	$\sigma_{z'}$	K_s	σ_{K_s}	$R-z'$	$z'-K_s$
530	2 22 49.63	−4 13 52.97	Q*	1.566	20.301	0.007	20.461	0.018	17.793	0.087	−0.160	2.668
281	2 23 51.10	−4 47 29.76	Q*	2.164	20.065	0.005	20.311	0.015	18.001	0.101	−0.246	2.310
91	2 23 58.66	−4 53 51.40	Q*	2.275	19.866	0.004	20.084	0.012	17.509	0.066	−0.218	2.575
271	2 24 13.46	−4 52 10.27	Q*	2.487	20.646	0.009	20.913	0.026	18.097	0.110	−0.267	2.816
375	2 25 37.03	−5 01 09.41	Q*	1.937	19.968	0.005	19.953	0.013	18.076	0.106	0.015	1.877

Table 3.2: Continued

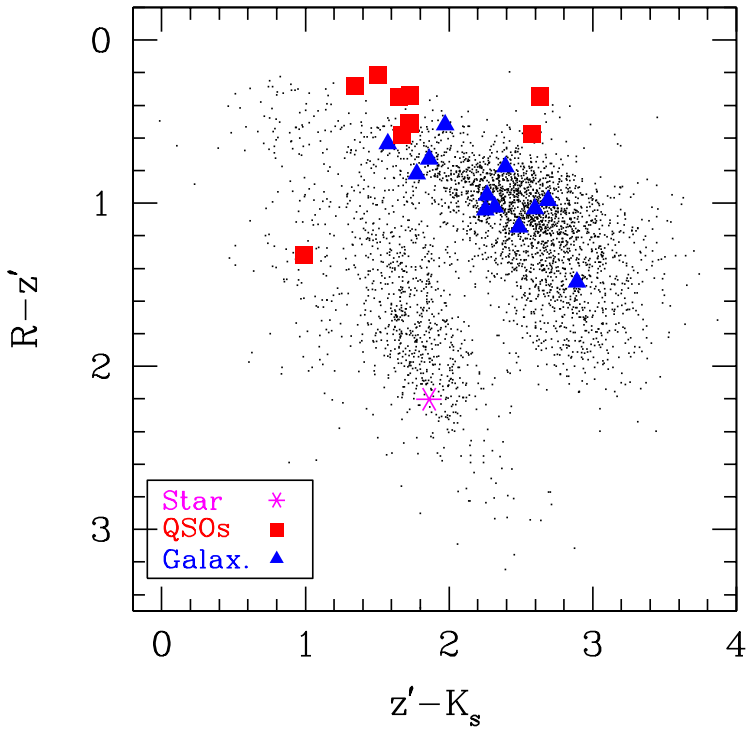

Figure 3.9: Distribution in the color–color diagram of the spectroscopically identified galaxies (triangles), stars (asterisks) and QSOs (squares). The galaxies occupy the "extended–source" region, demonstrating the efficiency of the criteria used for the morphological classification of the matched $Rz'K_s$ detections.

Fig. 3.10 shows the $(R - z')$ (lower panel) and $(z' - K_s)$ (upper panel) color index as a function of redshift. The galaxies and the QSOs are plotted as triangles and squares, respectively.

Fig. 3.11 shows the R–band (left panel) and z'–band (right panel) stellarity index for the spectroscopically identified galaxies (dashed line) and QSOs (solid line). The stellarity index for the QSOs/galaxies form two well separated distributions. Fig. 3.12 shows the z'–band versus R–band stellarity index for the 2dF spectroscopically identified galaxies (triangles) and AGN (squares). The good correlation of the stellarity indices in the two bands, for the QSOs and galaxies, is obvious.

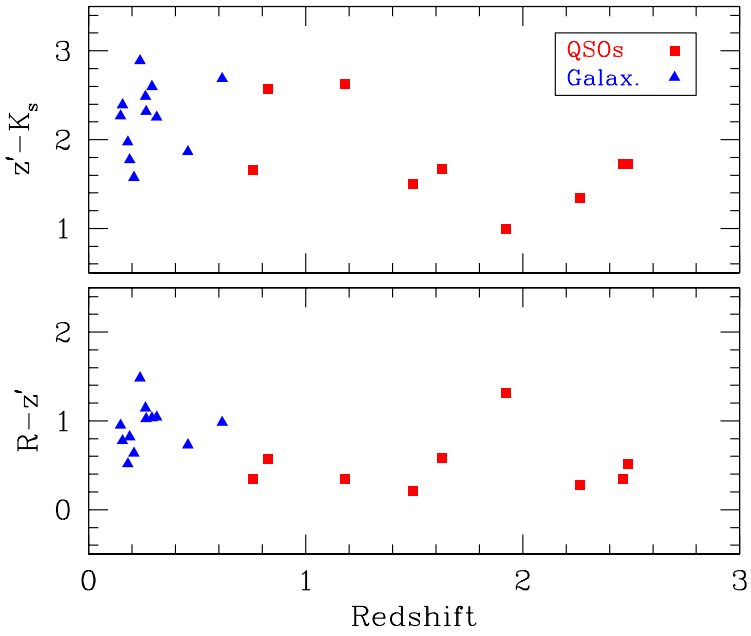

Figure 3.10: Color index as a function of redshift for the spectroscopically identified QSOs (squares) and galaxies (triangles).

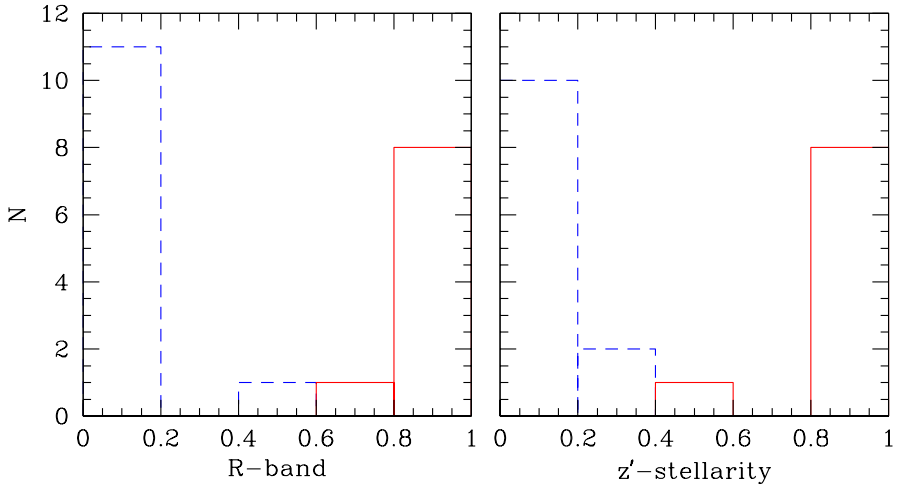

Figure 3.11: Histogram for the R–band (left panel) and z'–band (right panel) stellarity index for the spectroscopically identified galaxies (dashed line) and QSOs (solid line).

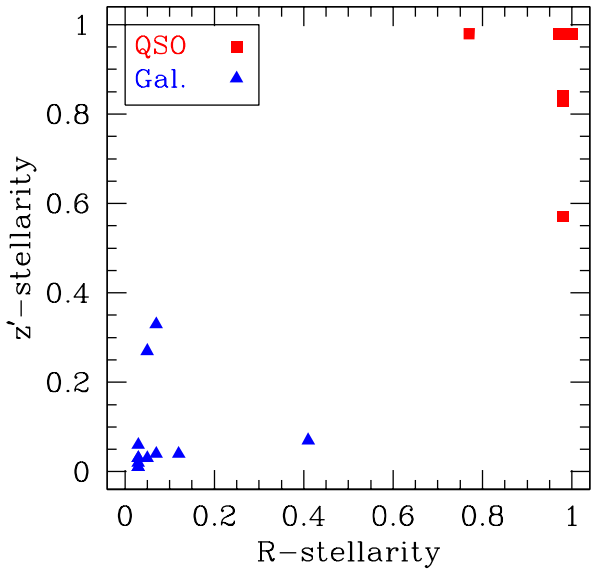

Figure 3.12: z'–band versus R–band stellarity index for the 2dF spectroscopically identified galaxies (triangles) and QSOs (squares).

3.3 Constraining the boundaries of the stellar locus

Characterization of the two distinct regions in the color–color plane, as "point–like" and "extended", has been based on a combination of morphological and magnitude criteria. An additional element, apart from the 2dF spectra, that can bring valuable information concerning the boundaries between the point–like/extended locus, are simulated stellar colors, through the use of spectral templates. In other words, by re–producing the $(R-z')$ and $(z'-K_s)$ colors of stars of various spectral types, and by examining their distribution in the $Rz'K_s$ plane, we can set better constraints on the stellar locus boundaries. The color index CI between two filters, A and B, for an object whose spectral energy distribution (SED) is described by a function f_λ, is computed according to the formula:

$$CI_{AB} = m_A - m_B = -2.5 \log \frac{\int_0^\infty S_\lambda(A) f_\lambda d\lambda}{\int_0^\infty S_\lambda(B) f_\lambda d\lambda} \tag{3.8}$$

where S_λ denotes the combined telescope–detector–filter sensitivity, convolved with the atmospheric transmission.

To simulate the stellar colors we used the spectral templates from the Pickles library (Pickles 1998), which contains ~ 130 spectra in the wavelength range 1150–25,000 Å. The spectral combination was done in two stages. First, the library was formed from the ultraviolet, optical and near–infrared sources in the wavelength range 1150–10,620 Å. At a second stage, the spectra were combined with additional infrared data in order to extend the library up to 25,000 Å. The resolution of the spectra used to generate the

library was 10 Å. The wavelength sampling in the catalog is 5Å. The number of spectra per spectral type range is presented in Table 3.3.

No of spectra	Sp. Type Range
45	O5–M6 V
14	B2–K3 IV
41	O8–M10 III
08	B2–M3 II
18	B0–M2 I

Table 3.3: Number of spectra (left column) corresponding to a given range of SEDs, from the Pickles library.

The S_λ term of Eq. (3.8), for the R, z' and K_s–band filters, was computed based on the real filters used for the CTIO and LCO observations, which were convolved with the quantum efficiency of the detector, the optical elements and the atmospheric transmission,

The distribution in the color–color diagram of the simulated stellar colors is shown in Fig. 3.13. In order to define the stellar locus boundaries, we first traced the distribution of the simulated colors with a second order polynomial for $z' - K_s \leq 1.8$, and with a cubic spline for redder values (thick line). The $\pm 3\sigma$ error–bar, both in the vertical and horizontal direction, around each point of the line tracing the simulated colors, will define the shape (i.e. horizontal/vertical width) of the stellar locus at the given position. Based on the standard deviation histograms presented in Fig. 3.6, the mean value for the magnitude errors for the R, z' and K_s–band data is of the order of 0.02, 0.02 and 0.08, respectively. Using as a reference the fitted line that traces the positions of the simulated colors, we computed the borders of the stellar region using the 3σ error–bars for the $\sigma_{(R-z')}$ and $\sigma_{(z'-K_s)}$ uncertainties. Leftward to the stellar locus shown in Fig. 3.13 the identification for about 15 different spectral types of stars can be also seen. The black dots denote the full $Rz'K_s$ catalog.

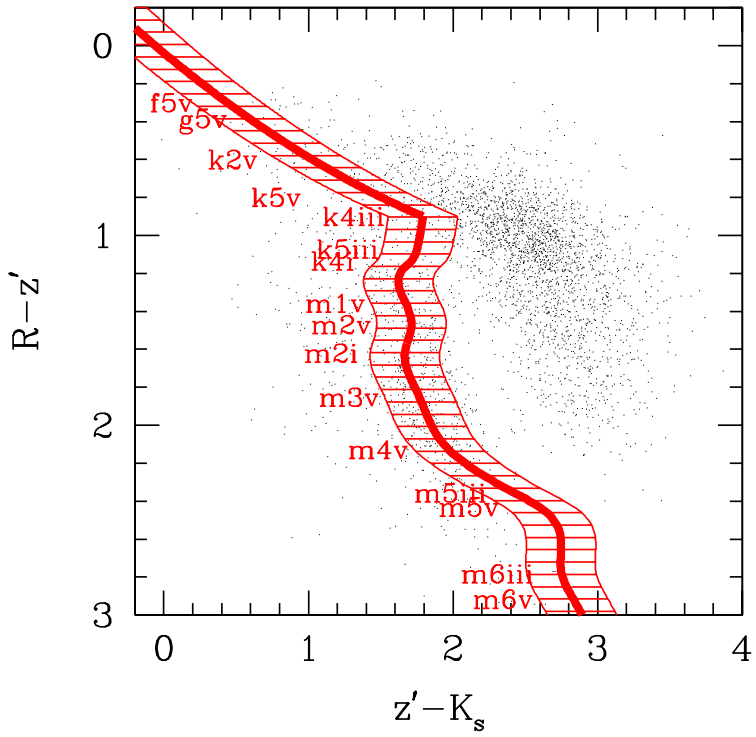

Figure 3.13: Distribution of simulated stellar colors in the color plane, based on SEDs from the Pickles library (Pickles 1998). Based on individual points it was possible to set the broader boundaries which define the stellar locus (see text for details).

3.4 Selection of the QSO candidates

The study of the objects with 2dF spectral identification and the work on the simulated stellar colors have played a significant role in the classification of the three populations present in the $Rz'K_s$ color–color diagram, i.e. stars, quasars and galaxies. The simulated stellar colors confirm that the region occupied by point–like objects is also, largely, the region of objects displaying stellar colors. At the same time, the shape of the locus implies that the boundaries separating the point–like from the extended sources have to respect a "break" around $(R-z',z'-K_s)= (1,2)$ and possibly a second one at about (2,2). The 2dF spectra have also confirmed that the galaxies occupy the region of objects with extended morphology.

The flow chart presented in Fig. 3.14 describes the logical steps followed for selecting the QSO candidates. According to their IR colors, the objects are split in two categories, objects having stellar–like or galaxy–like colors. Objects belonging to each category are then also classified, on the basis of their morphology. More analytically:

- About 860 sources (corresponding to $\sim 25\%$ of the matched $Rz'K_s$ catalog), having stellar colors and point–like morphology, were classified as **stars**.

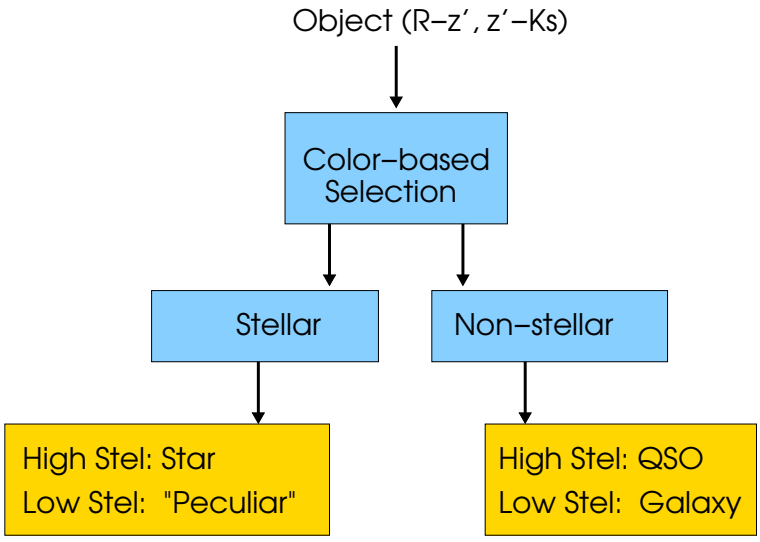

Figure 3.14: Flow chart describing the source classification procedure.

- Another $\sim$ 150 objects (4.5%) with stellar colors but extended morphology were characterized as "**peculiar**". Due to the low number of objects, each source was inspected in the $Rz'K_s$ images, in order to inquire more on the color–morphology discrepancy. For the majority of the sources the discrepancy was due to erroneous photometry, for one of the reasons mentioned in section 3.1. Only 22 objects were finally considered as "interesting".

- Similarly, $\sim$ 2300 objects (68%) having non–stellar colors and extended morphology were considered as **galaxy candidates**. The distribution in the color–color diagram of objects characterized as stars, galaxies and "peculiar" is shown in Fig 3.15.

- 96 sources (2.5%) with non–stellar colors and point–like morphology were considered as the *quasar candidates*. The objects were visually inspected in the corresponding R, z' and K_s images. Three candidates had to be removed from the catalog, since their galaxy–like colors were attributed to inaccurate photometry. Another eight sources were matched to 2dF spectroscopically identified quasars (marked as "Q" in Table 3.2). Hence, *the final list of quasar candidates contained 85 entries*. The distribution in the color–color diagram of the 85 QSO candidates and the QSOs with spectroscopic identification is presented in Fig. 3.16. A histogram of the R (left panel), z' (middle panel) and K_s–band (right panel) magnitudes of the QSO candidates is shown in Fig. 3.17. A list of the sources with their coordinates, magnitudes and other relative information can be found in Appendix B.

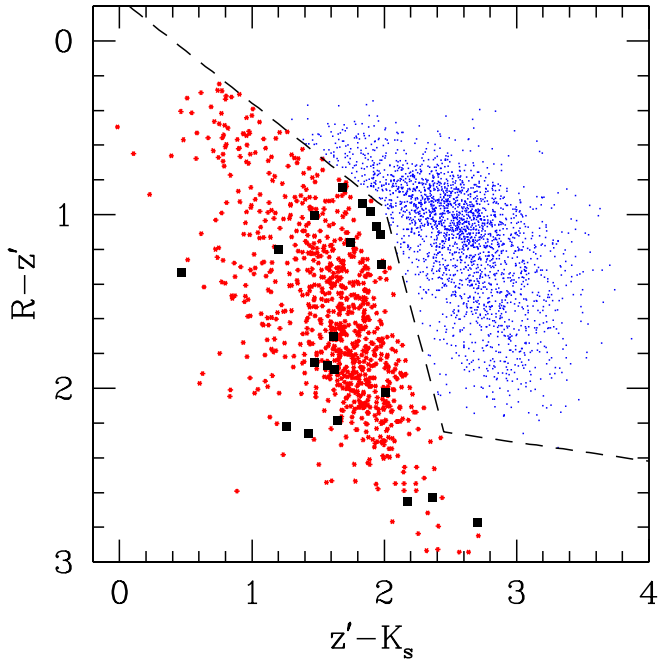

Figure 3.15: Distribution in the color–color space of the objects characterized as stars (red dots), galaxies (blue dots) and the objects characterized as "peculiar" (filled squares).

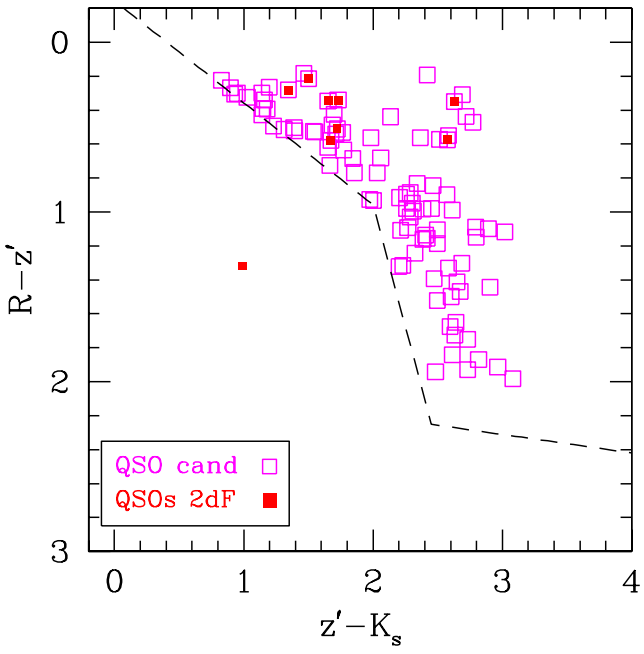

Figure 3.16: Distribution in the color–color space of the QSO candidates (open squares) and the nine qsos with spectroscopic identification (filled squares).

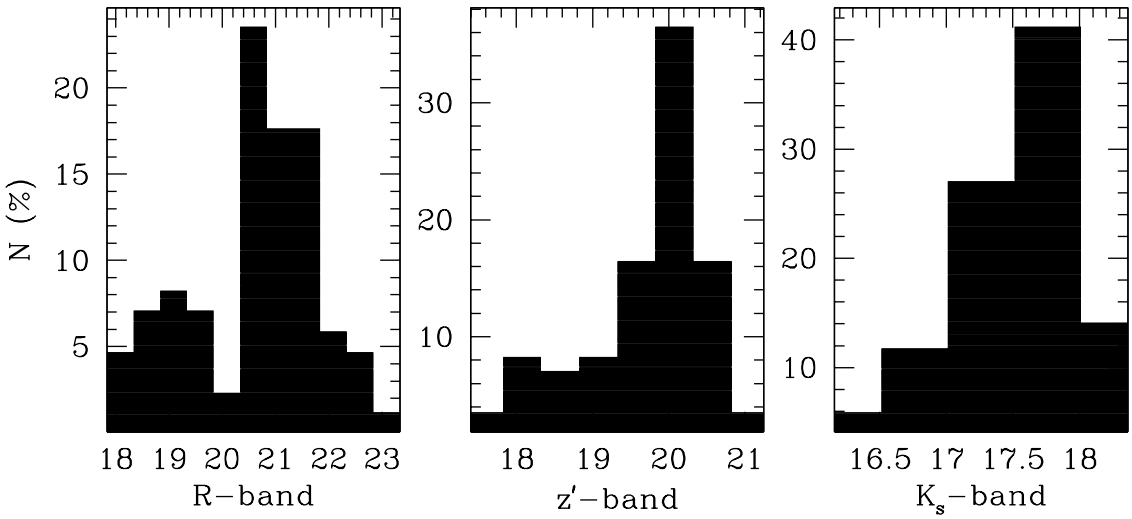

Figure 3.17: Histogram (in percentage) of the R (left panel), z' (middle panel) and K_s–band (right panel) magnitudes for the 85 QSO candidates. The binsize is 0.5 magnitude.

Chapter 4

Multi–wavelength properties of the quasar candidate sample

A general introduction on the physical processes in active galactic nuclei as a function of wavelength has already been given in chapter 1. The goal of chapter 4 is to study the multi–wavelength properties of the quasar candidates, using available catalogs from the legacy programs associated to the XMM–LSS survey. In the following sections we will explore the X–ray, optical and infrared characteristics of the KX–selected candidates, hoping to better understand their emission in the different wavebands, associate the nature of the objects to the related physical processes and finally get an insight on the type of sources selected by the KX technique.

4.1 X–ray properties

4.1.1 The XMM–LSS & XMDS surveys

A general description of the XMM–LSS survey (Pierre et al. 2004) has already been given in section 1.7. The XMM–Newton satellite will observe a total of 8×8 deg^2. Thanks to its design, XMM–LSS will generate the most complete and deepest, X–ray selected, map of the large scale distribution of matter at redshifts $0 < z < 1$, by studying some 800 clusters of galaxies in two redshift bins. It will, therefore, be possible to probe the nature and amount of dark matter, the initial fluctuation spectrum and other fundamental cosmological parameters. With its sensitivity, the XMM–LSS survey will also identify $\sim$ 200 QSOs/AGN per deg^2, out to redshifts of $z \sim 4$. By mapping the spatial distribution of AGN/QSOs, as outlined by the cluster/group population, we hope to arrive to a better understanding of the origin of active galaxies in terms of the initial density perturbation and galaxy interactions.

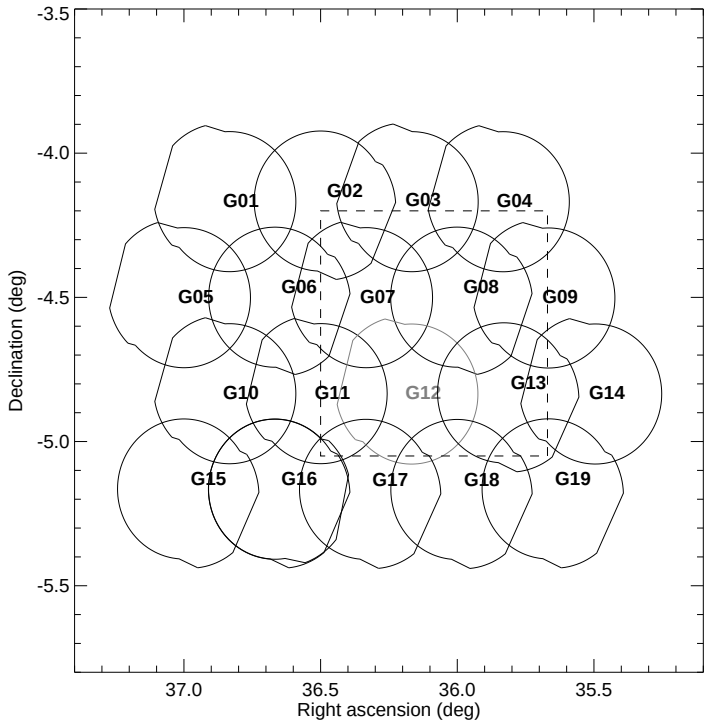

Figure 4.1: Location of the XMDS pointings on the sky, for the MOS1+MOS2+pn mosaic images (solid lines). Pointing G12 was background dominated and was finally not used (in gray color). The dashed rectangle represents the area covered by the Las Campanas observations.

The XMDS (XMM Medium Deep Survey,) is an X–ray survey which lies in the heart of the XMM–LSS field, covering a contiguous area of approximately 2 deg^2. For compensating for its smaller area, the XMDS survey goes much deeper than the XMM–LSS observations (the typical total exposure time of the survey's pointings is $\sim 20 - 25$ ksec, while for the XMM–LSS it is 10 ksec). Multi–wavelength (Optical/NIR/Radio) surveys (imaging and spectroscopy), associated to XMDS, will help to maximize our understanding of the high–redshift universe by targeting the X–ray selected sources. More details about the survey can be found in Chiappetti et al. (2005), while updates information related to XMM–LSS and XMDS can be found at the following URL:
http://vela.astro.ulg.ac.be/themes/spatial/XMM/LSS/index_e.html

The XMDS observations consist of 19 overlapping pointings (typical duration in the range 20–25 ksec), covering a contiguous area of ~ 3 deg^2. The XMDS observations were performed between July 2001 and January 2003, using the European Photon Imaging Camera (EPIC) on board the XMM–Newton satellite. EPIC consists of two metal–oxide (MOS) CCD arrays, plus a third CCD camera, the *pn*, which differs from the first two arrays in terms of semiconductor technology. An extended description of the MOS and *pn* cameras can be found in Turner et al. (2001) and Strüder et al. (2001), respectively. The

energy range of the EPIC instruments is between 0.1 to 12 keV. Thanks to the excellent relative astrometric calibration between the three detectors ($\sim 1''$), it is possible, for each energy band, to build up a mosaic image from the MOS1, MOS2 and *pn* frames. The location of the XMDS pointings on the sky, together with the area covered by the Las Campanas observations are given in Fig. 4.1.

4.1.2 The X–ray data

Although the EPIC instruments consist of CCD arrays, the direct output of the XMM observations is not images but event files, where each photon recorded by the detector is listed along with its arrival time, position and some information related to the cloud of free electrons generated by the photon arrival on the detector. Operating the cameras continuously avoids the use of shutters; for a space born mission this is a significant advantage, since the risk of technical failures is considerably reduced.

Starting from the event files, it is possible to generate images using standard pre–reduction tasks from the XMM Science Analysis System (XMM–SAS). Detailed information on the data preparation can be found in Baldi et al. (2002). At a first step, the event files are cleaned for hot pixels and soft proton flares. Based on the clean event files, the registered entries are separated, depending on the energy of the in–falling photons, in order to generate MOS1, MOS2 and *pn* images at different energy bands. The energy width of the different bands is user–defined and depends on the science we want to extract from the X–ray data. For our case, images were generated for five energy bands, designated as: A (0.3–0.5 keV), B (0.5–2 keV), C (2–4.5 keV), D (4.5–10 keV) and CD (2–10 keV). For each science frame, a corresponding exposure map, that takes into account the spatial variation of the quantum efficiency of each detector, the mirror vignetting and the field of view of each instrument, is also generated.

Fig. 4.2 shows the science frames (left column) and exposure maps (right column), for the G08 pointing, for the detectors MOS1 (upper row), MOS2 (middle row) and *pn* (lower row). The maximum (minimum) exposure time white (dark blue) region for the MOS1 and MOS2 exposure maps is of the order of 20 (7) ksec, while for the *pn* $\simeq$ 15 (5) ksec. The color scaling of the science frames has been chosen on purpose, to demonstrate the peculiarity of the X–ray data: the maximum number of photons, per pixel, registered on the MOS1, MOS2 and *pn* detectors is just 14, 12 and 32, respectively. The CCD arrays composing the EPIC detectors are clearly identified in the exposure maps. The MOS1 and MOS2 detectors are rotated by 90 degrees relative to each other.

The object detection is performed on the stacked images (MOS 1 + MOS 2 + *pn*), to take advantage of their improved signal–to–noise ratio (in order to preserve Poisson statistics, because of the very low number of counts, the detection is not performed on the time–divided mosaics). By removing the sources from the science mosaics a so called "cheesed" image is generated. This image is fit with a low order polynomial, to produce a background map. For any detection, the source counts S and its error σ_S are computed

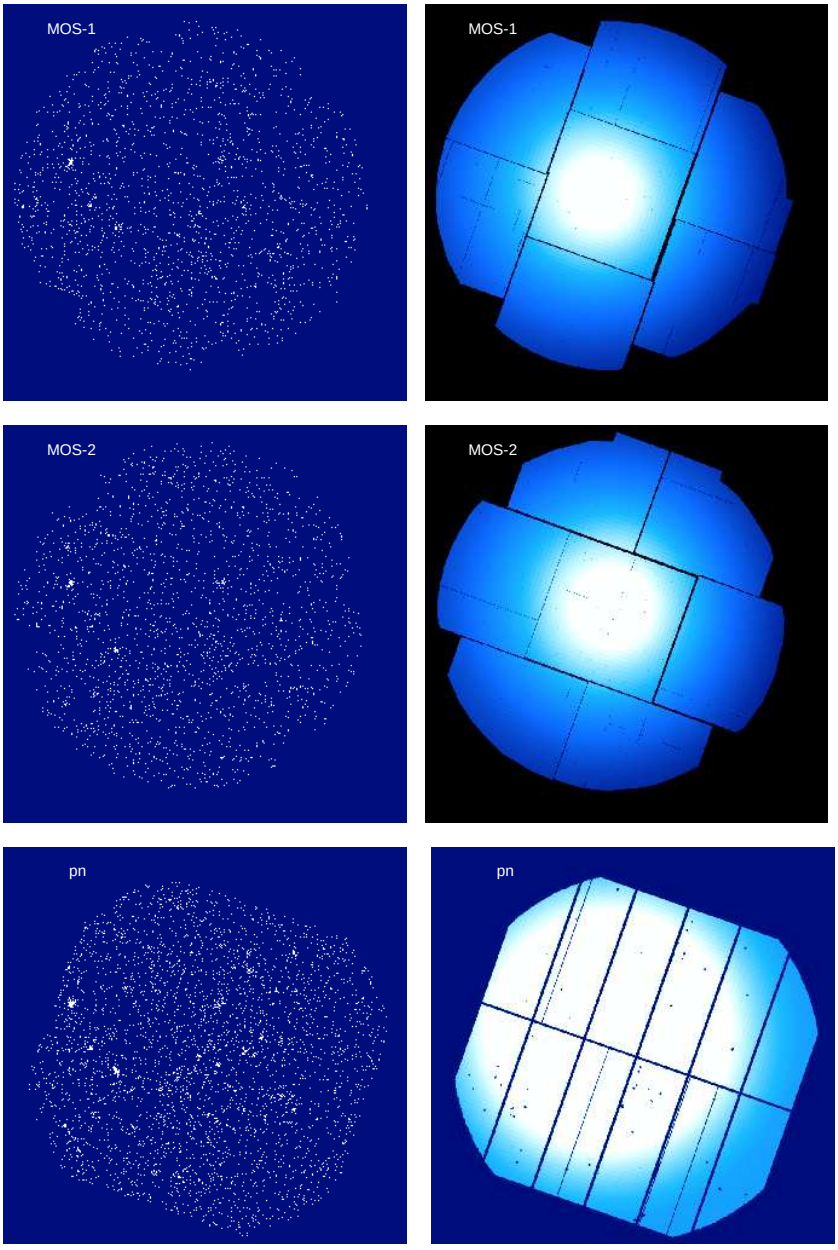

Figure 4.2: Science frames (left column) and exposure maps (right column), for the G08 pointing, for the detectors MOS1 (upper row), MOS2 (middle row) and *pn* (lower row). The maximum exposure time (white region) for the MOS and pn exposure maps is of the order of 20 and 15 ksec, respectively. In the exposure maps, the CCD arrays composing the EPIC detectors are clearly identified. The MOS1 and MOS2 detectors are rotated by 90 degrees relative to each other.

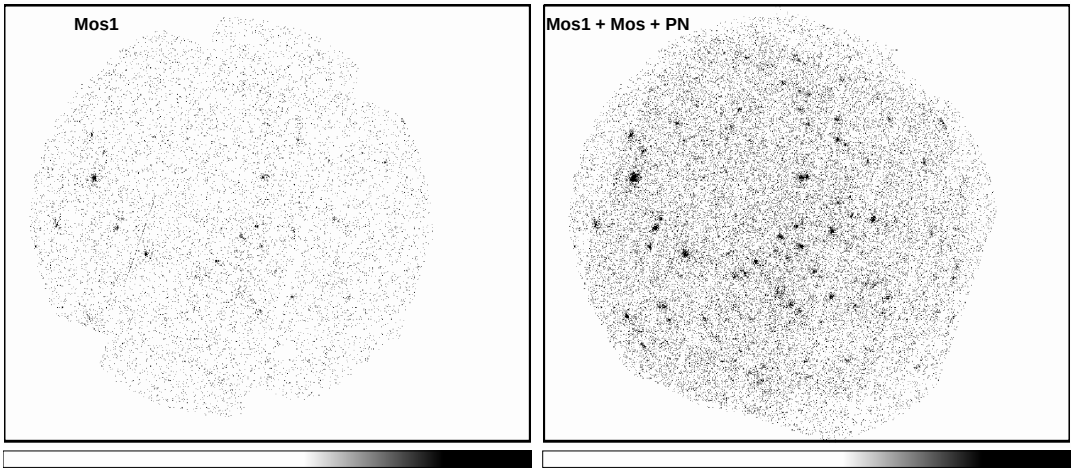

Figure 4.3: MOS1 (left panel) and mosaic image (MOS1 + MOS2 + *pn*) for the G08 pointing. Thanks to the excellent relative astrometry between the three cameras, it is possible to coadd the science frames in order to reach fainter sensitivities.

according to the formulas:

$$S = \frac{\mathrm{counts_{img}} - \mathrm{counts_{bkg}}}{\alpha} \qquad \sigma_S = \frac{1 + \sqrt{\mathrm{counts_{img}} + 0.75}}{\alpha} \tag{4.1}$$

where $\alpha = 0.68$, and $\mathrm{counts_{img}}, \mathrm{counts_{bkg}}$ represent the number of counts (measured in the same area) in the science mosaic and background map, respectively, using a radius within which 68% of the total PSF's energy is encircled. Since the PSF is varying as a function of position (on/off optical axis), instrument (MOS/*pn*) and energy band, using a flexible aperture allows to measure the source counts independently from the background fluctuations. More details on the PSF variations can be found in Ehle et al. (2005). Fig. 4.3 shows the MOS1 (left panel) and mosaic image (MOS1 + MOS2 + *pn*) for the G08 pointing, clearly demonstrating that the co–added image favors the detection of fainter sources.

The count rate is defined as:

$$\mathrm{CR} = \frac{S}{T_{\mathrm{MOS1}} + T_{\mathrm{MOS2}} + T_{pn}} \tag{4.2}$$

where T_{MOS1},T_{MOS2} and T_{pn} are the total exposure times, for the three detectors, measured on the corresponding exposure maps.

The count rate–to–flux conversion factors are computed for each instrument using the latest response matrices and assuming a power–law spectrum with photon index $\Gamma = 1.7$ and average galactic column density in the XMDS direction $N_H = 2.61 \times 10^{20}\,\mathrm{cm}^{-2}$ (Chiappetti et al. 2005). The total conversion factor (CF) has been calculated using the exposure times and the conversion factors for the three EPIC detectors, according to the

formula:

$$\frac{T_{\rm tot}}{\rm CF} = \frac{T_{\rm MOS1}}{\rm CF_{MOS1}} + \frac{T_{\rm MOS2}}{\rm CF_{MOS2}} + \frac{T_{pn}}{\rm CF_{\it pn}} \tag{4.3}$$

where $T_{tot} = (T_{\rm MOS1} + T_{\rm MOS2} + T_{pn})$, and $\rm CF_{MOS1}, CF_{MOS2}$ and $\rm CF_{\it pn}$ are the conversion factors for the MOS1, MOS2 and *pn* detectors, respectively. The exact values for the conversion factors in the two energy bands are summarized in Table 4.1. The two MOS detectors, being identical, have exactly the same conversion factors, in both energy bands.

Energy band	$\rm CF_{\it MOS}$	$\rm CF_{\it pn}$
B	$4.7500\,10^{-12}$	$1.350\,10^{-12}$
CD	$2.2075\,10^{-11}$	$7.595\,10^{-12}$

Table 4.1: Conversion factors for the MOS and *pn* detectors, for the B (0.5–2 keV) and CD (2–10 keV) energy bands. The units are $\rm erg\,cm^{-2}s^{-1}$.

Because of the very few counts in the X–ray data, it is necessary to associate each source with a detection probability, that will express the chance for the source to be a background fluctuation. In the general case of a binomial distribution, a given experiment may yield the event A or the event $\bar{A}$ (not A) with the probability $P(A) = p$ and $P(\bar{A}) = q = (1 - p)$, respectively. If the experiment is repeated n times, and X is the number of times A is obtained, then the probability of X taking exactly a value k is given by:

$$P(X = k) = \binom{n}{k} p^k\,(1-p)^{n-k}; \qquad k = 0, 1, 2, ..., n; \qquad 0 < p < 1 \tag{4.4}$$

with the binomial coefficients

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)...(n-k+1)}{1, 2, .., k}$$

The Poisson distribution can be defined as the limiting case of the binomial distribution, where for a large number of independent experiments ($n \to \infty$) only a very small fraction $p \times n$ is expected to yield events of a given type A ($p \times n = \lambda = \text{const}$). Based on Eq. 4.4, it can be proved that[1] in the Poisson's regime, the probability for $X = k$ events of type A to occur is:

$$P(X = k) = \frac{\lambda^k\, e^{-\lambda}}{k!} \tag{4.5}$$

Based on the definition for the measurement of the source counts, the probability that the counts of a detection originate from a background fluctuation can be expressed by the

[1] Bevington & Robinson (2002), p. 23

formula:

$$\sum_{n=\mathrm{counts_{img}}}^{\infty} e^{-\mathrm{counts_{bgk}}} \frac{\mathrm{counts_{bkg}^{n}}}{n!} > P \tag{4.6}$$

where P is the probability threshold used to separate real from false detections. For the XMDS catalog, only detections with $P < 2 \times 10^{-4}$ in at least one energy band were considered as real sources.

The above analysis helps understand the particularity of the X–ray data. Due to the very few counts registered on the detectors, it is difficult (a) to separate the real from the spurious detections and (b) to properly estimate the sky background. This particularity is illustrated in Fig. 4.4, where the left panel shows a B–band X–ray mosaic image (MOS 1 + MOS 2 + *pn*) and the right panel a K_s–band image of the same area on the sky. In the latter image, the circles indicate two K_s detections with no X–ray counterpart. In the X–ray image, in between the two circles, there is a bright source with no IR counterpart (marked with an ellipse). The peak intensity for this source, after a total integration time of 36 ksec, is just nine counts.

Despite the difficulties one encounters when working in the X–rays, this approach is one of the more advantageous for probing the AGN population. For a typical XMM exposure of 10 ksec, it is possible to detect at a $5\,\sigma$ level, in a mosaic image, point–like sources brighter than $\sim 3\,10^{-15}$ erg cm^{-2} s^{-1} in the [0.5–2] keV band. At such low flux, we expect to detect approximately 300 X–ray sources per square degree, among which 15–20 will be clusters of galaxies, about 200 AGN and QSOs and the remainder stars and nearby galaxies (Lehmann et al. 2001). Hence, with such a low contamination rate, studying X–ray selected samples is almost equivalent to probing the properties of X–ray selected AGN.

Another important benefit when studying X–ray selected samples is that they are free from the selection biases from which optical surveys suffer. Due to the presence of dust, optical light undergoes reddening and extinction. The impact of reddening is that it alters the intrinsic properties of the parent population. On the other hand, since extinction makes objects look fainter, it produces flux–limited samples and generates an additional bias. Contrary to optical light, the X–rays emitted from the heart of the active nucleus are much less affected by dust. As a result, X–ray selected samples not only reveal a different AGN population, but also give an insight on the intrinsic properties of the AGN.

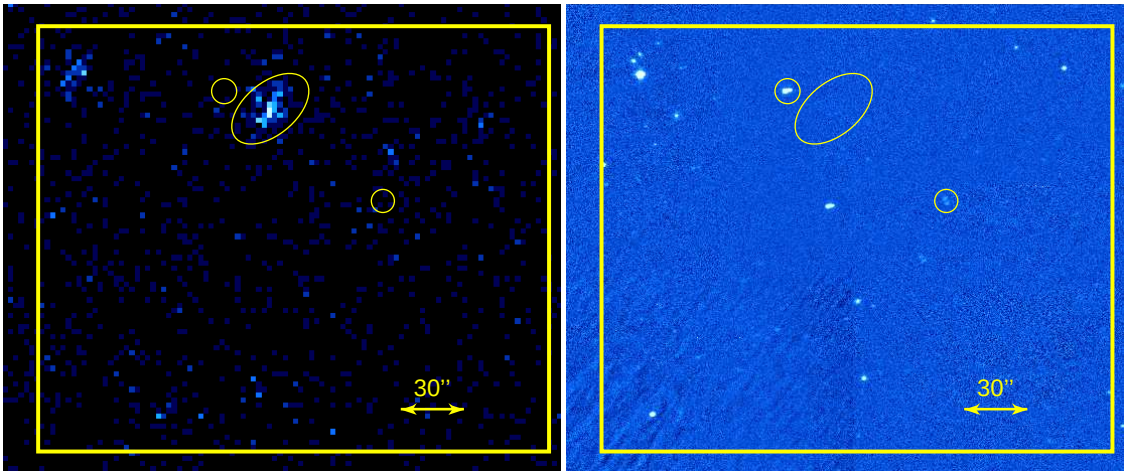

Figure 4.4: X–ray (left panel) and K_s–band (right panel) image of the same region of the sky. The X–ray image is a mosaic of the MOS1, MOS2 and *pn* B–band EPIC images. The two circles mark the position of two IR detections with no X–ray counterpart. In the left panel we clearly identify an X–ray detection in between the two circles, with no infrared counterpart (marked with an ellipse). The brightest of the pixels represented by the white color in the X–ray image has an intensity of just nine counts, after a total integration time of 36 ksec.

4.1.3 The X–ray catalog

The X–ray catalog matched to the $Rz'K_s$ detections was produced in Milan (version May 2005). A description of the pipeline and the properties of the catalog can be found in Chiappetti et al. (2005).

Before proceeding with the matching of the catalogs, it is essential to describe the main properties of the catalogs found in the XMM–LSS data base. First of all, because of XMM–Newton's large PSF ($\sim$ 6″), the term "point–like" does not have the same meaning as in optical/IR astronomy: extended sources are the clusters of galaxies, while point–like are all the other categories of astronomical objects (i.e. stars, galaxies and quasars). Hence, since we are interested in the X–ray properties of the quasars in the $Rz'K_s$ catalog, we only used the point–like source catalog. Another issue is that the pipeline that reduced the X–ray data produced separate catalogs for the different energy bands. As a consequence, it is possible that, for the same object, its (RA,Dec) vary by a few arcseconds in the different bands, depending on the object's morphology. A third point concerns the multiple entries of the same object, due to overlapping regions in adjacent pointings. Entries containing more than one source within a radius of 4″ were flagged and we verified that the counts of the different entries in the B and CD bands were indeed corresponding to the same detection. We then kept entries with more than 20 counts in the B (0.5–2 keV) band, to ensure that spurious detections were minimized (O.Garcet, private communication).

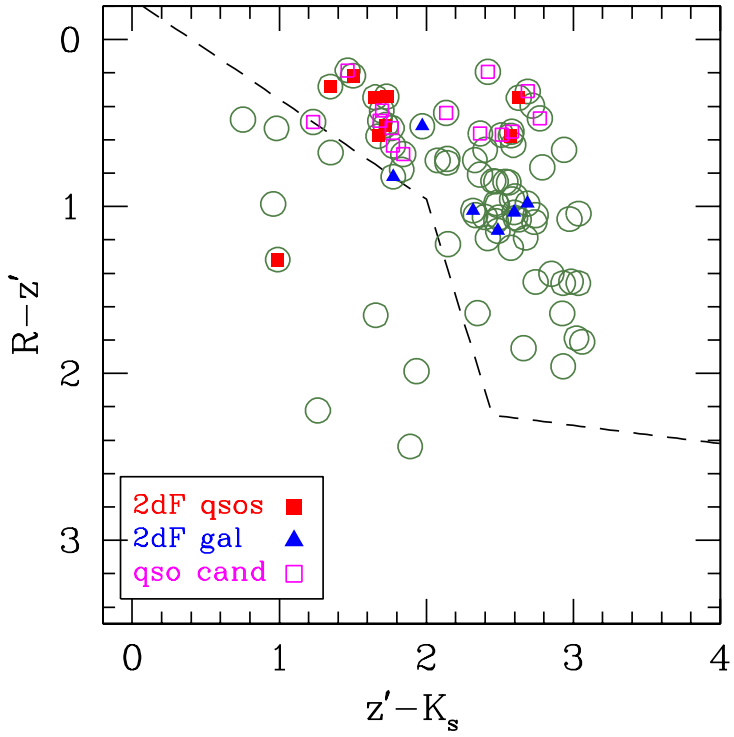

Figure 4.5: Distribution in the color–color diagram of 81 $Rz'K_s$ detections having an X–ray counterpart. The majority of the detections are found in the region occupied by galaxies and QSOs.

The B–band catalog complying to the above filters, corresponding to the Las Campanas area, contained some 370 sources. This catalog was matched to the $Rz'K_s$ catalog using a tolerance of 7″. Given the density of X–ray detections, the difference in the position in the various catalogs (a few arcseconds), the accuracy of the astrometric calibration ($\sim 1''$) and XMM's Point Spread Function (6″), we considered the 7″ to be a reasonable value. 81 sources were matched to a single X–ray detection, with 14 (out of 85) qso candidates, 6 (out of 12) galaxies and all 9 quasars with 2dF identification having an X–ray counterpart in at least one energy band.

The $Rz'K_s$ colors of the 81 sources matched to an X—ray detection are shown in Fig. 4.5. Precedented studies (see Figs. 3.8, 3.9 3.13), have demonstrated that rightward to the separating line in the $Rz'K_s$ color–color diagram is the region occupied by the objects with galaxy–like colors. As we expect the majority of the sources in the X–ray catalog to be AGN, we have one more element indicating that indeed quasars, in the IR color space, are located in the region occupied by galaxies.

We define the hardness ratio as $HR = (H - S)/(H + S)$, where H are the hard (2–10 keV) X–ray counts and S the soft X–ray counts (0.5–2 keV). According to this definition, objects lacking hard (soft) X–ray emission will have a hardness ratio of ≈ -1 (1). 14

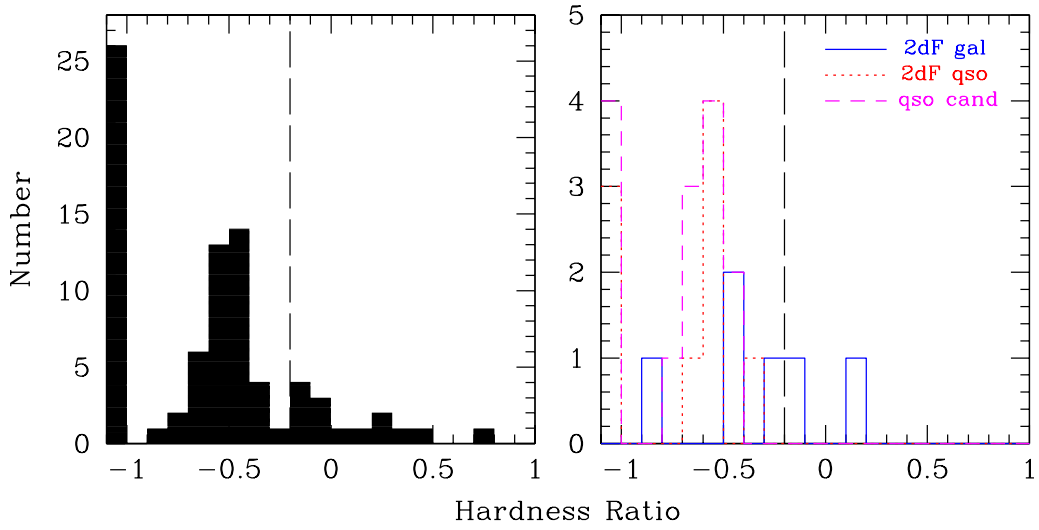

Figure 4.6: *Left panel:* Hardness ratio histogram for the 81 sources in the $Rz'K_s$ catalog with a single X–ray detection. The dashed line designates the limit between absorbed ($HR > -0.2$) and unabsorbed ($HR \leq -0.2$) sources. *Right panel:* Hardness ratio distribution for the 6 galaxies with 2dF identification (solid line), the 9 2dF QSOs (dotted line) and the 14 QSO candidates (dashed line) showing X–ray emission.

of the quasar candidates show X–ray emission in the soft band (corresponding to 16% of the total candidate sample), while 9 out of 14 (64%) show emission in both soft and hard energy bands. The hardness ratio histogram for all 81 objects with X–ray emission, together with the HR properties for the 2dF galaxies, quasars and the QSO candidates are presented in Fig. 4.6. According to Szokoly et al. (2004), AGN with $HR > -0.2$ are type II AGN, with significant part of their soft X–ray emission failing to get out of the dusty torus surrounding the AGN power house. From the HR distribution shown in the left panel we see that there are absorbed sources contained among the 81 $Rz'K_s$ sources showing X–ray emission. On the other hand, the lack of quasar candidates with a hardness ratio exceeding the −0.2 limit (right panel), implies that the sample KX–selected QSOs mainly consists of unobscured sources, which in terms of "optical identification" would be Type–I QSOs.

Additional information regarding the properties of our X–ray selected sample can be obtained by studying the X–ray to optical flux ratio, which, according to Chiappetti et al. (2005), is computed as:

$$\log(F_X/F_R) = \log F_X + 5.51 + R_{AB}/2.5 \tag{4.7}$$

where R_{AB} is the R–band magnitude of the sources in the AB system and F_X the X–ray flux in the CD (2–10 keV) energy band. The constant derives from the conversion of the AB magnitudes intro monochromatic flux and from the integration of the monochromatic flux over the filter band width, assuming a power–law spectral energy distribution with

index $\Gamma = 1.7$. Based on the explanations provided in paragraph 3.1.1, the R(Vega) magnitudes were converted to R(AB) as follows:

$$R_{AB} = R_{\rm vega} + 0.239$$

The conversion from counts to flux was performed according to Equations 4.2 and 4.3. The exposure times of each source for the MOS1, MOS2 and *pn* detectors were provided by the XMM–LSS data base, whilst the conversion factors were taken from Table 4.1.

The X–ray-to-optical flux ratio as a function of hardness ratio, for the 2dF galaxies, quasars and the QSO candidates, is plotted in the left panel of Fig. 4.7. All objects have $0.1 < \log(F_X/F_R) < 10.0$, which, according to Della Ceca et al. (2004) is the region occupied by the AGN. In the same plot, the region occupied by stars showing X-ray emission is indicated by the dashed line. The right panel shows the X–ray to optical ratio as a function of CD–band flux. The dashed lines indicate loci of constant R magnitude. This plot is also in very good agreement with the F_X/F_R versus F_X diagrams found in the literature (e.g. Fiore et al. 2003).

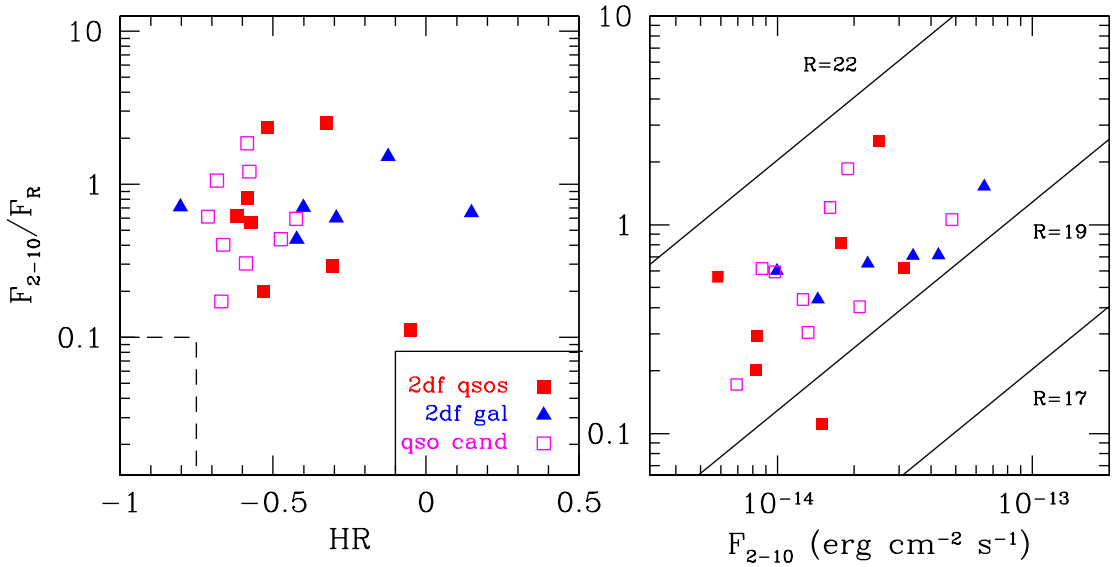

Figure 4.7: *Left panel:* X–ray-to-optical flux ratio for the QSO candidates (open squares), for the quasars (filled squares) and galaxies (triangles) with 2dF spectroscopic identification , as a function of hardness ratio. The dashed line defines the area occupied by stars (Della Ceca et al. 2004). *Right panel:* X–ray to optical flux ratio as a function of the CD band flux.

Stamp X–ray images of the sources presented in Table 4.2, for the hard (CD) and soft (B) energy bands, are shown in Figs. 4.8 and 4.9, respectively. The pixel size is $0.25''$ and the FoV for each image $1' \times 1'$. North is up, East is to the left. Depending on the number of pointings in which each source is present, individual subimages of $1' \times 1'$,

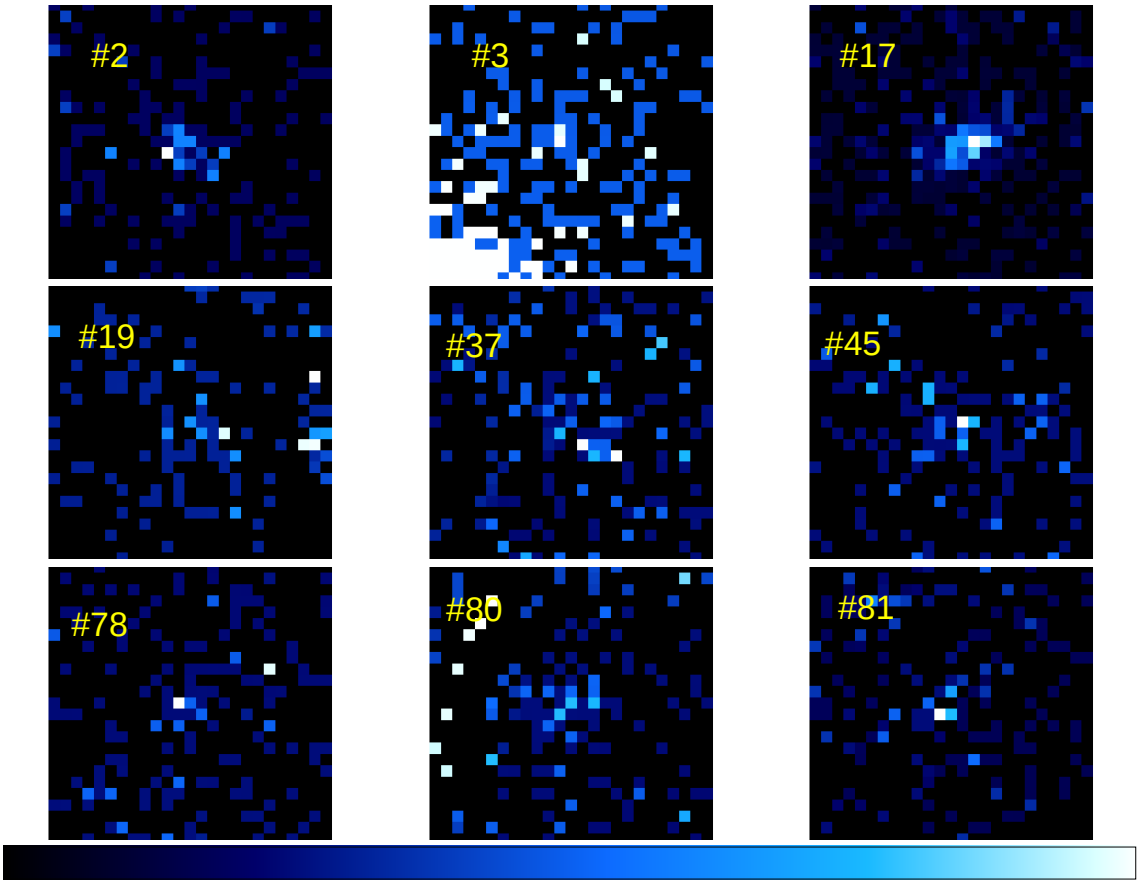

Figure 4.8: Stamp images for the 8 QSO candidates with X–ray emission in the hard band. The pixel scale is 2.5″pixel^{-1} and the field of view 1 × 1 arcminute. North is up, East is to the left. All stamp images are shown using the same scale, with exception object #3, for which a different scale was used for making visible the quasar emission. This was necessary, as the presence of a much brighter detection (seen south–ward of the central source) was biasing the statistics used for properly scaling the specific stamp–image.

centered around the (RA,Dec) coordinates of the source, have been extracted from the B and CD images of the given fields. If present in more than one pointings, the extracted sub–images were averaged, for improving the S/N ratio.

Table 4.2 summarizes the X–ray properties of the KX–selected QSO candidates and of the objects with 2dF spectroscopic identification.

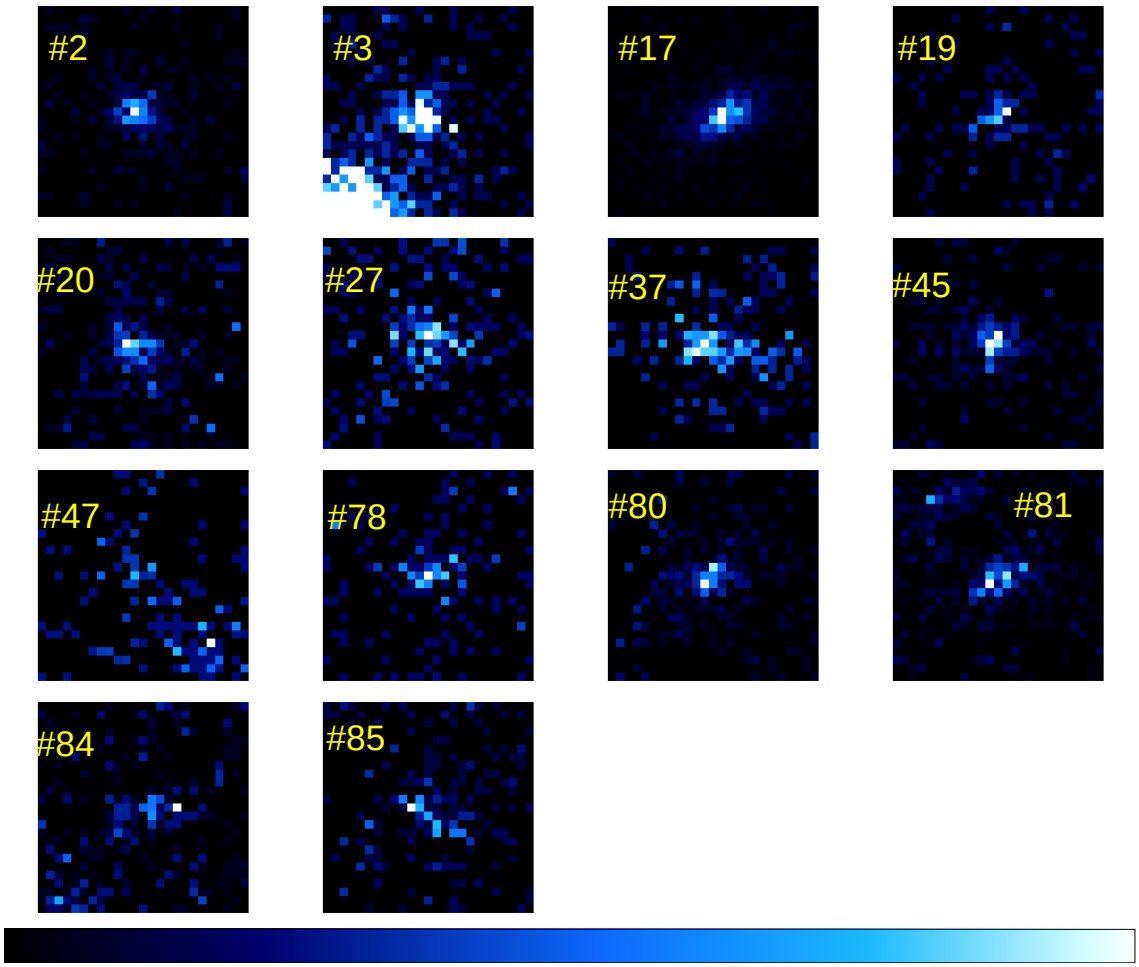

Figure 4.9: Stamp images of the 14 quasar candidates showing X–ray emission in the soft energy band, sharing the same properties as in Fig. 4.8. The identification of the quasar candidates is given in the upper left part of the stamps, with the exception of #81, due to the presence of another source close to the QSO candidate.

ID	RA (J2000)	Dec (J2000)	Type	B	CD	B+CD	$F_{0.5-2}$	F_{2-10}
	h m s	° ′ ″					erg s^{-1} cm^{-2}	erg s^{-1} cm^{-2}
735	2 22 44.40	−4 33 47.02	2dF	244.0350	64.0763	308.1113	1.29e-14	1.77e-14
736	2 22 47.90	−4 33 30.20	2dF	127.1680	64.6222	191.7902	8.84e-15	2.50e-14
97	2 23 26.47	−4 57 06.30	2dF	43.8462	39.6000	83.4462	8.70e-15	1.50e-14
100	2 23 54.82	−4 48 15.19	2dF	73.3729	−1.0000	73.3729	1.10e-14	–
251	2 24 29.14	−4 58 08.11	2dF	231.2730	54.8486	286.1216	2.30e-14	3.15e-14
23	2 25 14.40	−4 47 00.38	2dF	35.9802	19.1184	55.0986	2.81e-15	8.29e-15
41	2 25 40.61	−4 38 25.30	2dF	50.7523	13.7957	64.5480	3.87e-15	5.84e-15
11	2 25 56.83	−4 58 53.29	2dF	109.1030	33.3524	142.4554	5.20e-15	8.21e-15
02	2 22 42.60	−4 30 18.14	C	247.4970	66.4224	313.9194	1.17e-14	1.61e-14
03	2 22 42.96	−4 33 14.54	C	198.0980	33.2661	231.3641	9.97e-15	8.71e-15
17	2 22 58.80	−4 58 52.14	C	457.9600	86.1830	544.1430	4.47e-14	4.83e-14
19	2 23 04.20	−4 44 35.30	C	76.1010	30.7380	106.8390	4.49e-15	9.78e-15
20	2 23 06.00	−4 33 23.62	C	109.4160	−1.0000	109.4160	7.70e-15	–
27	2 23 25.80	−4 22 54.05	C	28.5507	−1.0000	28.5507	2.97e-15	–
37	2 23 40.92	−4 22 55.27	C	79.1324	15.6951	94.8275	6.33e-15	6.92e-15
45	2 23 50.64	−4 31 57.94	C	172.0840	44.6629	216.7469	9.62e-15	1.32e-14
47	2 23 52.44	−4 18 21.35	C	25.4118	−1.0000	25.4118	2.57e-15	–
78	2 25 15.24	−4 40 08.69	C	82.6132	29.4764	112.0896	6.24e-15	1.26e-14
80	2 25 34.68	−4 24 01.40	C	193.4140	50.7085	244.1225	1.30e-14	1.89e-14
81	2 25 37.20	−4 21 32.69	C	203.0440	41.1917	244.2357	1.73e-14	2.10e-14
84	2 25 39.36	−4 22 27.77	C	70.7539	−1.0000	70.7539	5.82e-15	–
85	2 25 55.56	−4 39 18.07	C	48.6572	−1.0000	48.6572	5.45e-15	–

Table 4.2: X–ray properties for 9 spectroscopically identified quasars (2dF) and 13 QSO candidates (C) with X–ray emission, found in the Las Campanas field. The first column provides the identification number of the source, according to Tables 3.2 or B.1. The second and third column give the right ascension and declination of the optical counterparts of the X–ray detections, respectively. The units for the right ascension are hours, minutes, seconds and for the declination degrees, arcminutes, arcseconds. Column four indicates the type of object. Columns five, six and seven give the X–ray counts for the B $(0.5-2$ keV), CD $(2-10$ keV) and B+CD energy bands, respectively. The -1 value is used to flag the lack of detection in a given energy band. Columns eight and nine give the flux in the B and CD bands, respectively

4.2 Optical properties

4.2.1 The CFHT Legacy Survey

Thanks to the multi–wavelength observations covering the XMM–LSS area, the LCO field was observed with the 3.6 m CFHT telescope (Hawaii), as part of the CFHT Legacy Survey (CFHTLS). The CFHTLS is a major project, with more than 450 nights over 5 years devoted to surveying the sky using the wide field imager MegaPrime equipped with MegaCam. MegaCam consists of 40 2048 × 4612 pixel CCDs, covering a field–of–view of one square degree, with a resolution of 0.187 arcsecond per pixel (Fig. 4.10). CFHTLS will cover in total 170 square degrees in three patches of 49 to 72 square degrees through the whole filter set (u^*, g', r', i', z'). The final transmission curves of these five filters are shown in Fig. 4.11.

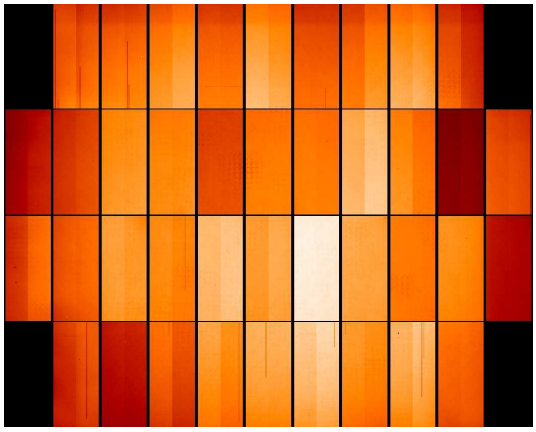

Figure 4.10: The MegaCam CCD mosaic.

The CFHTLS observations have not covered the Las Campanas area in a homogeneous way. Saturation effects, due to the presence of bright stars, reduce the effective area available. Moreover, due to the strategy followed during the CFHTLS observations, additional regions on the sky also lack CFHTLS coverage. As a result, approximately 40% of the LCO area has not yet been observed by CFHTLS. The exact overlapping between the two surveys is shown in Fig. 4.12. The squares denote the positions of the 85 quasar candidates. 30 sources, corresponding to 35% of the total sample, fall in regions without CFHTLS coverage.

The CFHTLS catalog covering the LCO area contained the positions and the photometry[2] (with their corresponding photometric accuracy) for ∼72,500 entries. For 27% of the sources the catalog contained no r'–band photometry, because of technical problems at the time of the observations. Due to the huge amount of data obtained with MegaCam, the

[2]The CFHTLS data are calibrated in the AB system.

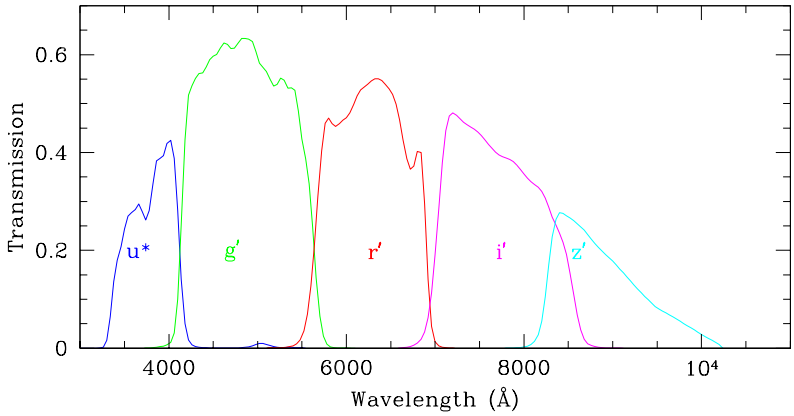

Figure 4.11: Final transmission curves for the MegaCam $u^*g'r'i'z'$ filters.

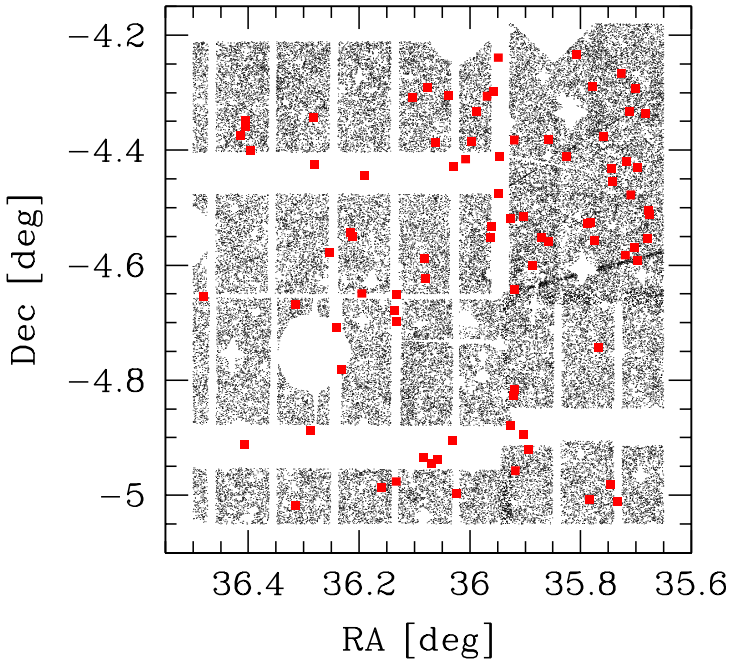

Figure 4.12: Field coverage of the Las Campanas area with the CFHTLS observations. The squares indicate the positions of the 85 candidates.

CFTLS pipeline processes the obtained images in an automatic way. As a consequence, a small fraction of the catalog entries are false or spurious. In order to remove these entries before any further catalog exploitation, we considered that the photometric information in the given band was meaningful only if the photometric accuracy was better than 0.5 mag. Using this threshold, the magnitude limit for the u^*, g', r', i', z' bands is 25.36, 26.6, 25, 24.7 and 24.7, respectively.

4.2.2 The optical color–color diagram

The $Rz'K_s$ and CFHTLS catalogs were correlated using a tolerance of $1''$, a value fairly reasonable for the astrometric precision of the catalogs and the object density of the CFHTLS survey. In this way 60% of the $Rz'K_s$ sources were matched to a single optical counterpart. Given the CFHTLS coverage of the Las Campanas area, this percentage agrees well with the expected value. For 28 of the 55 quasar candidates falling in regions with CFHTLS coverage we finally obtained photometry in all five u^*, g', r', i', z' bands, while the remaining 27 lack r'–band measurements. Fig. 4.13 shows four color–color diagrams, produced using the CFHTLS photometry of the $Rz'K_s$ sources. The corresponding colors of the spectroscopically identified quasars and galaxies and the quasar candidates (filled squares, triangles and empty squares, respectively) are overplotted. Due to the lack of r'–band photometry for some objects, the number the points is not the same in all plots. The colors of the spectroscopically identified objects presented in Fig. 4.14, are in full agreement with the results presented in Hatziminaoglou et al. (2005). Based on this, we can have a first evaluation regarding the selection of the quasar candidates. In all plots $30-40$% of the objects fall in the region occupied by the verified quasars. However, due to the lack of r'–band photometry for half of the objects, these numbers are only approximative and a more systematic study will be needed in order to better evaluate the KX selection.

Fig. 4.14 shows the same color–color diagrams as in Fig. 4.13, but for the full (i.e. not correlated) CFHTLS catalog covering the Las Campanas area (black dots). Obviously, the number density of the latter catalog is much higher compared to that of the $Rz'K_s$, which is flux limited to $K_s = 18$ (we remind that the limits for the R and z'–bands are 23.5 and 22, respectively). Overplotted, in red color, are the objects contained in the $Rz'K_s$ catalog. The distribution of the latter in the color–color diagrams of Fig. 4.14 clearly demonstrates that our infrared survey favors the detection of objects spanning a much more limited color range, with respect to the objects selected when combining the $u^*g'r'i'z'$ optical bands.

4.2.3 Reddening signature in the QSO candidates

The CFHTLS photometry can provide essential information on the candidate sample, since it allows to associate the optical properties of the sources with their properties at other wavelengths, such as in the X–rays and in the infrared. By better describing the characteristics of the candidates, we aim at extracting any possible information that will allow to identify the real quasars in the sample. The KX technique favors the selection of reddened and obscured quasars (Warren et al. 2000). Any reddening signature or correlation between X–rays and optical colors will lead to a better characterization of the quasar candidates.

A good criterion for studying the reddening is by measuring the $B-K$ color index of

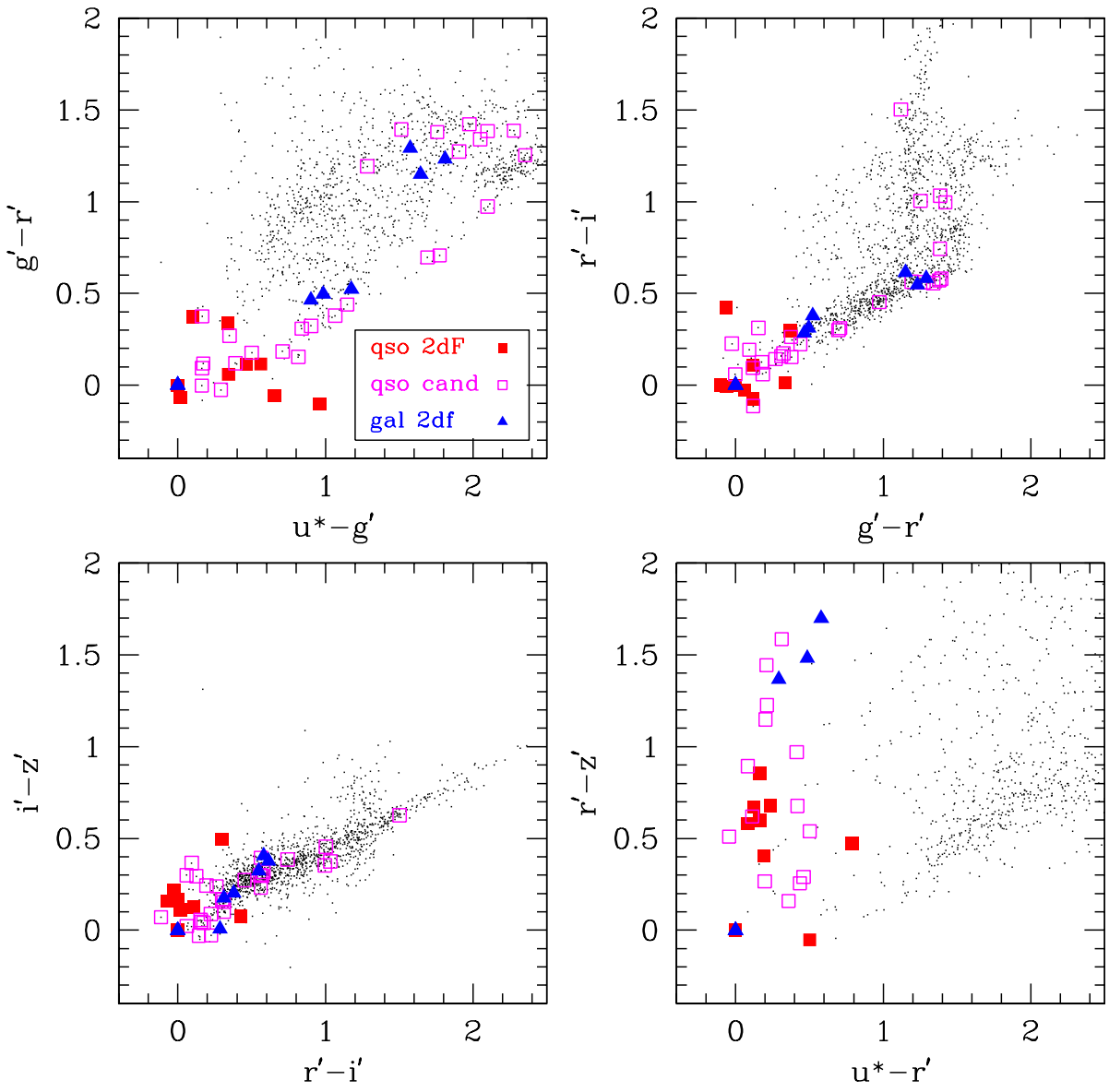

Figure 4.13: Color–color diagrams, based on the u^*, g', r', i', z' photometry for the $Rz'K_s$ sources with available CFHTLS data. The colors of the spectroscopically verified quasars, galaxies and the quasar candidates are overploted as filled squares, triangles and empty squares, respectively.

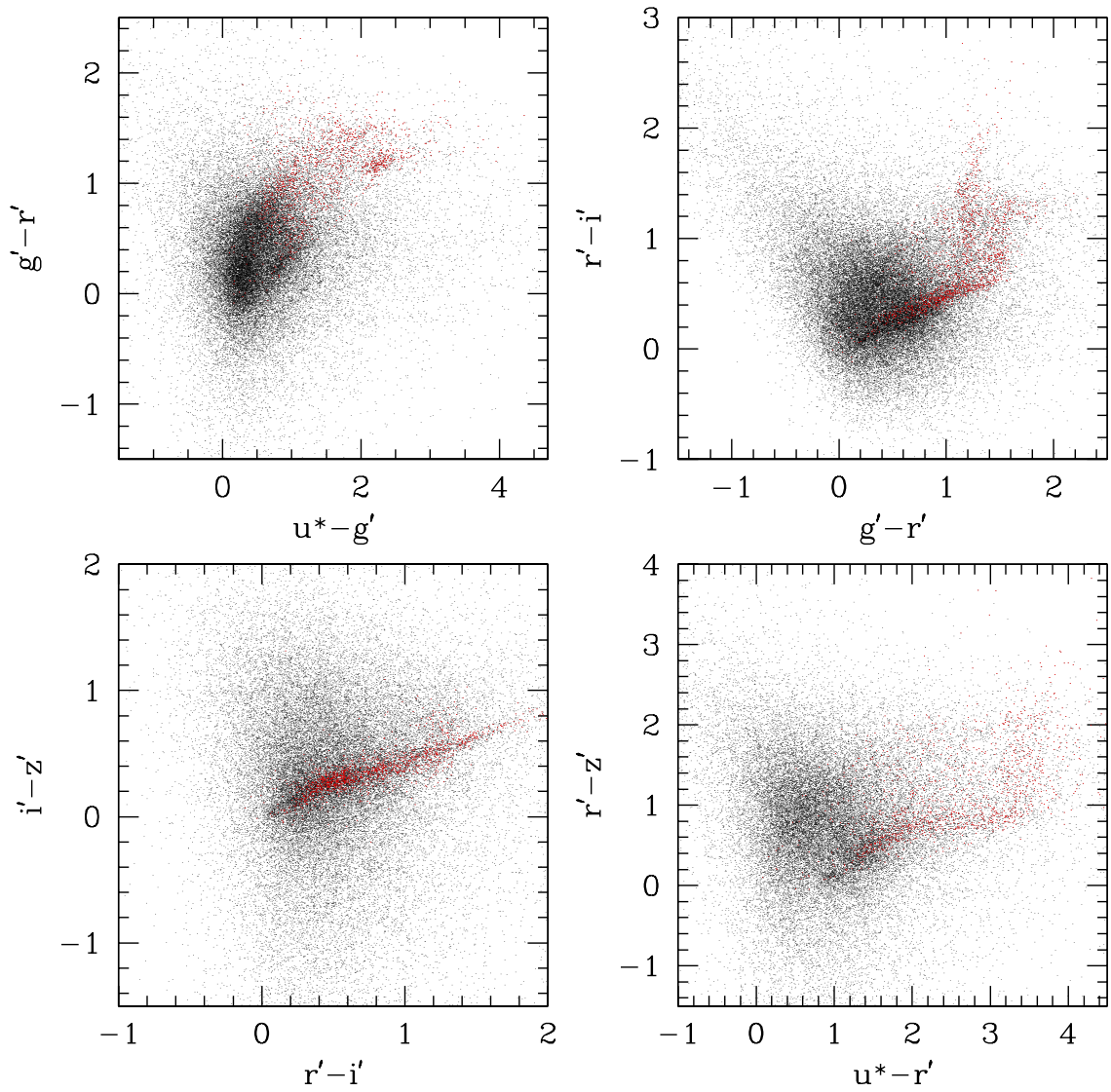

Figure 4.14: Color–color plots for the full CFHTLS catalog covering the LCO area (black dots). Overplotted, in red color, are the objects contained in the $Rz'K_s$ catalog.

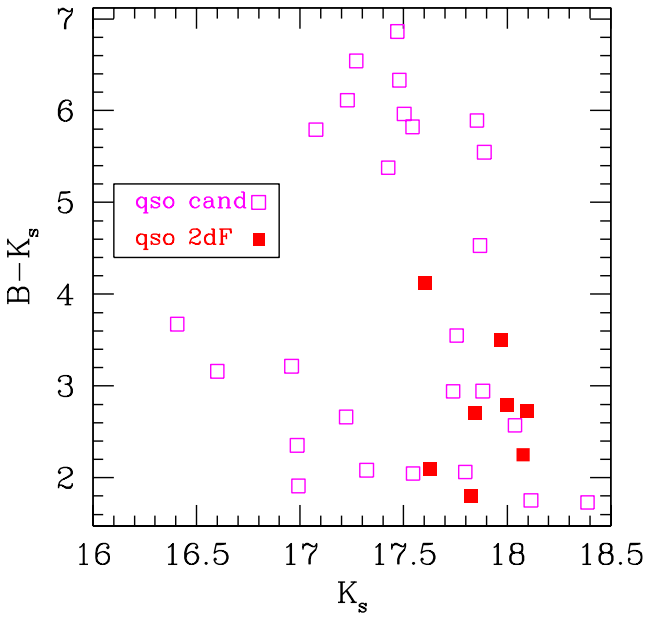

Figure 4.15: $B - K_s$ color index as a function of K_s–band magnitudes, where the spectroscopically confirmed quasars and the QSO candidates are plotted as filled and open squares, respectively.

the objects under question (Richards et al. 2003). Although the g' filter, the one closer to the B–band, could also serve for this study, we preferred to use the latter one for comparing our results to the ones presented in the literature (e.g. Webster et al. (1995)). In order to transform the CFHTLS photometry to Johnson B–band magnitudes we used the formula (Smith et al. 2002):

$$B = g' + 0.47\,(g' - r') + 0.17 \tag{4.8}$$

We have implemented the above transformation to both the quasar candidates and the spectroscopically confirmed QSOs with available g' and r' photometry. The distribution of the $B - K_s$ color indices as a function of the K_s–band magnitudes, for the quasar candidates (open squares) and the 2dF QSOs (filled squares) is shown in Fig. 4.15. From the plot it is clear that among the candidates there is a group of highly reddened sources, with $B - K_s > 5$. Quasars are quite variable objects and comparing observations taken at different epochs encompasses the risk of introducing observational biases (Benn et al. 1998). Nevertheless, the fact that we are comparing two single–epoch data sets minimizes this effect, and it would be rather unexpected for quasars to show such extreme magnitude differences, due to internal variability. Thus, Fig. 4.15 indeed demonstrates that there is a highly reddened population among the quasar candidate sample.

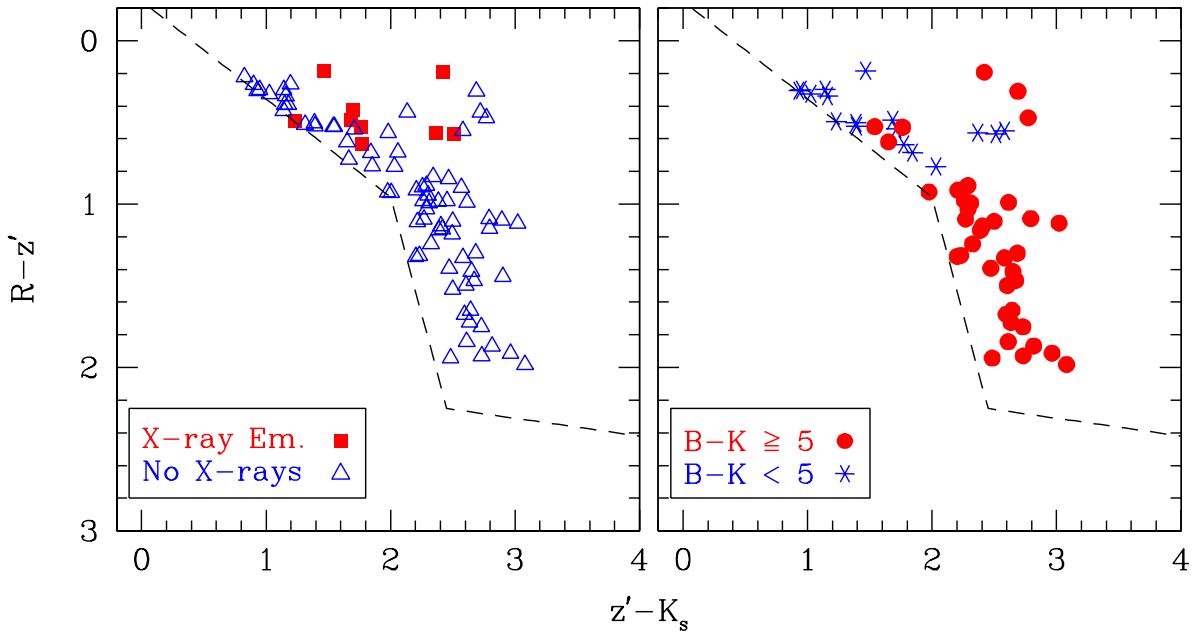

Figure 4.16: *Left Panel:* $Rz'K_s$ colors of the 85 QSO candidates with (filled squares) or without (empty triangles) X–ray emission in the B (0.5–2 keV) energy band. The dashed line separates the stellar from the non–stellar locus. *Right Panel:* Color distribution for 28 quasar candidates, flagged according to their $B - K_s$ color index.

4.2.4 Reddening–obscuration correlation

Given the conclusion derived from Fig. 4.15, the next logical step is to study a possible correlation between reddening in the optical and obscuration in the X–rays. A first effort to extract such a connection is shown in Fig. 4.16. The left panel shows the distribution in the $(R - z')$ versus $(z' - K_s)$ color–color diagram of the quasar candidates with and without X–ray emission in the [0.5−2 keV] energy band (filled squares and empty triangles, respectively). This plot demonstrates that the candidates having an X–ray counterpart are found in the bluer (for the $R - z'$ color) part of the locus. The 2dF spectra we have at our disposal reveal that this region is occupied by type I quasars. The right panel of Fig. 4.16 shows the distribution in the color–color diagram of the 28 quasars candidates for which it was possible to compute their $B - K_s$ color index. Sources undergoing low $(B - K_s < 5)$ or high $(B - K_s \geq 5)$ reddening are plotted as asterisks or filled circles, respectively. By comparing the two panels we see a clear correlation between the lack of X–ray emission and high reddening. A well plausible scenario explaining the lack of X–ray emission for the QSO candidates with $R - z' > 0.8$ could be that we are looking at a type II population with no X–ray emission. In such a case, the absorption of the soft X–ray emission by the torus surrounding the AGN 's central engine is expressed as reddening when observing in the optical and near–infrared wavelengths.

4.3 Infrared properties

4.3.1 The Spitzer Wide–Area Infrared Extragalactic survey (SWIRE)

The SWIRE survey (Lonsdale et al. 2003) is the largest of the Legacy Programs to be performed with the Spitzer Space Observatory[3]. It is a wide–field imaging survey, pointing to seven high–latitude fields, totaling 60−65 square degrees. On–board Spitzer there are two imaging instruments:

- IRAC (Infrared Array Camera (Fazio et al. 2004)) is a four–channel camera that provides simultaneous 5.2×5.2 arcminute^2 images at 3.6, 4.5, 5.8 and 8 microns. Two adjacent fields of view are imaged in pairs (3.6 and 5.8 microns; 4.5 and 8.0 microns) using dichroic beam–splitters. All four detector arrays in the camera are 256×256 pixels in size, with a pixel size of $1.2'' \times 1.2''$. The two short wavelength channels use InSb detector arrays, operating at ∼15K. The detectors for the two longer wavelength channels use Si:As technology and operate at ∼6K. In a 30 second exposure on a point source, the IRAC saturation levels are ∼13, 13, 92 and 48 mJy at 3.6, 4.5, 5.8 and 8.0 μm, respectively.

- The MIPS (Multi–band Imaging Photometry for Spitzer (Rieke et al. 2004)), which allows to perform broad band photometry, centered nominally at 24, 70, and 160 microns. Due to the low spatial resolution of the latter two channels we only made use of the data obtained with the $24\mu m$ band. Consequently, we limit the description of the MIPS camera to this detector only. The 24 μm channel consist of an array of 128×128 pixels, constructed using Si:As technology. The pixel size is $2.55''$, resulting in a field of view of $5.4' \times 5.4'$. A 5σ detection for a 500 second exposure time will detect a source at 110 μJy.

The final system's throughput (in number of electrons per incoming photon) for the four IRAC and the first of the MIPS detectors is presented in Fig. 4.17.

XMM–LSS is one of the SWIRE's observed fields and, thanks to the data obtained, it has been possible to extend the multi–wavelength study of the quasar candidates up to the mid–infrared wavelengths (namely up to 24 μm). The SWIRE coverage of the LCO area has been pretty good, with field losses of the order of only 10%. Fig 4.18 shows the exact field coverage, with the quasar candidates marked as squares. As it can be seen in the graph, there are only three objects falling in the region not covered by the SWIRE observations. The SWIRE catalog covering the LCO area contains ∼20,000 sources, all of which have emission in the 3.6 and 4.5 μm bands. For longer wavelengths, however, the number decreases dramatically, with only ∼ 15% of the sources having emission in at

[3]Spitzer is a NASA space mission, launched in 2003, with aim to study the Universe at IR wavelengths.

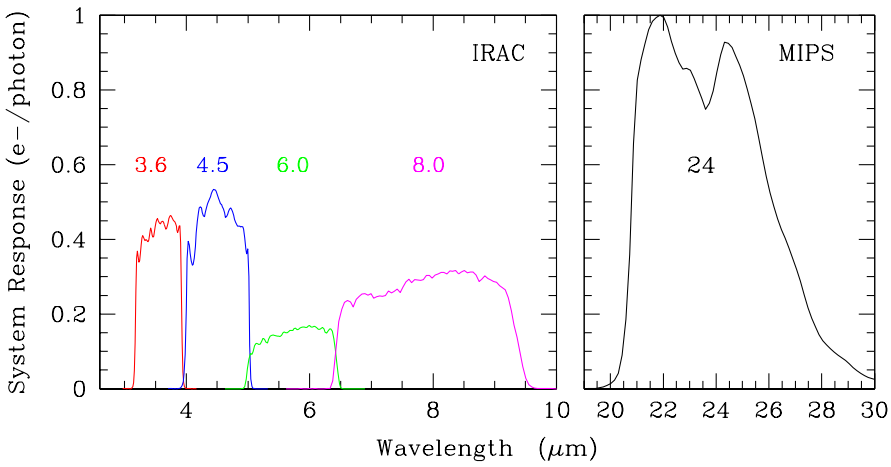

Figure 4.17: Final transmission curves for the Spitzer IRAC 3.6, 4.5, 5.8, 8 μm and MIPS 24 μm channels.

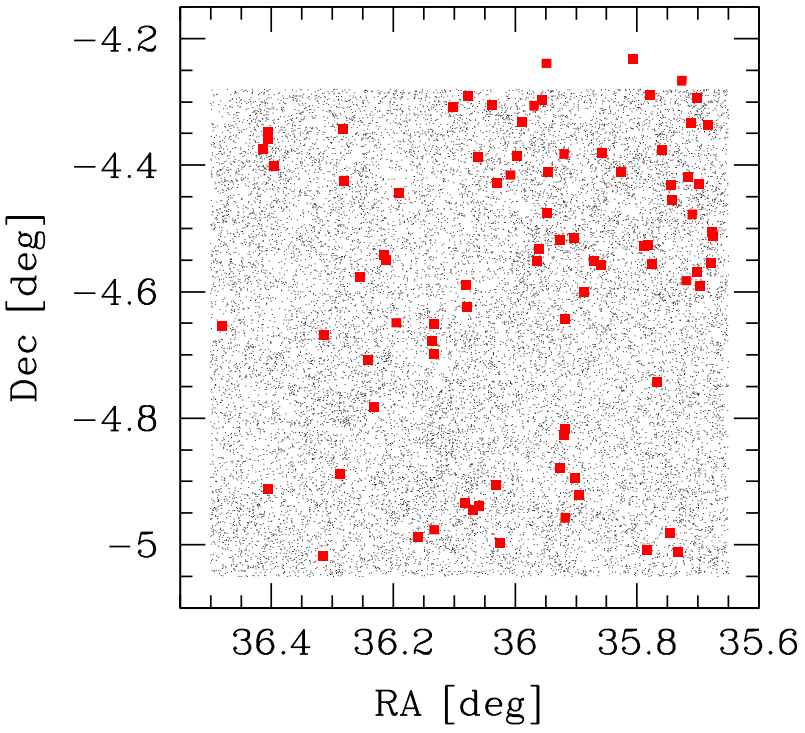

Figure 4.18: Field coverage of the LCO area by the SWIRE survey. The squares indicate the positions of the quasar candidates.

least one more of the 5.8, 8.0 or 24 μm channels. The number of sources having emission in all five channels reduces to only 748 sources (∼4%). The SWIRE data have been

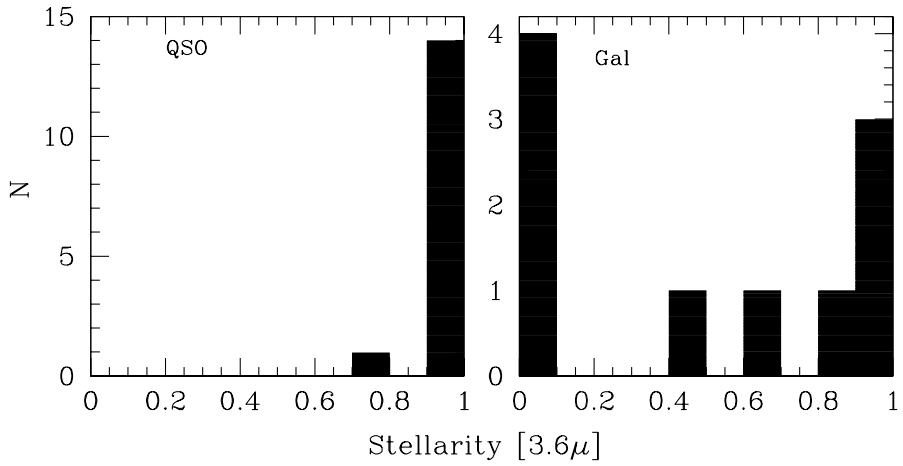

Figure 4.19: Stellarity index histograms, for the IRAC 3.6μm band, for the spectroscopically identified quasars (left panel) and galaxies (right panel).

processed using SExtractor (Bertin & Arnouts 1996) and therefore the catalog contains various types of photometry (see Annex A). The detection limit for the 3″ photometry (the one we will use for our study) in the 3.6, 4.5, 5.8, 8.0 and 24 μm energy bands is 6, 4.7, 20.8, 16.4 and 298 μJy, respectively.

4.3.2 The near and mid–infrared colors of quasars, stars and galaxies

Given the resolution of the Spitzer channels, the $Rz'K_s$ and SWIRE catalogs were matched to a tolerance of 1.5″. Thanks to the excellent coverage of the LCO area by Spitzer observations, it was possible to find a single infrared counterpart for 90% of the LCO sources. Regarding the QSO candidates, for all 82 objects with SWIRE coverage a single counterpart was found as well.

A first attempt to identify the quasar candidates among the sources, on the basis of SWIRE colors and morphological criteria, gave no results. The reason for this is clearly demonstrated in Fig. 4.19, which presents the stellarity histogram for the 3.6 μm band, for the spectroscopically identified galaxies and quasars. Galaxies, like quasars, can equally have a high stellarity index, make it impossible to separate the point–like from the extended sources lying in the same locus of the color–color diagram.

A thorough study of the near and mid–infrared properties of active galactic nuclei, using Spitzer photometry, has been performed by Eisenhardt et al. (2004), Lacy et al. (2004), Hatziminaoglou et al. (2005) and Stern et al. (2005). They have studied the colors of type I and type II quasars, using various combinations among the IRAC and MIPS photometric bands. They have demonstrated that a robust selection criterion for

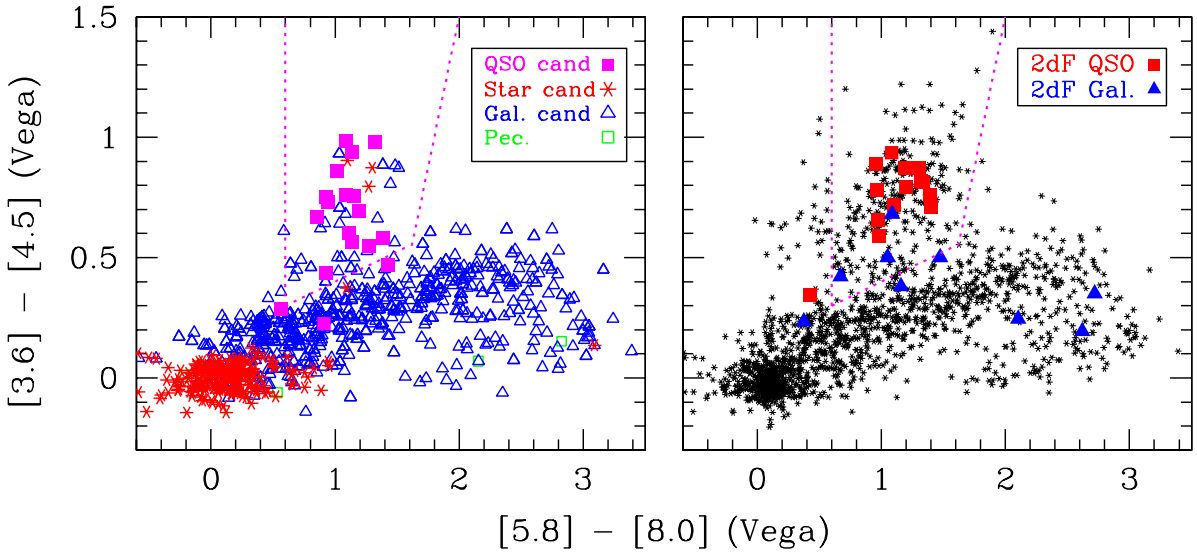

Figure 4.20: *Left panel:* Color–color diagram for the SWIRE counterparts of the $Rz'K_s$ matched sources. *Right panel:* Color–color diagram for the uncorrelated SWIRE catalog covering the LCO area. The spectroscopically identified quasars and galaxies are plotted as squares and triangles, respectively.

separating stars and galaxies from AGNs is by plotting the ([3.6] − [4.5]) color index as a function of the ([5.8] − [8.0]) one. For the different types of objects in the catalog, this combination or IRAC bands takes full advantage of features which are unique for each population, such as the different continuum emission and the redshifted emission lines. All 82 quasar candidates show emission in the 3.6 and 4.5 μm bands. Nevertheless, the number of objects with emission in one of the 5.8, 8.0 or 24 μm reduces to 25 (30%), 23 (28%) and 16 (20%), respectively. 18 objects (22%) show emission from 3.6 to 8.0 μm, while only 15 sources (18%) show emission in all five bands.

The color–color diagram for ∼2900 $Rz'K_s$ sources with a SWIRE counterpart is presented in the left panel of Fig. 4.20. The quasar candidates, together with the sources classified as galaxies, stars and the objects with peculiar colors are plotted as filled squares, open triangles, asterisks and open squares, respectively. The right panel shows the same color–color diagram for the uncorrelated SWIRE catalog covering the LCO area. The spectroscopically identified quasars and galaxies are plotted as squares and triangles, respectively. The area defined by the dotted line will be discussed soon after.

These two plots are of major importance for understanding the properties of the matched $Rz'K_s$ sources, since the different populations occupy distinct regions in the color–color plane, specifically:

- The majority of the point–like sources characterized as stars reside in the quite re-

stricted region around (0,0). This is explained by the fact that the 3.6 μm–8.0 μm wavelength region samples the Rayleigh–Jeans tail of the blackbody spectrum describing the stellar SEDs, where the stellar colors are very close to zero in the Vega system.

- The objects classified as galaxies occupy the area with $0 \leq [5.8] - [8.0] \leq 3$ and $0 \leq [3.6] - [4.5] \leq 0.5$. According to Eisenhardt et al. (2004) and Stern et al. (2005), starburst galaxies and galaxies lacking stellar formation are indeed found in this area, something also confirmed by the galaxies with 2dF spectroscopic identification. The high contrast regarding the vertical and horizontal spread of the QSO locus can be explained as follows:
  - The most important features in starburst galaxies, for rest–frame wavelengths spanning 1–10 μm, are the 1.6 μm bump, as a composite of the stellar SEDs, and three polycyclic aromatic hydrocarbon (PAH) emission lines at 3.3, 6.2 and 7.7 μm.
    - For low redshifts, between $0.2 \leq z \leq 0.5$, the first of the PAH features enters the 4.5 μm IRAC band, making galaxies look redder (explaining the positive $[3.6] - [4.5]$ color index). For this redshift range, however, the 6.2 and 7.7 μm PAH features are shifting through the IRAC 8.0 μm band. Since the 6.2 and 7.7 μm features are much stronger than the 3.3 μm one, they produce a much wider spread on the of $[5.8] - [8.0]$ color index.
    - For redshifts $z \geq 1$ the 1.6 μm feature shifts from the 3.6 μm IRAC band to the 4.5 one, producing a modest reddening on the $[3.6] - [4.5]$ color index, while at $z \approx 1.3$ the 3.3 μm PAH feature causes a similar effect to the $[5.8] - [8.0]$ index.
    - Finally, for redshifts $z \geq 2$, the 1.6 μm bump enters the IRAC 4.5 μm, band, also reddening the $[3.6] - [4.5]$ color.
  - For passive galaxies, with old stellar populations, the main feature is the 2.35 − 2.5 μm CO absorption due to red supergiants.
    - For redshifts in the range $0.36 \leq z \leq 0.8$, it shifts through the 3.6 μm IRAC band, making passive galaxies look bluer.
    - For redshifts in the range $0.8 \leq z \leq 1.3$, it enters the 4.5 μm band, producing redder $[3.6] - [4.5]$ colors.
- However, the most striking feature in Fig. 4.20, is the isolation of the quasar locus from the stars and galaxies. As it can be seen in the right panel, the 2dF quasars reside in a branch centered at about $[5.8] - [8.0] \sim 1$ in the horizontal axis, and extending in the vertical direction between $0.4 \leq [3.6] - [4.5] \leq 1.3$. This was first suggested by Eisenhardt et al. (2004) and was confirmed by Stern et al. (2005), on the basis of spectroscopic data from the AGES survey (Kochanek et al. in preparation). Regarding the shape of the area, the lack of strong PAH features reduces the dispersion of QSO colors in the horizontal direction. Additionally, since

for $\lambda \leq 5$ the QSO SED is described by a power–law, while stellar SEDs have a convex shape, quasars will have a redder $[3.6] - [4.5]$ color index with respect to stars (and consequently galaxies).

In order to isolate the AGN region in the color–color plane, Stern et al. (2005) used empirical criteria, by studying the IRAC colors of spectroscopically identified quasars. The final equation they adopted, which visually defines the area marked with the dotted line in Fig. 4.20, is:

$$\begin{aligned} ([5.8] - [8.0]) > 0.6 \ \wedge \quad & ([3.6] - [4.5]) > 0.2([5.8] - [8.0]) + 0.18 \\ \wedge \quad & ([3.6] - [4.5]) > 2.5([5.8] - [8.0]) - 3.50 \end{aligned} \tag{4.9}$$

where $\wedge$ is the logical AND operator. The left border (see Fig. 4.20) is meant to prevent contamination by high redshift galaxies, while the right boundary is defined by the [5.8]-[8.0] distribution of AGN colors. omitting a large number of AGNs Based on an analysis of ∼680 sources found in the AGN area, with spectroscopic identification, Stern et al. (2005) report that the expected percentage of broad–line AGNs, narrow–line galaxies and stars is of the order of 77%, 6%, 16% and 1%, respectively. However, as it is clearly shown in Fig. 1 of the Stern et al. (2005) paper, the fact that AGNs reside in the area defined by Eq. 4.9 does not exclude the possibility of finding AGNs elsewhere, both broad–line and narrow line.

Based on this analysis, we are now ready to study the distribution of the 82 QSO candidates having a SWIRE counterpart, in the IRAC color—color plane. The number of candidates with photometry in all four bands reduces to 18. 15 sources are found in the area defined by Eq. 4.9, while the remaining three are slightly away, thus are still considered as good QSO candidates. 18 good quasar candidates correspond to 22% of the initial sample, meaning that the contamination rate is ∼78%. This information provides a more robust evaluation concerning the KX–selection, with respect to the first evaluation performed with the CFHTLS data. Because of their point–like morphology, the contaminants have to be stars. As we see in the left panel of Fig. 4.20, a very small percentage of galaxies is also found in the AGN area, probably Seyfert type, or starbursts with ([3.6]-[4.5]) color redder than the average value. Finally, there are also three point–like objects found in the AGN area, classified as stars according to the KX–excess criteria. Should these objects be quasars, this would mean that the KX–excess failed in detecting only three cases, giving a success rate to the method $\sim 85\%$ (18/21).

Chapter 5

Spectral Energy Distributions & Photometric Redshifts

In chapter 4 we studied the multi–wavelength characteristics of the quasar candidates. Thanks to the CFHTLS and SWIRE data, it has been possible to derive a rough estimation on the number of real quasars and contaminants. The purpose of the current chapter is to perform a complete characterization of the sample. Due to the lack of spectroscopic data, this goal will be achieved by studying the spectral energy distribution (SED) of the quasar candidates, that is, study their flux as a function of wavelength. The differences in the SED of the objects in the sample will allow to identify the real quasars among the candidates and therefore obtain a precise idea on the contamination in the KX–selected sample. Even more, through the study of their SED, we hope to get an insight on the energy production mechanisms in the KX–selected quasars. Additionally, using the available multi–wavelength photometry, we will calculate the quasars' redshift.

Stellar SEDs have a convex shape and, as it was shown in paragraph 4.3.2, their mid–IR emission is too faint to be detected by the Spitzer satellite, resulting to lack of data at wavelengths long–wards 4.5 μm. Additionally, because of the infrared selection criteria, we expect the contaminants to be cool stars, whose emission will peak at fixed IR wavelengths. On the other hand, the physical processes taking place in quasars generate SEDs with different properties. The energy emitted from the central source is absorbed and re–processed by the dusty torus surrounding the active nucleus. Thus, we expect to detect mid–IR emission up to 24 μm.

"Classical" Type–I quasars, where the complete AGN structure (i.e. central source, accretion disk and torus) is seen face–on, are characterized by blue optical continua, i.e. their emission will decrease at longer wavelengths. After a minimum at $\sim$1 μm, which corresponds to the sublimation temperature of the graphite (1700 K) or silicate (1350 K), the torus contribution will be identified as a bump, whose peak will depend on the quasar's redshift.

Intrinsically red and "reddened" objects will both have much flatter SEDs. For the former case, the black hole–accretion disk coupling fails to produce the amounts of energy we see in the UV–optical part of Type–I quasar SEDs. For the latter case, either due to the viewing angle from which we observe the QSO (the torus partially hides the black hole), or due to the amounts of dust (in the host galaxy) that scatter, absorb and re–emit the incoming radiation, the observed SEDs will look much smoother, free from the characteristic features we find in Type–I quasars.

In the extreme case where a quasar is seen edge on (Type 2), the central source and the accretion disk are practically hidden behind the torus. Hence, the contribution from these two components in the UV–optical part of the SED will be considerably reduced, while, on the other hand, we expect to observe strong emission when going to mid–IR wavelengths, due to the dust in the torus.

Based on the above, a visual inspection of the SEDs of the 85 QSO candidates allowed to unambiguously identify 17 real quasars. The rest of the candidates are stars with $Rz'K_s$ colors similar to those of quasars, or galaxies respecting the morphological criteria applied, contaminating our sample.

Because of the lack of spectroscopic data, we calculated the redshift of the real quasars using the "photo–z" technique, a method that gives a redshift estimation on the basis of broad–band photometry. For each quasar, a theoretical template, characterizing its spectral energy distribution, is properly shifted (i.e. dilated) in wavelength, corrected for any possible extinction effects and calibrated in flux, in order to match the observational photometric points available for the specific object.

The proper interpretation of the photometric redshifts encompasses some subtleties. For this reason, the construction of the observational SEDs and the comparison with HyperZ's output will first be tested on the quasars with 2dF identification. This will allow to experiment with the methods, having an a priori knowledge on what we expect to obtain. Once we have explored these tools, we will proceed with the selection of the real QSOs on the basis of their spectral energy distribution and measure their photometric redshift. These steps are essential in order to investigate on the properties of the KX–selected quasars, identify the contaminants and evaluate the K–excess method.

5.1 Photometric redshifts

5.1.1 Introduction

The idea behind the photometric redshift technique is to estimate the redshift of galaxies and AGN on the basis of available multi–wavelength photometry. The necessity for an alternative approach for measuring redshifts, other than spectroscopy, has become more than imperative, especially during the last years. The participation of large–aperture

telescopes in multi–λ surveys has generated huge amounts of high–quality photometric data. Because of the areas covered, however, it is impossible to obtain spectra for all objects present in the photometric catalogs. Moreover, the limiting magnitudes of the imaging surveys are by far deeper than the faintest objects that spectroscopic observations can probe, making impossible the redshift measurement of the fainter objects via spectroscopy.

Among the first who explored the photometric redshift technique were Bolzonella et al. (2000), with their code called HyperZ. The main concept describing HyperZ is relatively simple: using a number of spectral templates, which describe the SED of the objects we are interested in, the method translates each of the templates at various redshifts, convolving them with the transmission curves of the filters for which we have photometric data. By comparing the simulated magnitudes to the observed ones, the method finally decides on the best–fit template for each object, on the basis of χ^2 statistics. The main output of the method is the redshift at which every source is found, accompanied by a probability value, expressing the credibility of the derived value (which is essential, since HyperZ will always give an output value), and the name of the best–fit template.

The photometric redshift technique has a disadvantage, which, depending on whether we are dealing with galaxies or AGN, can be minor or major, respectively. This method–related drawback, concerns the precision of the estimated redshift, which can never be as good as the one measured on the basis of spectra. The reason for this is that we are not simply measuring direct information originating from the source, but rather estimate what we would expect to measure, in terms of photometry, by a template that is supposed to (but does not really) perfectly describe the object's SED. Thus, the weakest point of the method is the necessity for realistic templates, that represent as accurately as possible the object's SED. Other factors, such as the combination of photometric data possibly coming from different telescopes, and variability because of data obtained at different epochs, also influence the credibility of the induced results.

The photo–z approach works quite well for the case of galaxies, a success explained by two reasons. First of all, galaxies have predictable SEDs, since their energy production mechanisms are based on well–understood processes related to stellar evolution. Additionally, the plenitude of galaxies found at low redshifts, has played a major role in this study, since excellent quality photometric data and high signal–to–noise spectra have made it possible to probe all features in their spectral energy distribution. As a result, the available galaxy templates describe with a very credible way their energy–production mechanisms.

Contrary to galaxies, the results obtained with the photo–z technique for the case of QSOs are not always consistent with the value measured on the basis of spectra, with a difference between the two methods that can be as large as one, even two, orders of magnitude. The principal reason for this inconsistency is the complex architecture of the active nucleus, which is subject to projection effects altering our view on the intrinsic processes taking place in the heart of the AGN (e.g. Type–I and Type–II quasars, BL–Lac

objects, etc). A second reason is the strong light absorption by the torus surrounding the central engine, reprocessed and re–emitted at longer wavelengths, a mechanism which is still poorly understood. Finally, the high redshifts at which quasars are found, unfolds, in its turn two complexities: the apparent faint magnitudes of distant quasars make them difficult to detect and our knowledge on their intrinsic processes is still poor. As a result, photometric redshifts for high–z QSOs are based on templates generated using low–redshift members, which do not incorporate possible evolutionary effects relating the black hole, the active nucleus and the host galaxy. Secondly, due to their very small angular size, it is very difficult to disentangle the contribution of the host galaxy from the emission coming from the quasar itself. For all these reasons, the quasar templates do not globally represent the energy distribution of quasars, and thus the method succeeds or fails in calculating the real redshift depending on the case.

5.1.2 The HyperZ code

The main input to the program is a parameter file which contains information regarding various parameters, such as:

- The redshift range in which the program must search for solutions, the step used for the minimization process and whether the step is constant or variable.
- The type of magnitudes in the photometric catalog (Vega or AB).
- The type of reddening law (due to dust extinction inside the galaxies themselves) and the values for the minimum/maximum absorption in the V–band, A_V.
- The basic parameters describing the cosmological model, namely the matter density, the cosmological constant and the Hubble constant.

Apart from the numerical parameters, the parameter file contains information regarding:

- The collection of templates describing the spectral energy distribution of the objects.
- The transmission curves of the filters for which we have photometric data.
- A catalog containing the photometry and the corresponding magnitude errors for the objects for which we want to measure the redshift. By convention, objects not detected in a given band, due to their faint magnitudes, have to be flagged with a value equal to 99, while objects with no photometry available ("out of field objects") must be flagged with by setting their magnitudes equal to one.

A typical HyperZ parameter file, containing the actual values we used during the performed tests, looks as follows:

```
                    #Parameter file

FILTERS_RES         FILTERS.ugriz.dat    # filters' transmission
FILTERS_FILE        FILTERS.ugriz.param  # filters' file
TEMPLATES_FILE      Templates.list       # models file
FILT_M_ABS          4                    # filter for absolute magnitude
ERR_MAG_MIN         0.1                  # err_min
Z_MIN               0.00                 # minimum redshift
Z_MAX               3.0                  # maximum redshift
Z_STEP              0.01                 # step in redshift
CATALOG_FILE        Photom.ugriz.cat     # cat file
MAG_TYPE            1                    # 0 standard Vega mag, 1 AB mag
REDDENING_LAW       5                    # reddening law
AV_MIN              0.0                  # Av_min
AV_MAX              1.2                  # Av_max
AV_STEP             0.2                  # Av_err
```

Given the depth of the matched $Rz'K_s$ detections, and the highest spectroscopic redshift for the objects in our catalog, we set the upper redshift limit to $z = 3$. We also chose to use a constant redshift step, with $dz = 0.01$. The magnitudes in the photometric catalogs were properly converted to the AB system. We used the Calzetti et al. (2000) reddening law, with A_V spanning between 0 and 1.2. We also used a standard cosmological model, with $H_0 = 70\,\mathrm{km/s/Mpc}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$ (although none from HyperZ's output we used depended on the cosmological model). Detailed information on HyperZ and the exact definition of its parameters can be found in the manual pages of the program and the related publication (Bolzonella et al. 2000).

Concerning the photometric catalogs, we have, as previously described, three photometric data sets:

- The five–band CFHTLS data ($u^*g'r'i'z'$),
- The $Rz'K_s$ photometry
- The SWIRE data (3.6, 4.5, 5.8, 8.0 and 24 μm)

The basic characteristics of the 13–band photometric data are briefly summarized in Table 5.1. Columns 2 and 3 present the central wavelength for each filter, in units μm, and its identification. Column 4 refers to the type of the photometric data (photometry in AB or Vega magnitudes, or fluxes in units μJy). Column 5 refers to the survey used for obtaining the various data–sets.

Regarding the templates, we used four local galaxy templates, available in HyperZ. These templates were actually taken from Coleman et al. (1980) and correspond to one

No	$\lambda(\mu m)$	Filter	Phot. Type	Legacy Program
(1)	(2)	(3)	(4)	(5)
1	0.374	u^*	AB	CFHTLS
2	0.487	g'	AB	CFHTLS
3	0.625	r'	AB	CFHTLS
4	0.700	R	Vega	CTIO
5	0.770	i'	AB	CFHTLS
6	0.910	z'	AB	CFHTLS
7	0.940	z'	AB	CTIO
8	2.200	K_s	Vega	LCO
9	3.600	IRAC–1	μJy	SWIRE
10	4.500	IRAC–2	μJy	SWIRE
11	5.800	IRAC–3	μJy	SWIRE
12	8.000	IRAC–4	μJy	SWIRE
13	24.000	MIPS–1	μJy	SWIRE

Table 5.1: Basic characteristics of the photometric data–sets used for measuring the photometric redshifts.

elliptical, one irregular and two spiral galaxies. Apart from these templates, we also used a library of SEDs provided to us by Dr. M. Polletta (Univ. of California, CASS, personal communication). An extensive use of these templates is made in Polletta et al. (2006), but the publication with an analytical description of the templates is in preparation. The list of all the above templates is presented in Table 5.2. The first and second columns provide the identification number and the name of the template, respectively. The name informs whether the given template has been generated on the basis of a simulation code (e.g. templates 1–3), or it corresponds to the SED of a given object, which is representative for a whole class (e.g. template No 17, for Markarian 231). Column 3 gives some basic information concerning the nature of the object (elliptical galaxy, spiral, etc). Column 4 provides a basic reference for each template.

A selection of the templates that were most frequently used by HyperZ are shown in Fig. 5.1. The horizontal axis represents rest–frame wavelengths, while the vertical axis fluxes, normalized to the monochromatic flux F_λ at $\lambda = 5500\,\text{Å}$. Both axes are in logarithmic scale.

All the templates presented in Table 5.2 extend from the UV to the infrared. Some additional information regarding the templates can be found below:

- Templates 01–03 have been generated using the Grasil code (Silva et al. 1998) and correspond to an elliptical galaxy of 2, 5 and 13 Gyr of age.

- Templates 04–09 correspond to spiral galaxies of different sub–classes, with the

No	Template	Spectral Type	Reference
(1)	(2)	(3)	(4)
01	Ell(2)	Ell (t= 2 Gyr)	Silva et al. (1998)
02	Ell(5)	Ell (t= 5 Gyr)	Silva et al. (1998)
03	Ell(13)	Ell (t=13 Gyr)	Silva et al. (1998)
04	S0	Spiral	Silva et al. (1998)
05	Sa	Spiral	Silva et al. (1998)
06	Sb	Spiral	Silva et al. (1998)
07	Sc	Spiral	Silva et al. (1998)
08	Sd	Spiral	Silva et al. (1998)
09	Sdm	Spiral	Silva et al. (1998)
10	NGC 6090	Starburst	Silva et al. (1998)
11	M 82	Starburst	Silva et al. (1998)
12	Arp 220	ULIRG[1]	Silva et al. (1998)
13	Seyf 2	Seyfert 2	Polletta et al., in prep.
14	Seyf 1.8	Seyfert 1.8	Polletta et al., in prep.
15	NGC 6240	Starburst + Seyfert 2	Polletta et al., in prep.
16	IRAS19254–7245 South	Seyfert 2 + ULIRG	Berta et al. (2003)
17	Mrk 231	Type–1 AGN + ULIRG	Berta et al. (2003)
18	QSO1–median	Type–1 QSO	Brotherton et al. (2001)
19	QSO1–low	Type–1 QSO	Brotherton et al. (2001)
20	QSO1–high	Type–1 QSO	Brotherton et al. (2001)
21	Torus	Type–2 QSO	Polletta et al., in prep.
22	QSO2	Type–2 QSO	Polletta et al., in prep.
23	CW–Im	Irregular	Coleman et al. (1980)
24	CCW–E	Elliptical	Coleman et al. (1980)
25	CCW–Sbc	Spiral	Coleman et al. (1980)
26	CCW–Scd	Spiral	Coleman et al. (1980)

Table 5.2: List of templates used as input to the HyperZ code.

corresponding PAH features[2] modified on the basis of IRS spectra[3].

- For templates 10–12, the PAH features were modified using ISO PHT–S spectra[4].
- Templates 18–20 have been constructed on the basis of an average Type–I QSO template taken from the Large Bright Quasar Survey (LBQS–Brotherton et al. 2001). Low–redshift QSO spectra, obtained with IRS on–board Spitzer, were scaled to the same level and combined. Sources with optical SEDs similar to the LBQS spectrum were kept. Since the dispersion in the IR was pretty large, the quasars were separated in three categories, one using the median flux of all sources, one using only the quasars with the highest infrared–to–optical (IR/Opt) ratios and one using the quasars with the lowest IR/Opt ratios. By fitting the observational

[2]Polycyclic Aromatic Hydrocarbon; features found between $\sim 6 - 15\,\mu$m.

[3]IRS is the InfraRed Spectograph on–board the Spitzer satellite.

[4]PHT–S refers to the spectroscopy mode performed with the PHT instrument on–board the ISO satellite.

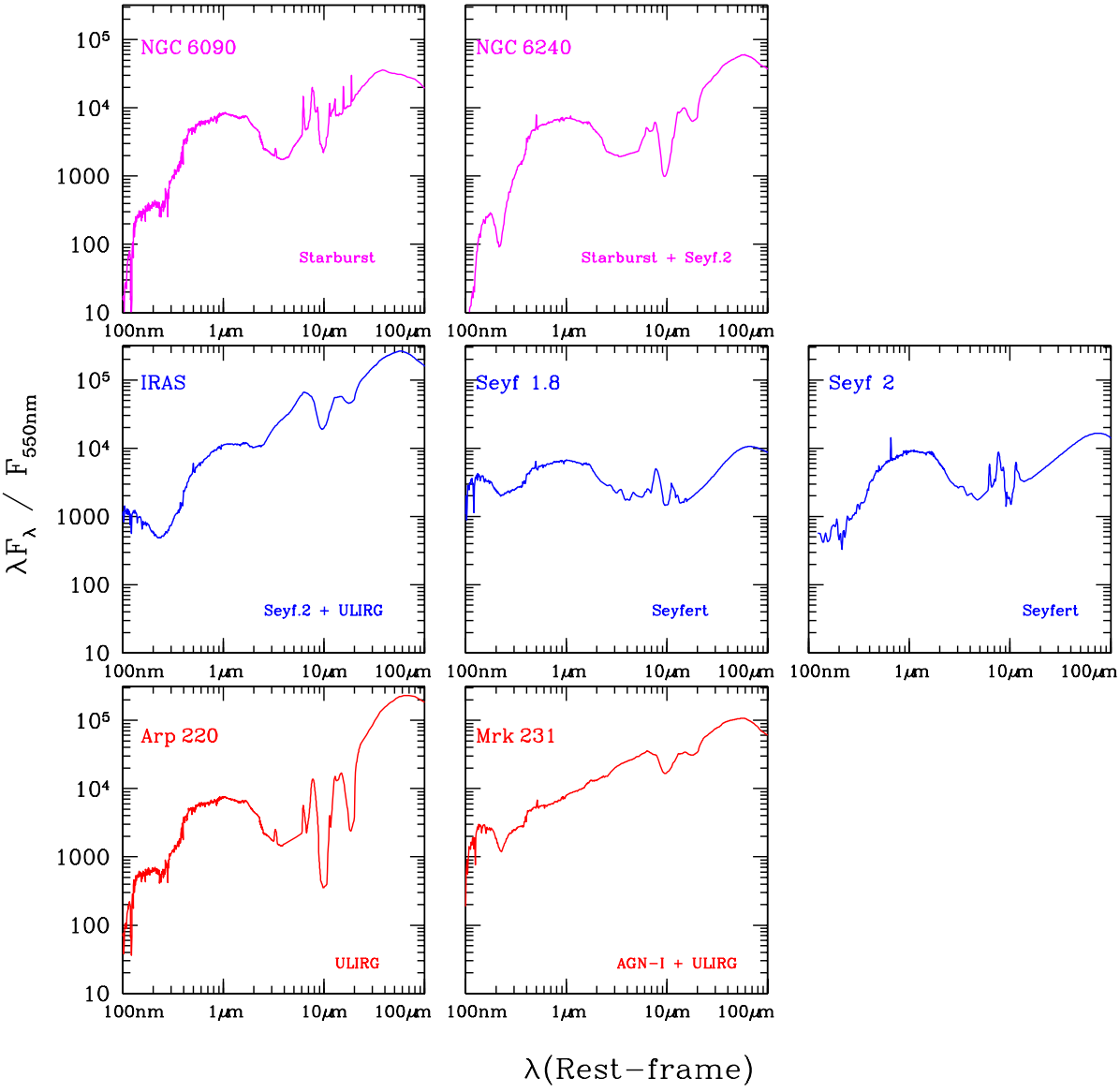

Figure 5.1: Template SEDs for different types of objects, used by HyperZ for calculating the photometric redshifts.

points for each of these categories, three average mid–to–far infrared QSO templates were generated. The average SED for each of these three categories was then put together with the LBQS spectrum, for producing a quasar template spanning from the ultra–violet to the far–infrared. Nevertheless, in order to reconcile the IR SED with the optical one, the last part of the optical spectrum was removed, at around ∼5500 Å, for the SEDs corresponding to the low IR/Opt and the median. As a result, some emission lines (among which H_α) are absent from these two templates.

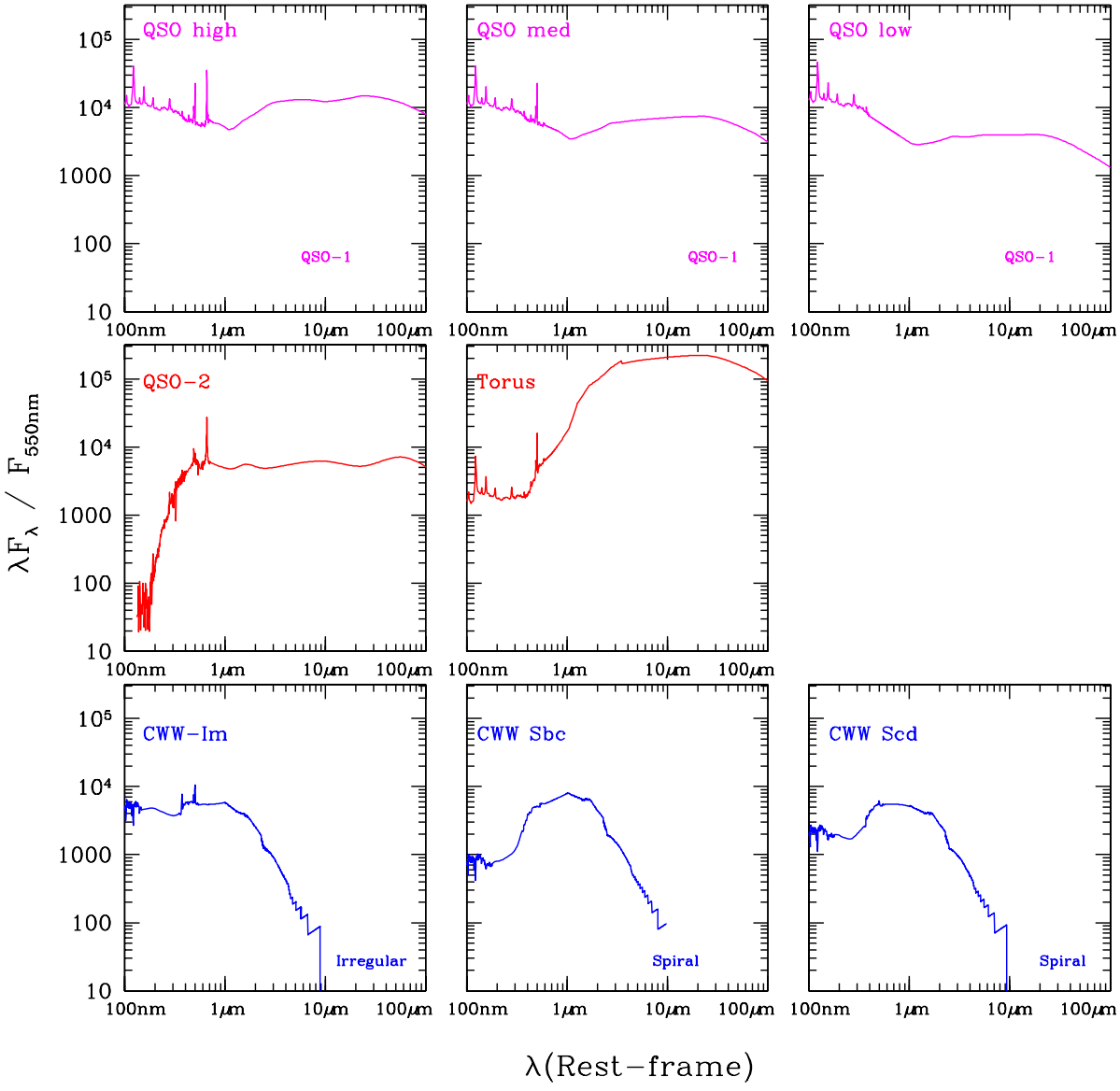

Figure 5.1: Continued

5.1.3 Photo–z versus spectroscopic redshift for the 2dF QSOs

Due to the numerous combinations available within HyperZ (e.g. filters, templates, type of extinction, etc.), we considered that it would be extremely useful to test the program on objects with known spectroscopic redshift, before trying to assess a redshift value for the quasar candidates. For this reason, we performed various tests on the spectroscopically identified (2dF) quasars present in the Las Campanas field (see Table 3.2), in order to find the combination of parameters that would give a photo–z value as close as possible to the spectroscopic redshift. Because of the incomplete coverage of the LCO area by

the CFHTLS observations, we decided to use, for the first time, the information available for all 14 spectroscopically identified QSOs and not only for the nine objects that were directly matched to the $Rz'K_s$ catalog (see section 3.2).

During the trials we studied the solutions obtained when using all available photometric data ($u^*g'r'i'z' + R\,z'_{ctio}\,K_s$ + 4 IRAC (3.6–8.0 μm) + 1 MIPS (24 μm) bands) as well as the solutions obtained when using combinations among the 13 photometric bands. Additionally, we tested various combinations of templates, since HyperZ limits the number of simultaneously used SEDs to 15 (see Fig. 5.1).

Before running the HyperZ code, all non–AB magnitudes, namely the R, K_s–band Vega magnitudes, were converted to the AB photometric system according to the following equations:

$$R\,(AB) = R\,(Vega) + 0.22 \quad (5.1)$$

$$K_s\,(AB) = K_s\,(Vega) + 1.89 \quad (5.2)$$

Similarly, the SWIRE fluxes, measured in μJy, were properly converted to AB magnitudes using the definition of the AB system:

$$m_{AB} = -2.5\,\log \int F(\nu)\,d\lambda - 48.60 \quad (5.3)$$

where the monochromatic flux $F(\nu)$ represents the flux in each of the IRAC or 24 μm MIPS bands, properly converted in units of erg s^{-1} cm^{-2} Hz^{-1}.

Once the photometric data were properly converted to the AB system, HyperZ was executed using different combinations of filters and, or templates. The results were carefully examined, in order to detect the combination that would give the best results. Apart from the primary solution, HyperZ also proposes alternative solutions, with a higher χ^2 and a lower probability. These solutions were equally studied, in order to better understand HyperZ's behavior.

Depending on the type of photometric data introduced as an input, HyperZ used different template "families" for finding the best photo–z, that is templates belonging to the same type of objects (e.g. starbursts, elliptical, etc). More specifically, the use of the CFHTLS data exclusively, oriented HyperZ to find the best solution among templates showing predominant features only in the optical part of their spectrum, rather than in the infrared. Similarly, using the infrared photometry inevitably "forced" the program to use different type of families, that could reconcile the emission both in the optical and in the infrared part of the object's SED.

The tests we performed demonstrated that the photometric redshift results were quite sensitive both to the combination of filters and templates. For some objects the calculated

redshift could be equally close and far away from the spectroscopic value, depending on the above combination. Using all available photometric data did not necessarily improve the photo–z, since differences in the calibration between the various data sets and the fact that the data were obtained at different epochs could potentially influence the derived values. Additionally, the photometry for the last two IRAC (5.8, 8.0 μm) and the MIPS (24 μm) bands rather complicated, than improved, the χ^2 minimization. The explanation is that the detection in these bands corresponds to processes in rest–frame wavelengths that are relatively poorly studied and as a consequence the average templates do not successfully describe the SED of all quasars. For all these reasons, we finally decided to use, when available, the CFHTLS photometry and the first two of the IRAC bands, since using these data–sets we cover wavelengths spanning from the optical to the infrared, with only two instruments and epochs of observations. Due to the lack of CFHTLS photometry, the two SWIRE bands were used together with the $Rz'K_s$ photometry.

Table 5.3 presents the photo–z results for nine of the spectroscopically identified QSOs, when using the CFHTLS photometry exclusively. Column 1 gives the spectral identification of the objects, as presented in Table 3.2. Columns 2, 3 and 4 give the spectroscopic redshift, the HyperZ value and their difference, respectively. Column 5 and 6 give the χ^2 and a percentage expressing the reliability of the calculated value. Columns 7 and 8 present the best–fit template and the corresponding A_V value. Column 9 is a flag, indicating whether the photo–z value closest to the spectroscopic redshift is the primary solution proposed by HyperZ (flag "P"), or it was found among the secondary solutions ("S").

From the values presented in column 4 of Table 5.3 we conclude that the five–band CFHTLS photometry provides quite satisfactory results. The average value on the absolute relative differences between the spectroscopic and photometric redshifts is $\Delta(z)/z = 0.096 \pm 0.040$. In order to estimate the best redshift value HyperZ used five templates, taken from three "families" (quasars, Seyferts and normal galaxies). Finally, for seven values out of nine the primary solution was the one closest to the spectroscopic redshift.

Table 5.4 presents the results for the same quasar sample, for the case that the photometric redshift has been computed using the CFHTLS and the two–band SWIRE photometry. The column identification is the same as in Table 5.3. The mean redshift difference is $\Delta z/z = 0.069 \pm 0.083$. By comparing the two tables it is clear that using the SWIRE photometry improved the photo–z value. Additionally, the selection of templates has considerably changed, since HyperZ correctly identified the objects as quasars.

Comparison between the photometric and spectroscopic redshifts presented in Tables 5.3, 5.4 is graphically illustrated in Fig. 5.2. The vertical and horizontal axis correspond to the photo and spectro–z, respectively. The left panel shows the results obtained using the CFHTLS photometry exclusively, while the right panel presents the results obtained using simultaneously the CFHTLS and the two IRAC–band photometry.

Table 5.5 presents the results for five 2dF QSOS for which no CFHTLS data are

Sp. ID	Spectro–z	Photo–z	$\Delta(z)$	χ^2	Prob(%)	Template	A_V	Sol.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
11	1.183	1.031	0.152	0.122	97.45	QSO–high	0.4	P
14	2.263	2.503	−0.240	0.178	94.99	I 19254s	0.2	P
41	2.483	2.274	0.209	0.103	98.14	Seyfert 1.8	0.0	P
97	0.826	0.784	0.042	0.097	98.34	QSO–high	0.0	P
271	2.487	2.390	0.097	0.379	82.36	Seyfert 1.8	0.0	P
281	2.164	2.377	−0.213	0.046	99.60	CWW–Scd	0.0	P
375	1.937	1.748	0.189	0.454	76.95	Seyfert 1.8	0.0	P
530	1.566	1.798	−0.232	0.144	93.38	QSO–high	0.0	S
736	1.629	1.538	0.091	0.359	78.27	CWW–Im	0.0	S

Table 5.3: Comparison between spectroscopic and photometric redshifts for nine spectroscopically identified quasars. The photo–z has been computed on the basis of CFHTLS photometric data exclusively.

Sp. ID	Spectro–z	Photo–z	$\Delta(z)$	χ^2	Prob(%)	Template	A_V	Sol.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
11	1.183	1.171	0.012	1.107	35.50	QSO–high	0.4	S
14	2.263	2.401	−0.138	0.363	90.26	QSO–low	0.0	P
41	2.483	2.448	0.035	0.712	63.96	QSO–low	0.0	P
97	0.826	0.874	−0.048	1.120	34.77	QSO–med	0.2	S
271	2.487	2.591	−0.104	0.833	69.80	QSO–med	0.0	P
281	2.164	2.394	−0.230	0.306	93.42	QSO–med	0.2	P
375	1.937	1.903	0.034	0.706	64.47	QSO–med	0.2	P
530	1.566	1.513	0.053	0.191	96.63	QSO–med	0.0	P
736	1.629	1.798	−0.169	0.339	88.92	QSO–low	0.4	S

Table 5.4: Comparison of the photometric and spectroscopic redshifts, in the case the photo–z has been derived using CFHTLS + SWIRE photometry.

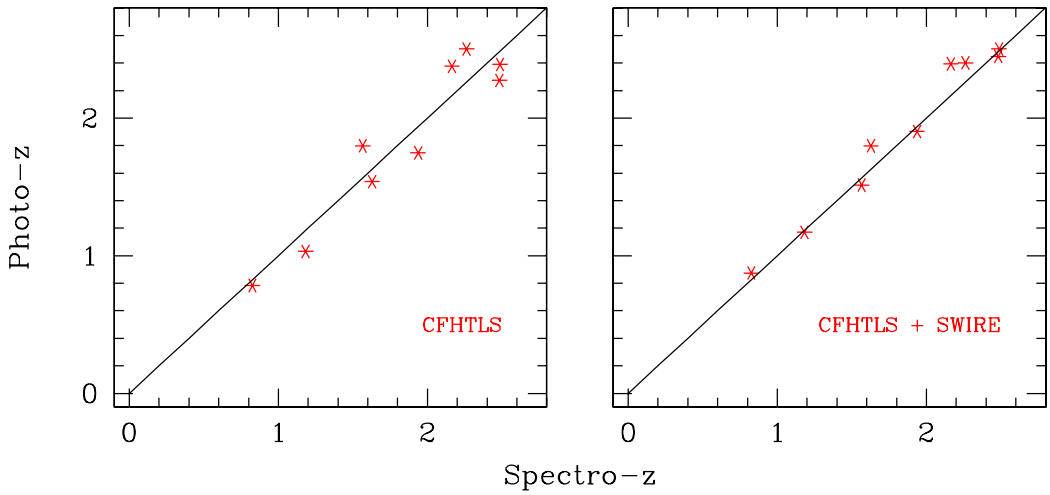

Figure 5.2: Photometric redshift as a function of spectro–z when using only CFHTLS photometry (left panel) or CFHTLS+SWIRE (right panel).

Sp. ID	Spectro–z	Photo–z	$\Delta(z)$	χ^2	Prob(%)	Template	A_V	Sol.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
23	1.924	0.719	1.205	5.984	00.01	QSO–med	0.0	P
251	1.497	1.064	0.433	0.665	61.62	QSO–med	0.0	P
91	2.275	2.483	−0.208	1.474	20.72	QSO–high	0.0	P
100	2.463	2.533	−0.070	0.776	54.07	QSO–med	0.0	P
735	0.760	0.864	−0.104	0.691	59.77	QSO–med	0.4	S

Table 5.5: Comparison of the photometric and spectroscopic redshifts, in the case where the photo–z has been derived using $Rz'K_s$ + SWIRE photometric data.

available. In this case the photo–z has been computed using the $Rz'K_s$ and the first two IRAC bands. Two cases have been isolated from the remaining three, since the photo–z value deviates significantly from the spectroscopic redshift and no secondary solution was found to approximate the real value.

O. Garcet (AGO, University of Liège) performed an independent study on the spectroscopic and photometric redshifts of an X–ray selected sample of AGN in the XMM–LSS field, with available CFHTLS–SWIRE photometry and 2dF spectra. His sample contained 631 point–like sources with more than 30 counts in the CD $(2-10)$ keV (hard) band, with a unique counterpart in both the optical and in the infrared. He defined two sub–samples on the basis of their hardness ratio, one with $HR < -0.2$ and the second with $HR > -0.2$. Concerning the unabsorbed sources ($HR < -0.2$), 98 entries have both CFHTLS and SWIRE photometry, with 17 having 2dF spectroscopic information. For

optimizing HyperZ's efficiency he followed the same strategy as we did, i.e. used various combinations among the vast variety of available templates (the same with those in our study), filters and HyperZ's parameters. Similarly to our conclusion, he also noticed a significant improvement when using the CFHTLS and the 3.6 and 4.5 μm bands, as 13 out of 17 sources had a $\Delta z \leq 0.2$ (efficiency $\sim 76\%$). Adding more IR filters ($> 4.5\,\mu$m) degraded HyperZ's efficiency. His sample of absorbed X—ray sources contained 29 objects with CFHTLS and SWIRE photometry, but only 3 with spectroscopic data, thus the success rate of that sample was poorly defined. Nevertheless, it's worth mentioning that he was able to tune the parameters in order to have 100% efficiency.

5.2 Spectral Energy Distributions

In order to get an insight to HyperZ's results, we decided to re–construct the spectral energy distributions of the 2dF QSOs, that is, represent the product λF_λ (flux, in units erg s^{-1} cm^{-2}) as a function of observed wavelength λ. This type of diagrams demonstrate how the quasar's emitted energy is re–distributed within a given spectral range. By over–plotting, for each object, the best–fit template chosen by HyperZ, appropriately redshifted to the observed–frame wavelengths, we can study to what extent it matches the observational points.

5.2.1 Calibration of the photometric points

Before generating the diagrams for the ten quasars with good photometric redshifts, presented in Table 5.4, we first calibrated their optical magnitudes to monochromatic fluxes (in units erg sec^{-1}cm^{-2}Hz^{-1}). Such a conversion was not necessary for the SWIRE data, since they were directly available as such, in μJy units.[5] This conversion was performed according to the formula:

$$F_\nu = F_\nu^{Vega}\, 10^{0.4(m_{Vega}-m)} \tag{5.4}$$

where m, m_{Vega} is the magnitude of the object we are interested in and that of Vega, in a specific energy band, respectively, and F_ν, F_ν^{Vega} their monochromatic fluxes at a specific frequency. By definition, Vega's magnitude is zero in any energy band (see Eq. 3.1), hence Eq. 5.4 is simplified to:

$$F_\nu = F_o\, 10^{-0.4m} \tag{5.5}$$

where F_o plays the role of the zero–point conversion factor, and is equal to Vega's monochromatic flux at the given frequency. Using Vega's spectral energy distribution,

[5] We remind the reader that 1 Jy is defined as 10^{-23}erg sec^{-1}cm^{-2}Hz^{-1}

as provided by HyperZ, we measured the conversion factor for the eight photometric bands we are interested in. The results are presented in Table 5.6.

Band	$\lambda[\mu m]$	F_o [Jy]
u^*	0.374	1878
g'	0.487	2938
r'	0.625	3162
R	0.700	2870
i'	0.770	2530
z'	0.910	2734
z'_{ctio}	0.940	2364
K_s	2.200	620

Table 5.6: Effective wavelengths and conversion factors.

To convert the CFHTLS and CTIO z'–band magnitudes, given in the AB photometric system, in monochromatic fluxes, it was first necessary to transform them to Vega magnitudes, since the conversion factors were computed using Vega's SED. This conversion was done according to the following equations[6]:

$$\begin{aligned} u^*(Vega) &= u^*(AB) - 0.346 \\ g'(Vega) &= g'(AB) + 0.092 \\ r'(Vega) &= r'(AB) - 0.171 \\ i'(Vega) &= i'(AB) - 0.401 \\ z'(Vega) &= z'(AB) - 0.554 \\ z'_{ctio}(Vega) &= z'_{ctio}(AB) - 0.549 \end{aligned}$$

Once this step was performed, each photometric point was converted to monochromatic flux, using Eq. 5.5 and the appropriate conversion factor taken from Table 5.6. By multiplying the monochromatic flux, for every photometric point, with the frequency corresponding to the effective wavelength of the specific energy band, the magnitudes were finally converted to fluxes in units $\mathrm{erg\,s^{-1}}$.

5.2.2 Calibration of the template SEDs

In order to compare every object's observational energy distribution with HyperZ's output, it is necessary to properly calibrate both the observational points and the best–fit

[6]For the CFHTS filters the conversion factors are available at the following URL: http://orca.phys.uvic.ca/ gwyn/cfhtls/photz/filters.html

template. The template's rest–frame wavelengths, λ, were shifted to the observer's frame, λ_o, according to the formula:

$$\lambda_o = \lambda\,(1+z) \tag{5.6}$$

To properly calibrate the template's intrinsic flux to the observed frame, we had to account for two effects:

- The dust extinction, produced inside the host galaxy and
- The absorption due to neutral hydrogen clouds, as the light emitted by distant quasars crosses the intergalactic medium (Lyman forest).

5.2.2.1 Reddening Law

HyperZ provides five different reddening laws. We decided to use the one derived by Calzetti et al. (2000), an empirical law which describes pretty well the reddening in starburst galaxies. The intrinsic flux, $f_{int}(\lambda)$, is associated with the observed flux $f_{obs}(\lambda)$, according to the formula:

$$f_{obs}(\lambda) = f_{int}(\lambda)\,10^{-0.4\,A_\lambda} \tag{5.7}$$

where A_λ is the extinction at wavelength λ, related to the color excess $E(B-V)$ and to the reddening curve $k(\lambda)$ by:

$$A_\lambda = k(\lambda)\,E(B-V) = \frac{k(\lambda)\,A_V}{R_V} \tag{5.8}$$

with $R_V = 4.05 \pm 0.80$ for Calzetti's law and A_V is derived from HyperZ's output. The reddening curve, $k(\lambda)$, is described as follows:

$$k(\lambda) = \begin{cases} 2.659\,(-2.156 + \frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3}) + R_V & 0.12\mu m \leq \lambda \leq 0.63\mu m \\ 2.659\,(-1.857 + \frac{1.040}{\lambda}) + R_V & 0.63\mu m \leq \lambda \leq 2.20\mu m \end{cases} \tag{5.9}$$

5.2.2.2 Intergalactic absorption – Lyman forest

Neutral hydrogen clouds, found along the line–of–sight between the observer and the quasar, produce a first drop in the object's flux at wavelengths blue–ward to the Lyα line (rest–frame wavelength $\lambda = 1216$Å) and a second one below Lyβ (rest–frame wavelength $\lambda = 1026$Å). Since these clouds are found at redshifts lower than that of the quasar, the redshifted absorption lines will produce a feature in the QSO spectrum, found blue–ward to the Lyα line, known as the Lyman forest.

Oke & Korycansky (1982) defined two depression factors, D_A and D_B, which describe the amount of absorption between Lyα and Lyβ, and the Lyβ and the Lyman limit (found at 912Å), respectively. The D_A factor is defined as:

$$D_A = 1 - \frac{f_{obs}(\lambda)}{f_{int}(\lambda)} \tag{5.10}$$

where $f_{obs}(\lambda)$ and $f_{int}(\lambda)$ are the observed and intrinsic fluxes per unit wavelength in the QSO rest–frame. D_B is defined in the same way as D_A. The values for the depression factors used by HyperZ were taken from Madau (1995).

5.2.3 The spectral energy distribution of the 2dF QSOs

The spectral energy distribution of the 14 quasars with 2dF identification, together with the (properly calibrated) best–fit template, chosen by HyperZ, are illustrated in Fig. 5.3. The product λF_λ is in units erg s^{-1} and the horizontal axis represents wavelengths in the observer's frame. The graphs were generated using HyperZ's output and the plots show the observational points used for the calculation of the photo–z *exclusively*, together with the appropriate template, redshifted to the observer's frame, corrected for extinction, intergalactic absorption and scaled in flux for matching the available photometry. Although the templates span a much wider wavelength range (see Fig. 5.1), the redshifted–calibrated templates produced by HyperZ are limited in the wavelength range of the photometric points used for estimating the photo–z.

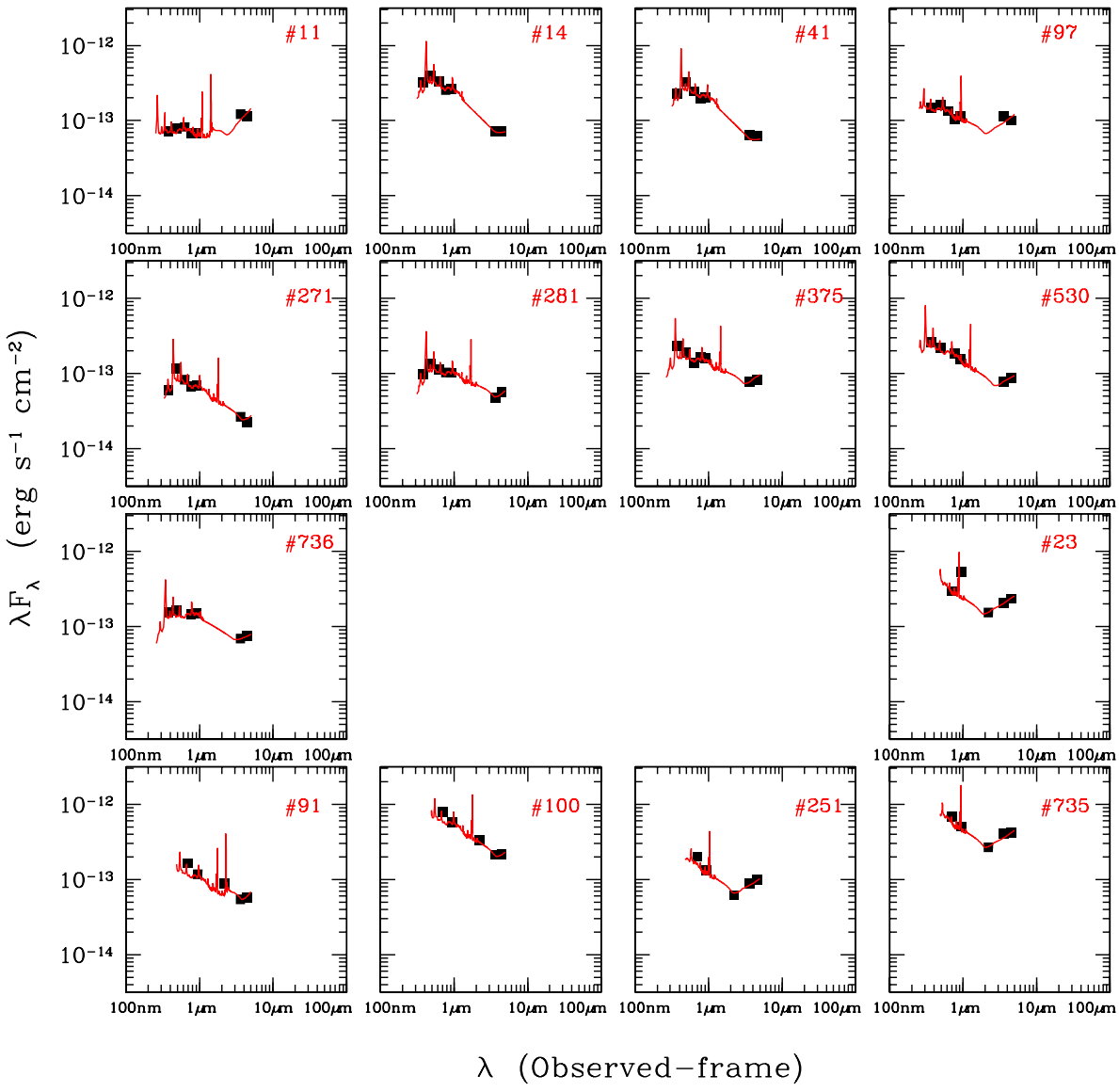

Figure 5.3: Spectral energy distribution for the 14 quasars with available spectroscopic redshift. Their 2dF identification is shown in the upper right part of each cell. For the first nine objects (11–736) the photo–z have been computed on the basis of $u^*g'r'i'z'$ and the first two SWIRE bands (3.6 μm, 4.5 μm). For the remaining five objects (23–735) the photometric redshifts have been computed using the $Rz'K_s$ measurements and the photometry from the first two SWIRE bands.

5.3 SEDs and photometric redshifts for the KX–selected quasars

5.3.1 Unveiling the real quasar population among the candidates

In order to identify the real QSOs among the KX–selected candidates we reconstructed the spectral energy distribution of the 85 sources, using their available photometry. The quasar identification was based on the fact that Type 1 objects have a much flatter SED (in λf_λ) when compared to stars, galaxies, or even their Type 2 counterparts. Nevertheless, Type-1 quasars are still characterized by a minimum in their UV–optical emission, found at about $1\,\mu$m at the rest–frame. The UV–optical slope corresponds to the accretion disk emission. The rise red–ward of the minimum corresponds to a black body emission of about $1300-1700$ K, the sublimation temperature of the silicate or graphite grains. Based on these characteristics features we managed to identify 17 candidates as real quasars. The rest of the objects were either stars with QSO–like infrared colors or galaxies with a morphological classification complying our selection criteria.

While generating the SEDs of the 85 candidates we noted that, for many of the non–QSO sources, the i' and z' CFHTLS photometric points were systematically shifted toward brighter fluxes (by a factor ~ 2), with respect to the global shape of the spectral energy distribution of the source. The study of the SEDs of the quasars with 2dF spectral identification had not revealed any similar effect, On the contrary, as we explained in section 5.1.3, the photometric redshifts closest to the spectroscopic values were obtained when using the CFHTLS data. Additionally, no calibrations problems were reported by the CFHTLS consortium. To minimize the influence of this effect on our analysis, we finally decided to discard the i' and z'–band CFHTLS photometric points when constructing the observational SEDs of the objects.

The observational SEDs of the 85 QSO candidates, constructed on the basis of optical (u^*, g', r), near–infrared $(Rz'K_s$) and mid–infrared (3.6–24 μm) photometry, are illustrated in Fig. 5.4. The quasars SEDs are presented as filled squares, while the photometry of the non-QSO candidates, with exception objects #1, #62, #68 and #77, is shown as open circles. The energy distribution of the four objects mentioned above shows some particularities and this is the reason for which they have been plotted as asterisks. A separate paragraph will be take dedicated to them, to discuss on their possible origin. The identifier of each candidate is shown in the upper right part of each cell.

A careful look at the energy distribution of the 17 KX–selected quasars leads to the conclusion that the sample contains three different types of objects:

- quasars showing strong emission both in the UV–optical part of the spectrum and in the infrared (e.g. #78). These objects are "classical" Type–I QSOs, i.e. face–on

quasars, whose blue optical colors are attributed to light coming directly from the accretion disk.

- quasars with a relatively flat emission in all wavelengths. These are intrinsically red, or "reddened" objects (e.g. #17), where the contrast between the optical and infrared emission is smoothed due to the presence of dust in the host galaxy, or even in the intergalactic medium.

- quasars with significantly low emission toward the blue part of the spectrum, which has a steep increase when going to redder wavelengths (e.g. #33). This is the case for type 2 quasars, which are seen edge–on. The main AGN components (i.e. the black hole and the accretion disk) are hidden behind the torus, which absorbs a large fraction of the incoming light and re–emits it in the IR.

Regarding the non–quasar SEDs, both galaxy and stellar profiles have a convex shape. The KX–selected galaxies are probably ellipticals, reaching their peak emission long–ward $1.6\,\mu$m, the maximum of the photospheric emission for stellar populations at rest–frame wavelengths. Stellar SEDs are much steeper compared to the galaxy ones and peak at $\sim 1\,\mu$m. Although spectroscopic observations will reveal the nature of the non–quasars, the above classification criteria give a rough, but realistic, estimation on the type of objects contaminating the KX–selected QSO sample.

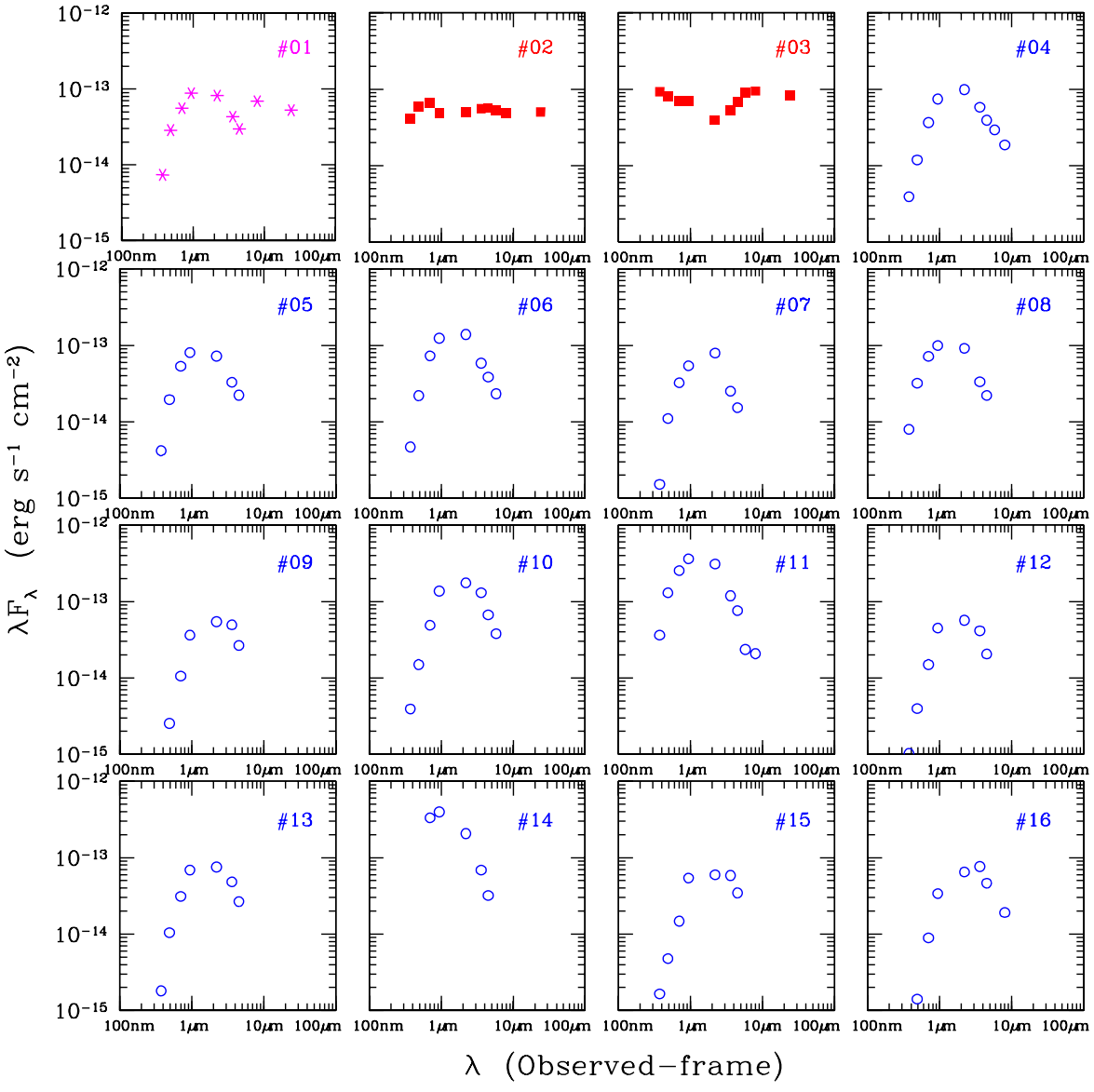

Figure 5.4: Spectral Energy Distributions for the 85 KX–selected quasar candidates: squares – real QSOs; circles – non QSOs; asterisks – non QSOs showing rise in wavelengths longward 4.5 μm).

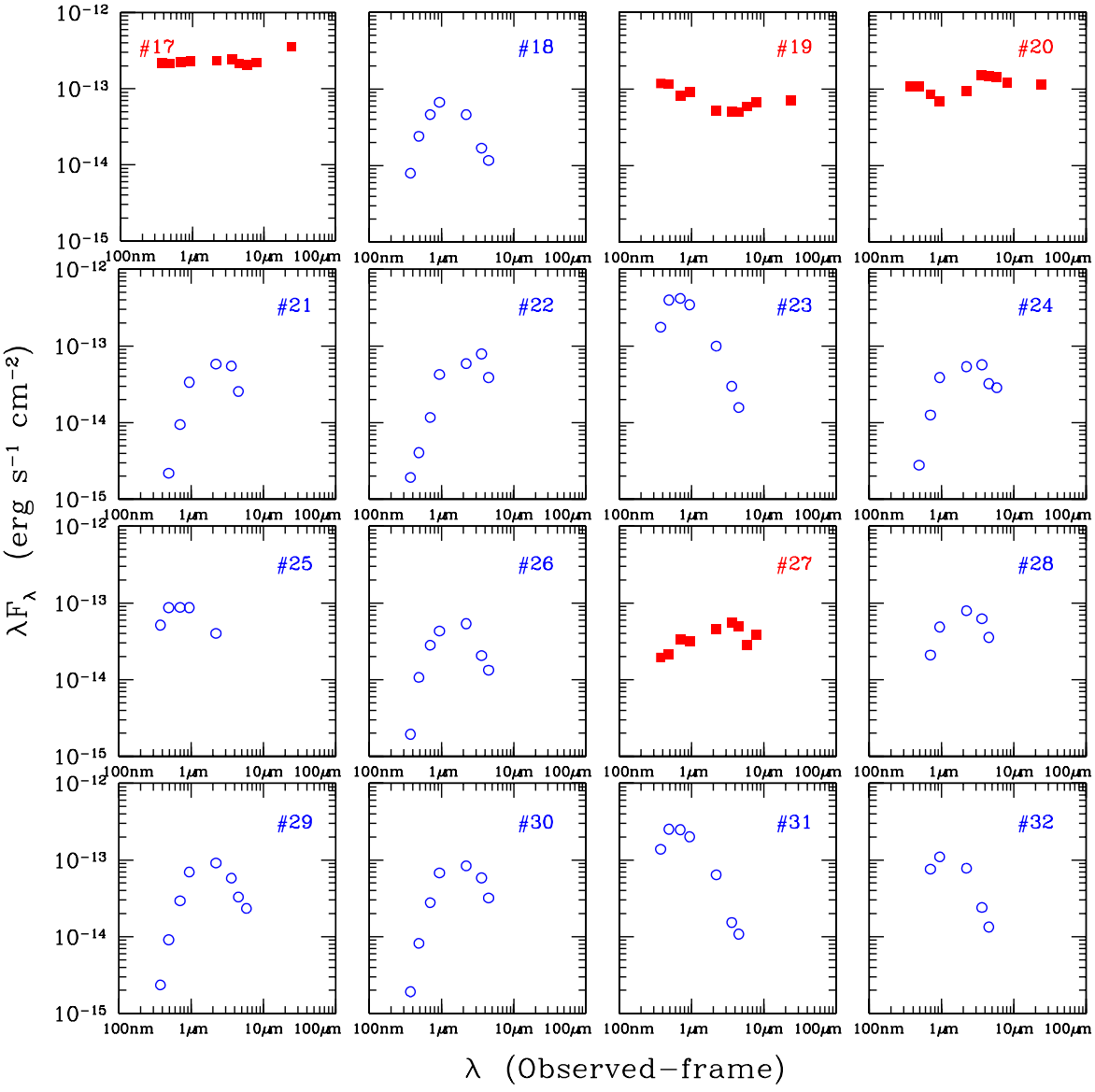

Figure 5.4: Continued.

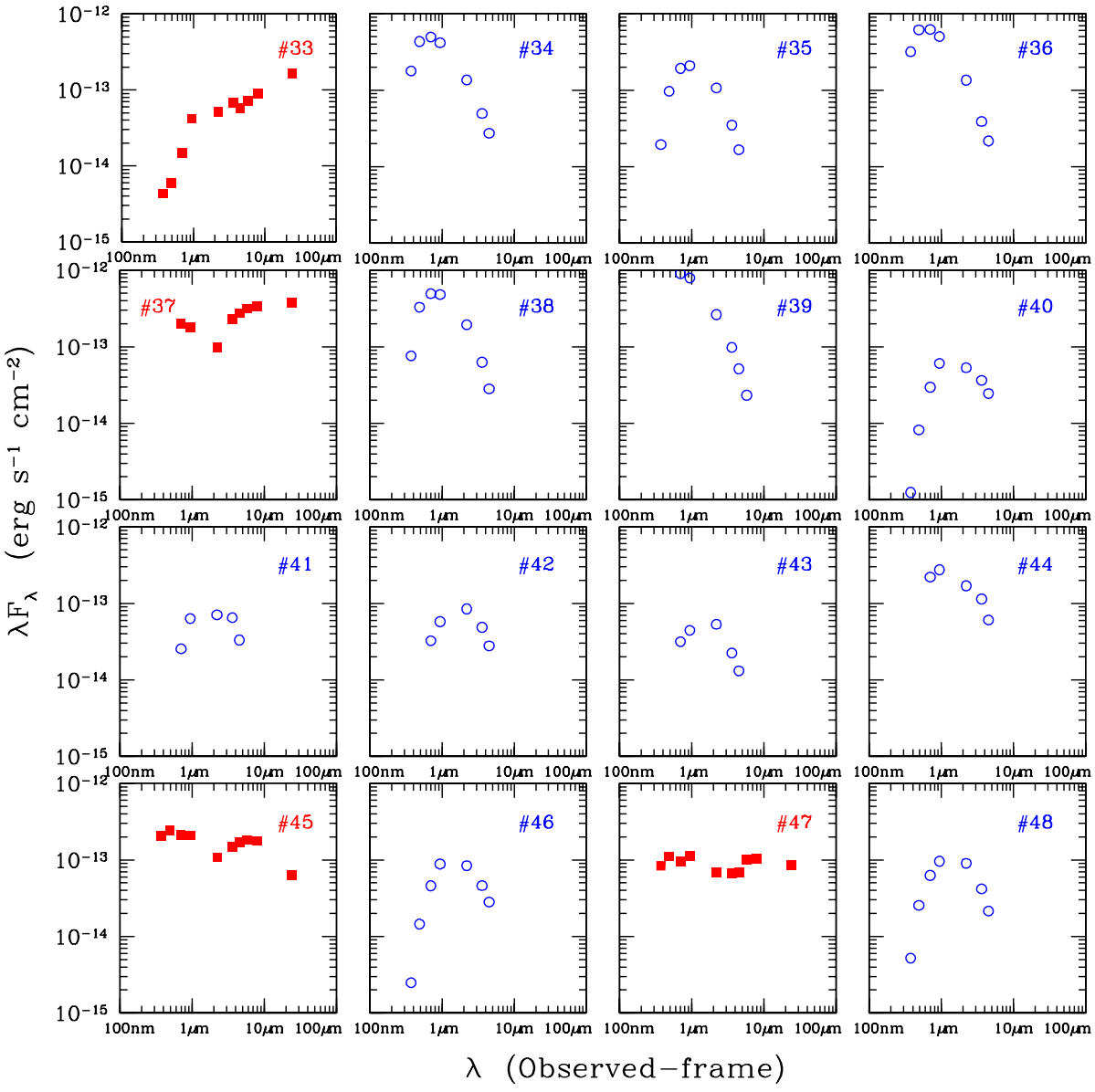

Figure 5.4: Continued.

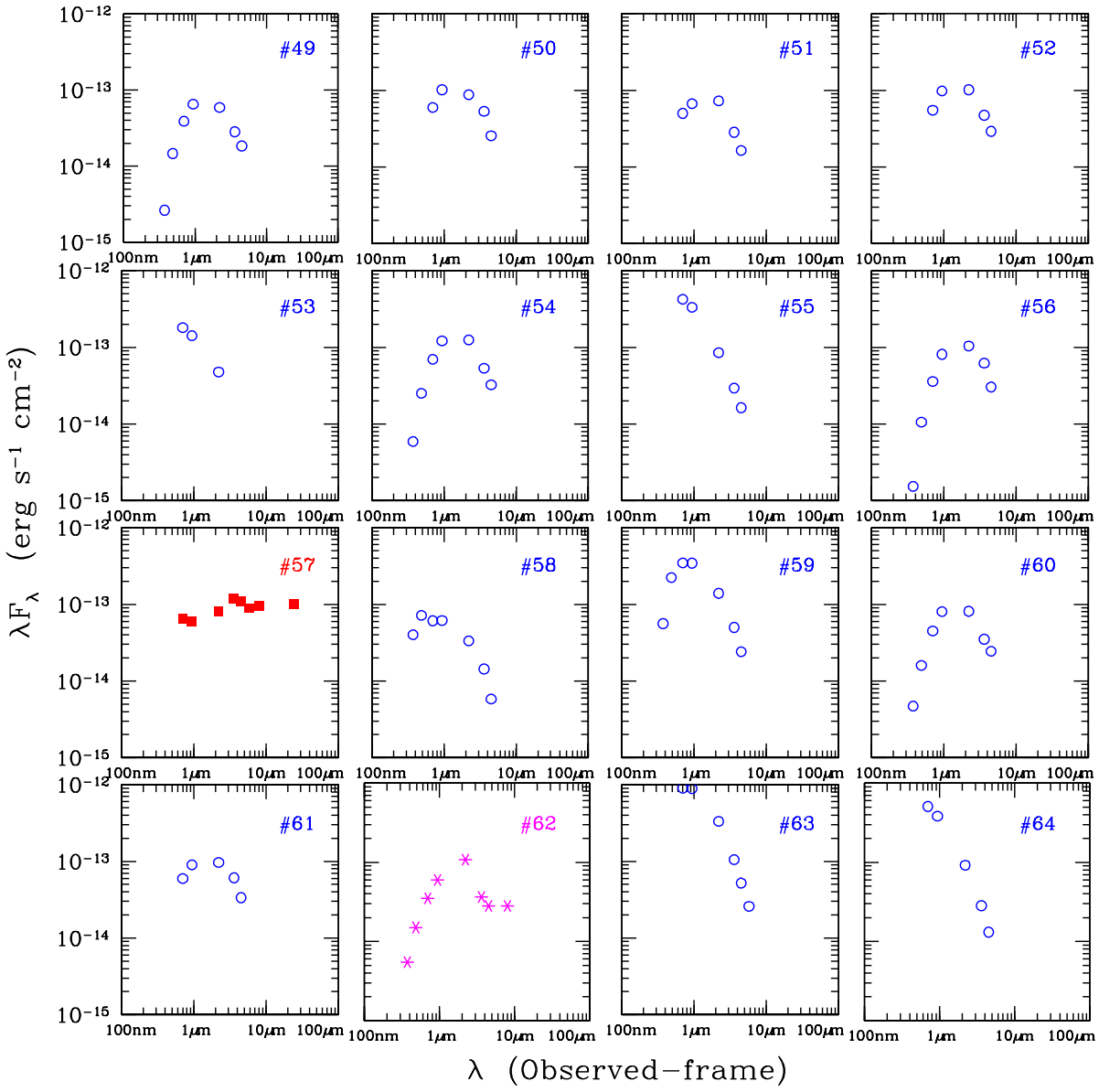

Figure 5.4: Continued.

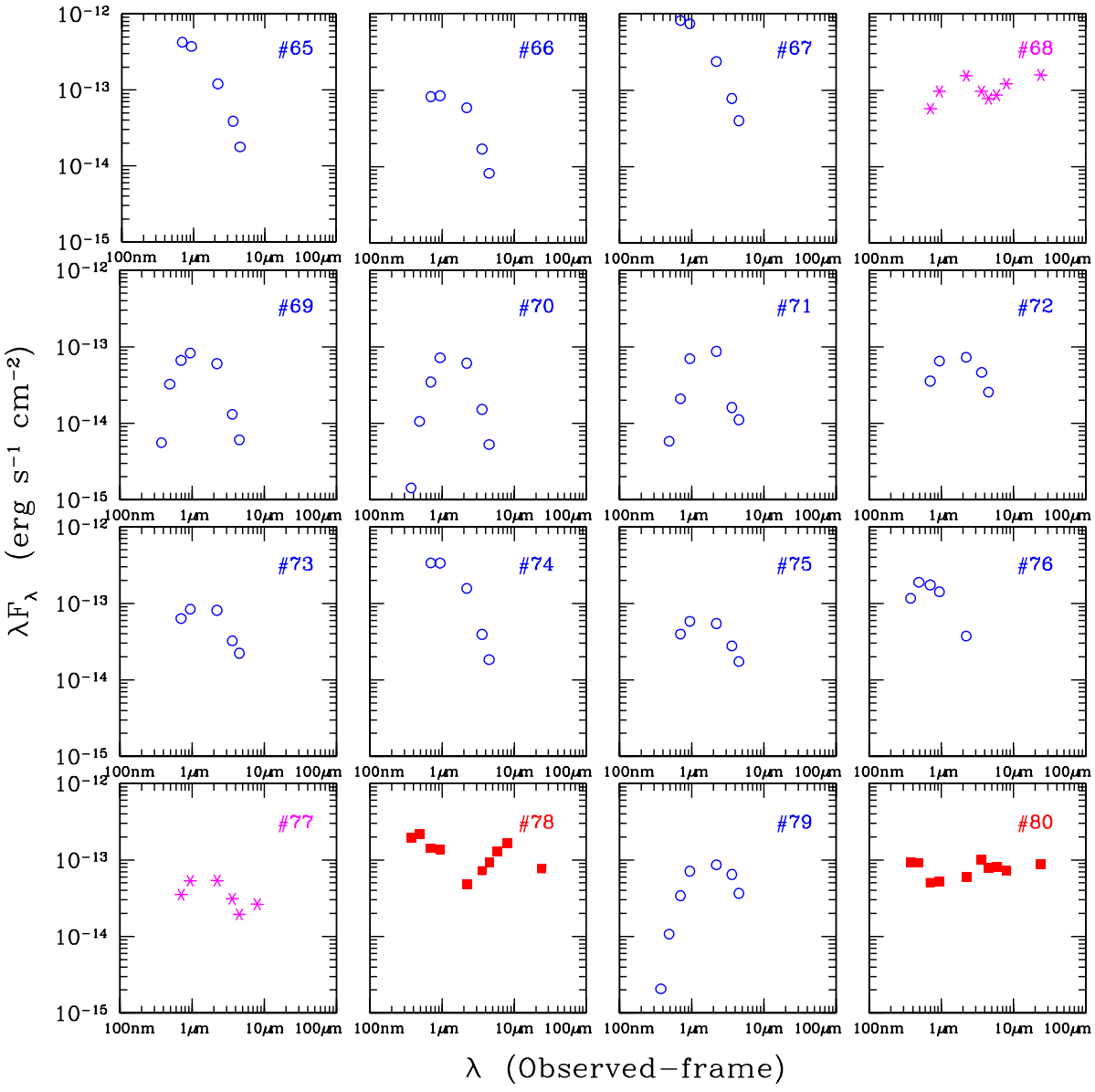

Figure 5.4: Continued.

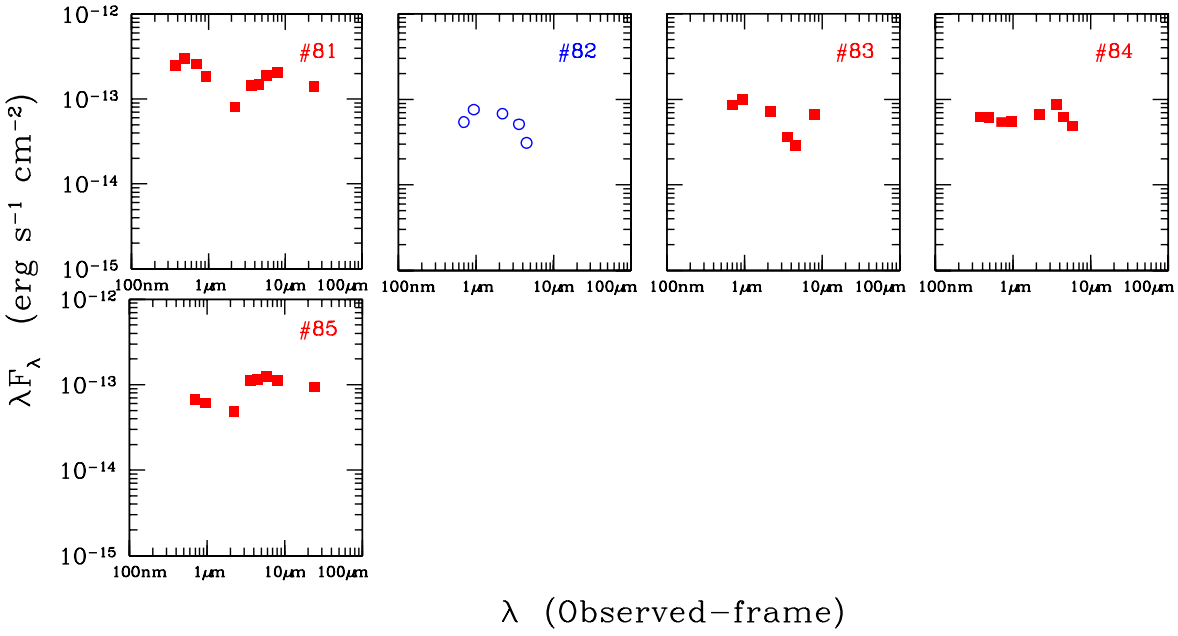

Figure 5.4: Continued.

5.3.2 Photometric redshifts for the KX–selected quasars

Using HyperZ we estimated the photometric redshift for the 17 identified quasars. The final decision on the object's redshift was taken by comparing both the primary and secondary solutions as obtained by various filter combinations. In the case where CFHTLS data were available, we compared the results obtained with CFHTLS to those obtained when combining the CFHTLS photometry with the first two of the IRAC bands. Due to the lack of CHTLS photometry, we compared the results obtained when using the $Rz'K_s$ photometry and when combining the $Rz'K_s$ and the photometry for the first two IRAC bands.

For some cases the two primary solutions (CFHTLS – CFHTLS + IRAC1,IRAC2 or $Rz'K_s$ – $Rz'K_s$ + IRAC1,IRAC2) were in good agreement, thus there was little doubt concerning the object's redshift. For other cases, however, more thought had to be given, since the two data sets gave quite different results. As it was explained in paragraph 5.1.3, the photo–z value derived when using the CFHTLS photometry can be quite close to the real redshift of the object, even if the best–fit template does not describe the object's SED. Thus, when the two primary solutions were quite different, we tried to find a solution among the results obtained when using the CFHTLS + SWIRE photometry, with a redshift close to the one given by the CFHTLS on its own. When looking to the CFHTLS + SWIRE solutions, very high A_V values ($A_V > 0.6$) combined with a "strange" template (i.e. not one among the templates chosen for the 2dF QSOs) were considered as less reliable. In this way, most solutions were finally excluded. In case of doubt among the

remaining solutions, we visually inspected each of the solutions, i.e. we overplotted the redshifted template, properly corrected in flux for reddening and intergalactic absorption, together with the photometric points. In this way it was finally possible to discard all but one solution, that best fit the available photometric data.

Table 5.7 presents HyperZ's output for the real quasars in our sample. Column 1 is the object's identification, taken from Table B.1. The second, third and fourth columns give the photo–z value, its χ^2 and the corresponding probability, respectively. Columns 5 and 6 give the best–fit template and the value for A_V, respectively. Finally, Column 7 informs on the data sets used for calculating the photometric redshift of the object.

ID	Photo-z	χ^2	Prob(%)	Template	A_V	Data
(1)	(2)	(3)	(4)	(5)	(6)	(7)
02	0.885	1.079	36.96	QSO–low	0.6	CFHTLS + Swire
03	1.762	0.218	95.48	QSO–high	0.0	CFHTLS + Swire
17	2.265	0.632	70.49	QSO–low	0.6	CFHTLS + Swire
19	2.301	0.190	97.97	QSO–low	0.2	CFHTLS + Swire
20	0.774	1.469	19.61	QSO–high	0.0	CFHTLS + Swire
27	1.439	0.809	54.26	CWW–Im	0.4	CFHTLS + Swire
33	0.556	1.495	18.75	Seyf1.8	1.2	CFHTLS + Swire
37	0.850	2.284	5.77	QSO–high	0.0	$Rz'K_s$ + Swire
45	1.232	0.272	95.01	QSO–low	0.4	CFHTLS + Swire
47	2.367	0.097	99.66	QSO–low	0.4	CFHTLS + Swire
57	0.110	0.722	57.68	QSO–low	1.0	$Rz'K_s$ + Swire
78	1.768	0.676	66.92	QSO–med	0.0	CFHTLS + Swire
80	1.641	0.841	53.80	CWW–Im	0.0	CFHTLS + Swire
81	0.870	0.378	89.32	QSO–low	0.2	CFHTLS + Swire
83	0.163	0.439	78.06	Seyf1.8	0.0	$Rz'K_s$ + Swire
84	1.584	0.714	63.79	CWW–Im	0.0	CFHTLS + Swire
85	0.625	1.242	29.07	QSO–high	0.2	$Rz'K_s$ + Swire

Table 5.7: Photometric redshifts for the 17 quasars found among the 85 KX–selected QSO candidates.

The photometric points and the redshifted–calibrated templates for the objects with CFHTLS and Swire photometry are shown in Fig. 5.5. Similar results for the remaining cases, where the $Rz'K_s$ photometry replaced the CFHTLS measurements, are illustrated in Fig. 5.6.

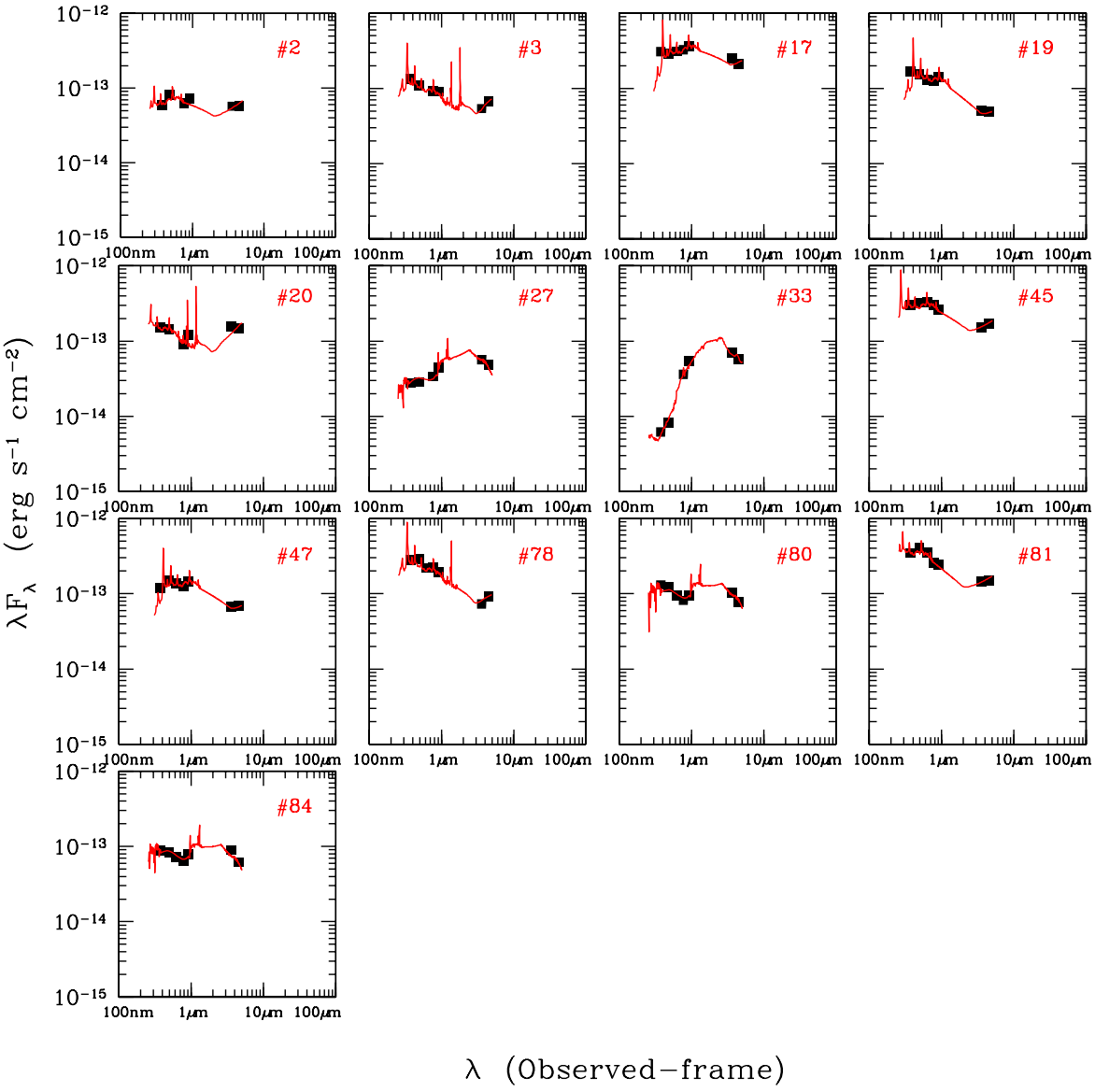

Figure 5.5: Spectral energy distribution for the real quasars among the KX–selected candidates. The photo–z has been estimated on the basis of CFHTLS and IRAC1, IRAC2 photometry. For each object, overplotted is the best–fit template found by HyperZ, redshifted in wavelength, properly corrected for extinction, and calibrated in flux in order to match the photometric points.

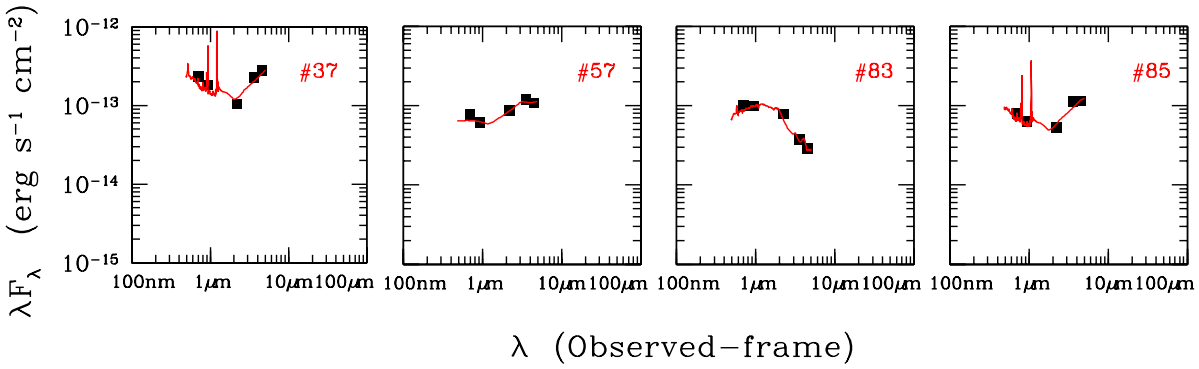

Figure 5.6: Same plot as in Fig. 5.5 for the objects for which the photometric redshift has been computed using the $Rz'K_s$ and IRAC1, IRAC2 photometry.

The visualization of the spectral energy distributions has proved to be an extremely useful tool for the identification of the contaminants in the KX–selected sample. Additionally, it has been possible to estimate the redshift of the real quasars using the HyperZ code. Both the construction of the objects' SED and the photo–z estimation have been carried out thanks to the existing multi–wavelength information, available within the context of the XMM–LSS project. Hence, although spectroscopic observations play an essential role in the full characterization of the properties of the quasars (spectro–z, identification of broad/narrow emission lines and measuring of their equivalent width), we have demonstrated that multi–wavelength photometry can significantly contribute to understanding the properties of the objects contained in a given sample.

Chapter 6

Evaluating the KX–method

The identification of the real quasars among the KX–selected candidates has marked the end of a major part of the data analysis. Thus, the time has come to perform an in–depth evaluation of the K–excess technique. The initial sample contained 93 candidates. For 8 objects we had a spectroscopic identification, while the remaining 85 were finally classified on the basis of their SEDs as QSOs, stars or galaxies. With this approach, another 17 objects were also identified as quasars, resulting in a total of 25 objects out of 93. In terms of contamination, i.e. false detections with respect to the total number of candidates, the KX–selected sample suffered at a level of $\sim 70\%$, a quite high percentage that could potentially hamper the success of a forth–coming spectroscopic survey. Therefore, if we want to refine the selection process, it is essential to understand whether this contamination originates from the applied morphological (Fig. 3.7) or color (Fig. 3.8) criteria applied.

Apart from the contamination issue, we also need to estimate the success rate of the method, i.e. the number of quasars selected by KX with respect to the total number of quasars that the $Rz'K_s$ catalog contains. By comparing the number of KX–selected quasars (normalized to surface unit and corrected for the detection probability corresponding to the magnitude limit of our catalog) to the ones predicted in the literature, we expect to obtain a rough idea on whether the KX has detected all QSOs contained in the $Rz'K_s$ catalog.

6.1 The KX–contaminants

The distribution, in the color–color diagram, of the 25 real quasars found among the 93 candidates is shown in the left panel of Fig. 6.1. The colors of the false candidates, together with those of M–type stars (M 3–M 9) (Dahn et al. 2002), have also been plotted, for comparison purposes. From the plot it is clear that all but one KX–selected quasars are found in the bluer part of the vertical axis, with $(R - z') < 0.8$. The single quasar found

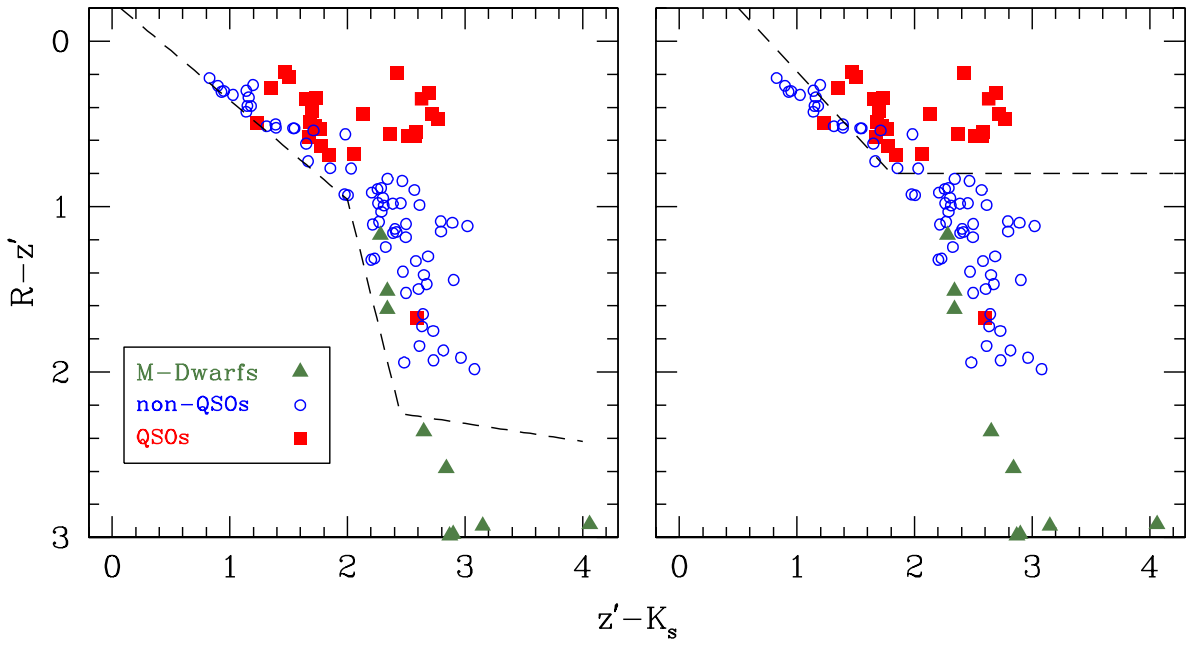

Figure 6.1: Distribution in the color–color diagram of the real quasars (filled squares) and the contaminants in the sample (open circles). The filled triangles correspond to M–type dwarfs (Dahn et al. 2002). *Left panel:* The dashed line corresponds to the actual color criteria applied for separating the stellar from the non–stellar sources. *Right panel:* Modifying the color criteria would strongly reduce the number of contaminants, with a minor impact on the number of missed QSOs.

at about $(R - z') \approx 1.7$ is the one corresponding to ID #33, which, based on its SED, is the only Type–II QSO contained in our sample. Although the 2dF–spectroscopically identified QSOs were also occupying the region where the majority of the KX–selected quasars has been found (see Fig. 3.9), their limited number could not allow to set more strict constraints on the color criteria we had applied.

The study of the SEDs of the KX–selected candidates has revealed that the sample contains galaxies and stars, the latter ones probably being M–type dwarfs, as their colors coincide well with the M–dwarf $Rz'K_s$ photometry found in the literature (Dahn et al. 2002). Revisiting the color criteria applied for separating the sources with stellar or non–stellar colors can help significantly improve the selection efficiency. A $(R-z') < 0.8$ color criterion would reduce the number of contaminants, having a minor impact on the number of missed quasars that the KX could have potentially selected (only object #33 would be missed, out of 25 QSOs). Therefore, using the $(R - z') \leq 0.8$ color criterion exclusively, with no further color selection in the $z' - K_s$ color index, would give a contamination rate of the order of $\sim 50\%$ ($\sim$ 22 non–QSOs out of 46 candidates).

For removing the contaminants with colors bluer than $(R - z') \leq 0.8$ one has to

be more careful, as the majority of these objects is found very close to the border–line separating the sources with stellar/non–stellar colors. The "blue contaminants" problem can be solved by shifting the separating line to the right, as proposed in the right panel of Fig. 6.1. Using this $(z' - K_s)$ color criterion would drastically reduce the contamination rate to only $\sim 30\%$ (8 non–QSOs out of 31 objects) with only one more QSO missed, as it can be seen in the corresponding color–color diagram.

Objects #1, #62, #68 and #77 of Fig. 5.4 are of particular interest, as their spectral energy distribution does not have a typical stellar or galaxy–like shape, that declines toward infrared wavelengths, neither shows the typical features found in a QSO–like SED. The rise in their infrared emission, at wavelengths $\lambda \geq 4.5\,\mu$m, could either be explained by starburst activity, should the object be a galaxy, or by the presence of a circumstellar disk around a low-mass ($M \sim 1\ M_{\odot}$) pre–main sequence star (T Tauri type). Given the point–like morphology of the objects, we decided to investigate further more on the latter scenario.

The circumstellar disks found around young stars emit most of their energy in the infrared, as a result of the disk's reprocessing of the radiation emitted by the stellar photosphere. Consequently, the signature of the presence of such disks comes from the excess observed at infrared wavelengths over the spectral energy distribution corresponding to the star's photosphere (Kenyon & Hartmann 1987).

Fig. 6.2 shows the typical SED for two T Tauri stars, taken from Gras–Velázquez & Ray (2005). The different symbols correspond to multi–epoch observations spanning a wide wavelength range (UBVRI, JHKL, ISO and IRAS photometry). The solid line represents the black body (BB) distribution of the stellar photosphere best fitting the observational points (spectral type M0 for CoKu Ser/G7 and K7 for V536 Agl). The dashed line represents an alternative BB curve, when a different model (e.g. temperature, alternative extinction, etc) also fits well the observations. The excess at wavelengths longward of $5\,\mu$m, due to the presence of a disk, is clearly visible.

Although the SEDs of the four objects resembled the ones presented in Fig. 6.2, the T Tauri scenario was finally discarded. One reason is the very faint magnitudes of the objects (between 20.8–21.4 in the R–band). The magnitudes of the T Tauri stars reported in Gras–Velázquez & Ray (2005), found at distances between 150–300 pc, were of the order of 11–12, for the same band. Riaud et al. (2006) also report on similar properties, with a typical magnitude $R = 11.35$ for a T Tauri star found at a distance of ~ 140 pc. This would mean that, being 10 magnitudes fainter, our objects would lie at distances ~ 100 times further away than the typical distance at which young star populations are frequently found (assuming similar absolute magnitudes). The second argument is the lack of any elements suggesting an over–density of interstellar clouds, contributing to the birth of new stellar populations. A visual inspection of optical images in the region these objects were found gave no signs of absorption, with the object surface density being comparable to that in regions away from the objects under question. In conclusion, objects #1, #62, #68 and #77 are probably starburst galaxies.

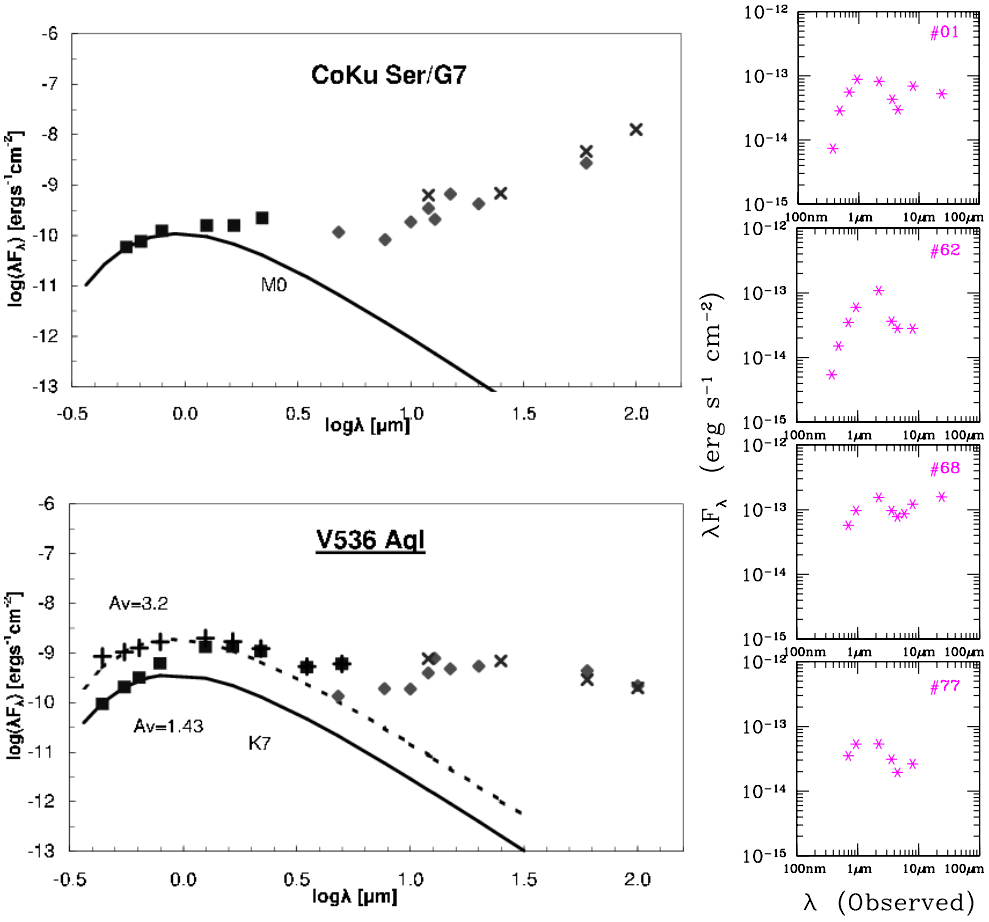

Figure 6.2: *Left Panel:* Typical spectral energy distributions for T Tauri stars, taken from Gras–Velázquez & Ray (2005). The excess in the IR part of the SED is due to the circumstellar disk around the star, that absorbs and re–emits the radiation coming from the star's photosphere. *Right Panel:* Spectral Energy distributions for four of the QSO candidates, which show similarities to the T Tauri SEDs.

6.2 Model predictions – Number counts

In paragraph 1.6.4 we raised the subject of selection biases and clearly stated that each survey selects objects complying with the survey's selection criteria. It is therefore obvious that the evaluation of the K–excess could not be completed unless we investigate on the possibility of having missed objects not respecting the color and morphological criteria we implemented for identifying the QSO candidates. For doing so we compared the number of KX–selected quasars with the numbers predicted using the model of Xu et al. (2006, in preparation), for the $Rz'K_s$ limiting magnitudes of our survey. This model, treating the evolution of dusty and E/S0 galaxies, was originally developed to address two major questions concerning galaxy populations in deep infrared (IR) surveys:

- whether normal late–type galaxies or starburst galaxies (including galaxies with obscured active galactic nuclei) dominate among sources in deep IR surveys, and
- how much do E/S0 galaxies contribute to the counts in deep mid–infrared (3-20 μm) surveys.

An IR–based model was selected, as opposed to models based on optical properties of AGN, because part of the selection of AGN candidates for this work has been based on the evidence of torus emission in the IR part of their SEDs.

For a detailed description of the model see Xu et al. (2003) and Xu et al. 2006 (in prep.) Hereafter we briefly outline its major characteristics. Their approach consisted in testing three models for dusty galaxy evolution: S1 assumes that starburst galaxies are the dominant population; S2 assumes that normal galaxies dominate and S3 is an intermediate model. Comparing the model predictions with a wide range of observational data collected from the literature, they found that none of these models could be ruled out. In Xu et al. (2006), galaxies with an active nucleus, one of the populations in the new model S4, have the same definition and evolution function as in the S1 model (Xu et al. 2003). These are IR galaxies having "MIR-excess" (rest frame): $f_{25\mu m}/f_{60\mu m} \geq 0.2$. Their evolution function is assumed to be the same as that of optical QSOs, which is a pure luminosity evolution function of the form (Boyle et al. 2000):

$$L^*(z) = L^*(0)10^{1.36z-0.27z^2} \quad (z \leq 7) \tag{6.1}$$

but with L^* now defined at 25 μm.

The model was run for the limiting magnitudes of the $Rz'K_s$ survey, namely $R = 23.5$, $z' = 22$, and $K = 18.0$ and resulted in an expected number of 222 AGN *of both types* (i.e. Type–1 and Type–2) per square degree.

Before comparing the number of KX–selected quasars with the model predictions, we first have to take under consideration that:

- The effective area of the survey was 0.68 square degrees.
- The $Rz'K_s$ catalog was cut at a threshold of $K_s = 18$ (Kron magnitude), which corresponds to a completeness of 80% (section 2.4).

Thus, the number of QSOs per square degree we expect to discover in the LCO survey is:

$\mathrm{N_{qsos}} = 25/0.68/0.8 \sim 46$

Before continuing with our analysis, it is important to emphasize that the models by Xu (2003, 2006) make no distinction between Type–1 and Type–2 objects, or any AGN sub–category, that could be attributed to orientation effects or a different morphology.

Nevertheless, although the models are based on certain assumptions, the difference between the model predictions and the number of detected KX–quasars is quite significant. For understanding the origin of this discrepancy, we will carefully revise the selection criteria applied and try to recover the missing population among the point–like and extended sources which have not been considered, until now, as quasar (and AGN, in general) candidates.

6.2.1 Point–like sources

In order to investigate on the possibility of having missed any QSOs/AGN among the point–like sources, we reproduced a SWIRE color–color diagram for the objects that were characterized as stars on the basis of their colors and point–like morphology. The results are shown in the upper panel of Fig. 6.3. As it can be clearly seen, three objects are located in the "AGN" region (as defined by Stern et al. (2005)). The study of their SEDs, shown in the lower panel of Fig. 6.3, revealed that objects #S01 and #S02 are quasars, while the third one is probably a starburst galaxy. The coordinates and $Rz'K_s$ photometry of these two new QSOs are presented in Table D.1.

Using the available CFHTLS and IRAC 1+IRAC 2 photometry, we estimated the photometric redshifts for the two quasars and for object S03. The obtained results are summarized in Table 6.1. As expected, objects S01 and S02 were fitted with a quasar template. S03 was indeed fit with a starburst template and is found at a much lower redshift. The observational points used for computing the photometric redshifts, together with the best–fit templates, are illustrated in Fig. 6.4.

For understanding the reason for which objects S01 and S02 were missed by the K–excess, we looked for their colors in the $Rz'K_s$ color–color diagram. As they were found very close (about 0.2 magnitude in each direction) to the separating line in the $(R - z')$ versus $(z' - K_s)$ color–color diagram (see left panel of Fig. 6.1), it is not surprising that objects S01 and S02 were not included in the catalog of quasar candidates.

ID	Photo-z	χ^2	Prob(%)	Template	A_V	Data
(1)	(2)	(3)	(4)	(5)	(6)	(7)
S01	1.384	0.593	73.65	QSO–low	0.00	CFHTLS + Swire
S02	2.344	0.068	99.88	QSO–med	0.00	CFHTLS + Swire
S03	0.110	1.582	17.60	M82	0.20	$Rz'K_s$ + Swire

Table 6.1: Photometric redshifts for the 3 point–like sources found in the AGN locus of the SWIRE color–color diagram.

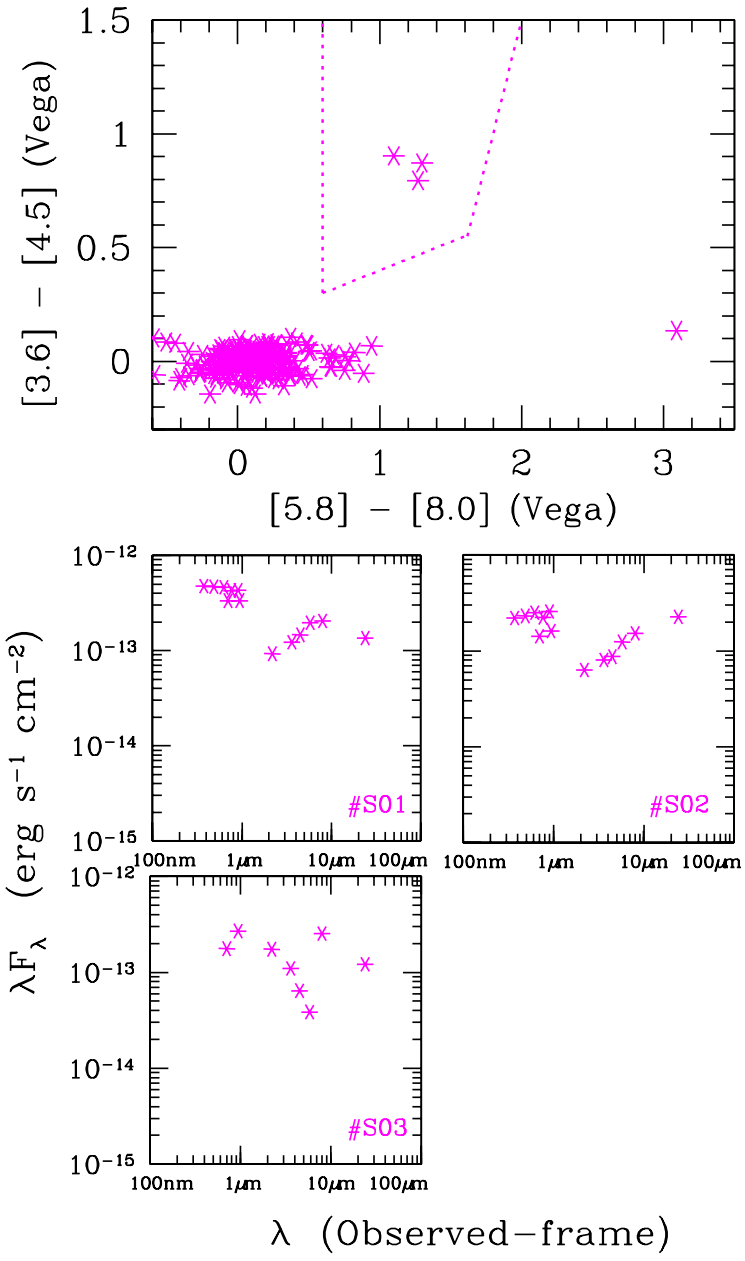

Figure 6.3: *Upper panel:* SWIRE color–color diagram for the sources in the $Rz'K_s$ catalog which were classified as stars. Three of these sources are found in the AGN locus (Stern et al. 2005). *Lower panel:* Spectral energy distribution for the three sources found in the AGN locus. #S01 and #S02 have QSO–like SEDs, while #S03 is probably a starburst galaxy.

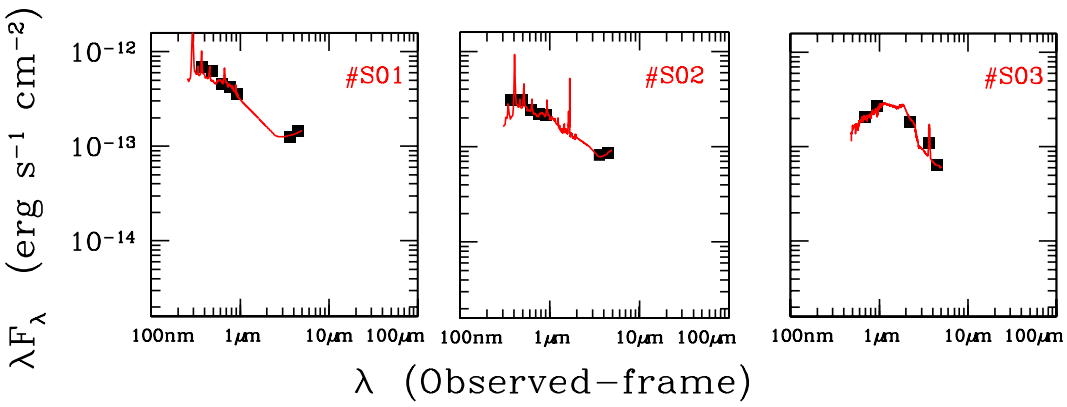

Figure 6.4: Photometric redshifts for the three point–like sources found in the AGN locus in the SWIRE color–color diagram (see Fig. 6.3). Objects S01 and S02 are QSOs, S03 a starburst galaxy.

6.2.2 Extended sources

An additional search for missing AGN took place among the sources with extended–like morphology. For this purpose we examined three different sub–samples:

- The sources showing X–ray emission (32 objects).
- The sources located in the AGN locus in the SWIRE color–color diagram presented in Fig. 4.20 (27 objects).
- The sources which show emission at 3.6, 4.5 and 5.8 μm (774 objects). This criterion serves at detecting AGN with emission below the sensitivity limit of the SWIRE survey for wavelengths longward of 5.8 μm. Although these sources are not present in the SWIRE color–color diagram, three bands are sufficient for detecting a rise in the infrared part of the AGN's spectral energy distribution, as a signature of the torus emission.

For all objects in the three sub–samples we constructed their spectral energy distribution using the available photometry. Based on their SEDs, we identified:

- 8 AGN (1 Type–1 and 7 Type–2) in the X–ray sub–sample.
- 14 AGN (4 Type–1 and 10 Type–2) in the AGN–locus sub–sample.
- 25 AGN (2 Type–1 and 23 Type–2) among the sources with emission in the first three IRAC bands.

All 8 AGN with X–ray emission were also included in the sub–sample of the sources occupying the SWIRE AGN locus. Additionally, 7 objects with emission in the first

three IRAC bands were also found in the latter sub–sample. Thus, the total number of single entries finally reduces to 32 (3 Type–1 and 29 Type–2). Information regarding the astrometry and photometry of the additional AGN found in the $Rz'K_s$ catalog is given in Table D.2.

Due to the lack of spectroscopic data, we calculated the photometric redshift of these AGN following the same methodology as in section 5.1.3, i.e. when CFHTLS data were available, the photo–z was computed using the $u^*g'r'i'z'$ and the first two IRAC bands (3.6 and 4.5 μm), otherwise we used the $Rz'K_s$ and IRAC 1–IRAC 2 filter combination.

The observational SEDs, together with the best–fit (redshifted, corrected for absorption and flux–calibrated) templates for the 32 AGN found among the sources with extended morphology are shown in Fig. 6.5. The first three cells correspond to the objects which have been characterized, from the features in their SEDs, as Type–1 (#1, #3 and #121), while the rest are Type–2 objects. The photometric points used for the photo–z estimation are plotted as filled squares. The emission in the 5.8, 8 and 24 μm bands, although not used in the photo–z estimation, is also plotted (open squares), for demonstrating the evolution of the energy distribution toward mid–IR wavelengths. With the exception of objects #7 and #9, the corrected rest–frame templates seem to agree well with the observed emission at $5.8 - 24\,\mu$m. Although this implies that the redshift estimation for these two objects might be erroneous, it corresponds to a minor fraction of the population ($\sim 6\%$) and has no significant impact on the conclusions that we will draw from our study. The photometric redshifts and additional information from HyperZ regarding the 32 AGN is summarized in Table 6.2. Although the energy distribution of ten objects was better fit with a quasar template, when we refer to the objects in Table 6.2 we will use the more general term "AGN", as in our analysis we call quasars those AGN with a point–like morphology.

Separating the Type–1 and Type–2 populations is straightforward, given the striking differences in the shape of their SEDs. As it has been already explained, the origin of these differences is related to the attenuation of the light emitted from the accretion disk in the active nucleus, which is then absorbed and reprocessed by the dusty torus. The A_V coefficients found by HyperZ also agree with the "visual" identification of the two populations. More precisely:

- Only one out of three Type–1 AGN (object #121) experiences a relatively high extinction in the optical, with $A_V = 0.4$. The object's energy distribution (Fig. 6.5) indeed reveals the presence of dust, that modestly reddens the object's emission in the UV & optical (fitted by a quasar template).

- Concerning the Type–2 population, 50% of the sources have an A_V coefficient with values between 0.6 and 1.2 (the maximum value set in HyperZ), with most of them associated with a Type–1 quasar template which undergoes major UV–optical absorption. Regarding the remaining 50%, for the majority of the sources HyperZ has chosen a Type–2 template (quasar or Seyfert) which has the dust absorption effect

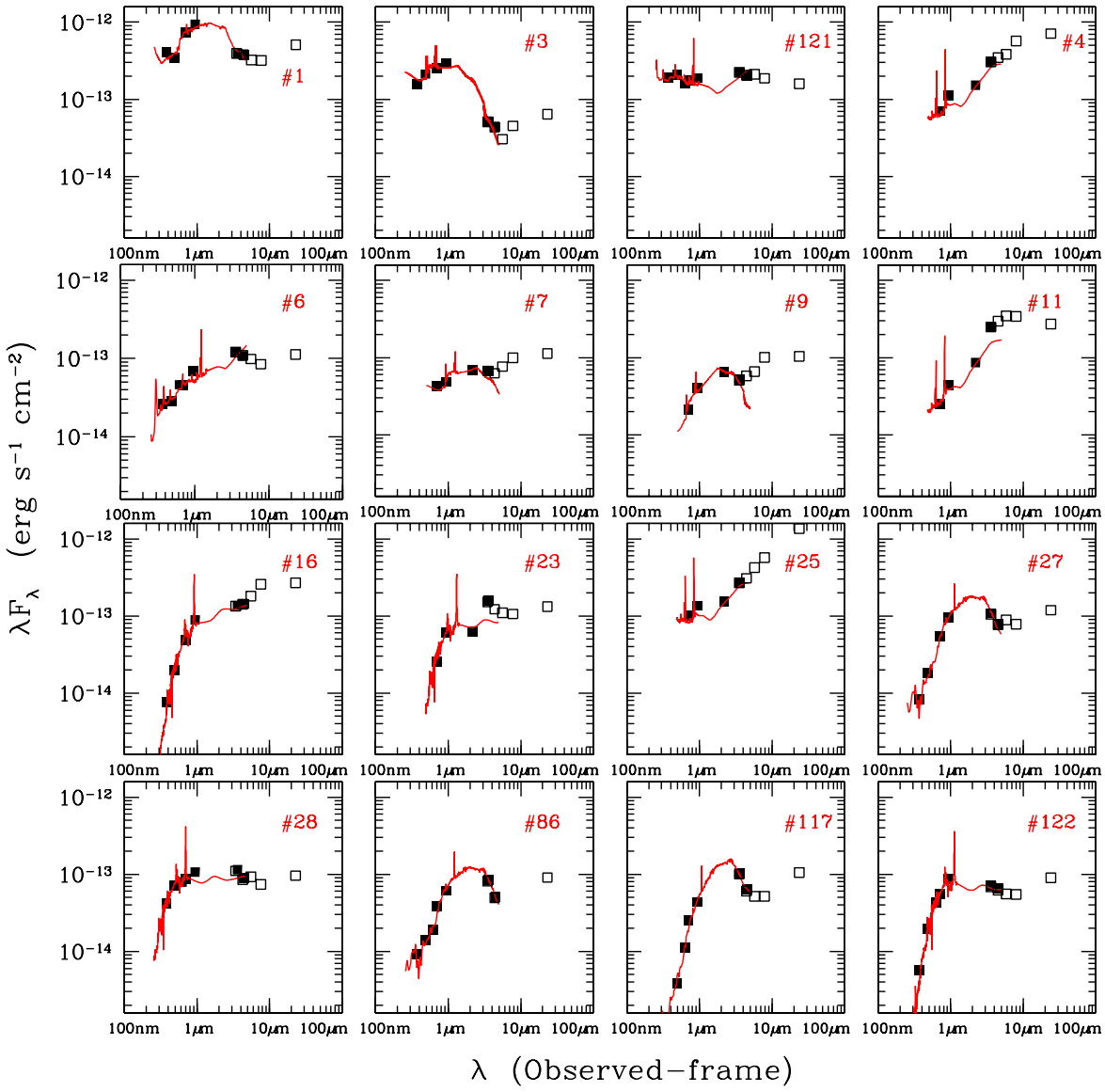

Figure 6.5: Spectral energy distribution for an additional sample of 32 AGN, found among the extended sources in the $Rz'K_s$ catalog. For each object the best–fit template, properly dilated in wavelength, corrected for absorption and calibrated in flux, is overplotted as a solid line. The filled squares indicate the photometric points used for computing the photometric redshift. The open squares correspond to emission in the 5.8, 8.0 and 24 μm bands (see text for more details).

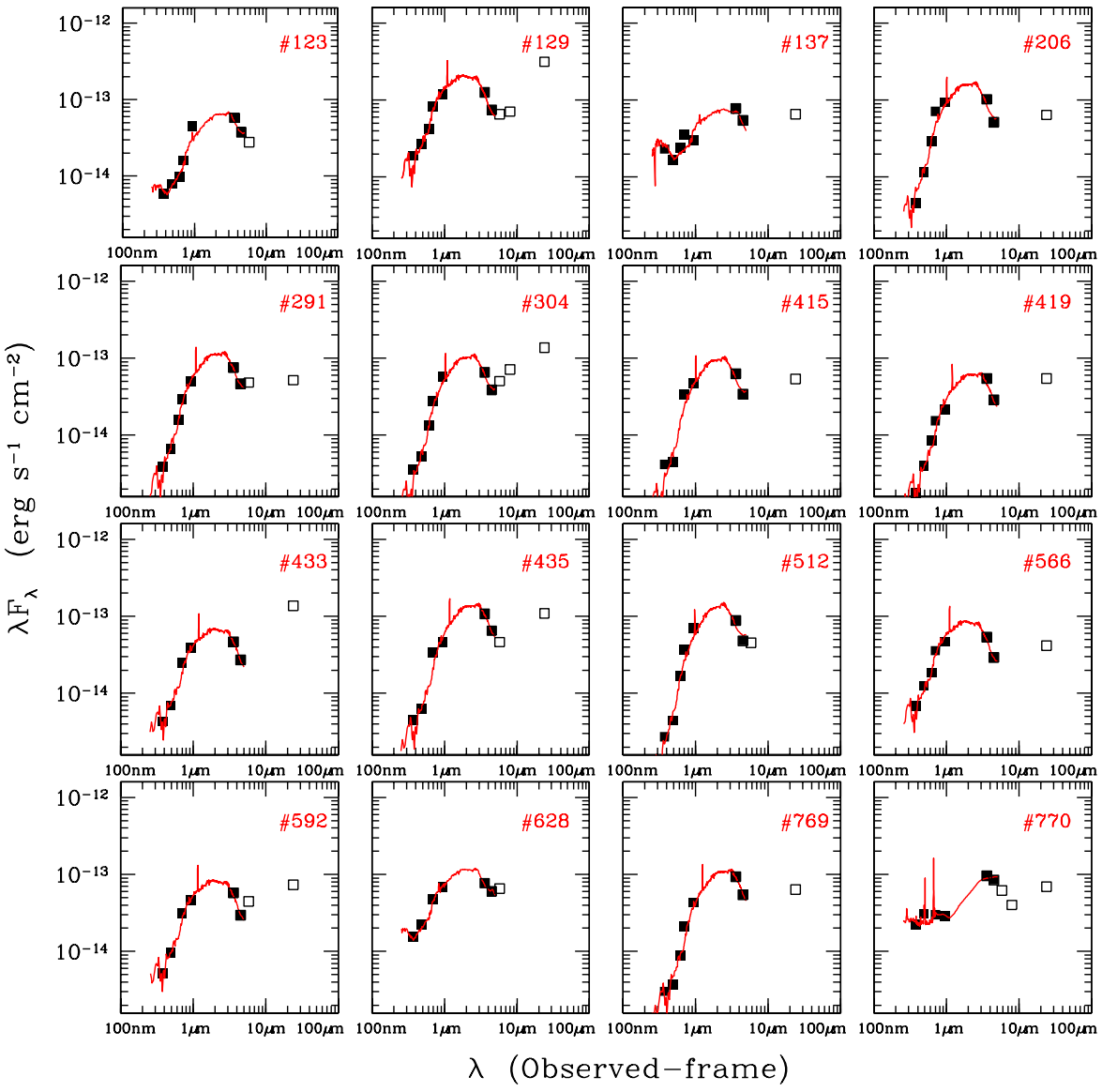

Figure 6.5: Continued

incorporated (see also Fig. 5.1).

ID	Photo–z	χ^2	Prob(%)	Template	A_V	Data
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	0.451	0.928	46.14	Seyf 1.8	0.00	CFHTLS + SWIRE 2
3	0.374	0.798	55.09	CWW–Im	0.00	CFHTLS + SWIRE 2
121	0.655	0.414	87.04	QSO–med	0.40	CFHTLS + SWIRE 2
4	0.276	0.199	93.92	QSO–high	0.80	$Rz'K_s$ + SWIRE 2
6	1.456	1.060	38.44	QSO–med	1.20	CFHTLS + SWIRE 2
7	1.491	0.154	96.12	CWW–Im	0.20	$Rz'K_s$ + SWIRE 2
9	0.781	0.055	99.43	CWW–Im	1.20	$Rz'K_s$ + SWIRE 2
11	0.255	0.299	87.86	QSO–high	1.20	$Rz'K_s$ + SWIRE 2
16	0.390	0.677	64.08	QSO–2	1.00	CFHTLS + SWIRE 2
23	0.989	0.353	84.18	QSO–2	0.40	$Rz'K_s$ + SWIRE 2
25	0.272	0.197	94.03	QSO–high	0.40	$Rz'K_s$ + SWIRE 2
27	0.722	0.454	81.08	Seyf 2	0.20	CFHTLS + SWIRE 2
28	0.063	0.790	57.78	QSO–2	0.20	CFHTLS + SWIRE 2
86	0.853	0.479	82.44	Seyf 2	0.00	CFHTLS + SWIRE 2
117	0.630	0.145	98.14	Seyf 2	1.20	CFHTLS + SWIRE 2
122	0.710	0.203	97.59	QSO–2	0.00	CFHTLS + SWIRE 2
123	0.885	1.237	28.37	Seyf 1.8	0.80	CFHTLS + SWIRE 2
129	0.676	0.765	59.75	Seyf 2	0.00	CFHTLS + SWIRE 2
137	1.289	2.061	5.43	Seyf 1.8	0.20	CFHTLS + SWIRE 2
206	0.534	1.393	21.30	Seyf 2	0.60	CFHTLS + SWIRE 2
291	0.651	0.193	97.89	Seyf 2	0.60	CFHTLS + SWIRE 2
304	0.570	1.017	41.17	Seyf 2	0.80	CFHTLS + SWIRE 2
415	0.536	2.270	4.48	Seyf 2	0.80	CFHTLS + SWIRE 2
419	0.826	0.808	56.36	Seyf 2	0.40	CFHTLS + SWIRE 2
433	0.813	0.141	98.26	Seyf 2	0.00	CFHTLS + SWIRE 2
435	0.804	0.723	60.63	Seyf 2	0.60	CFHTLS + SWIRE 2
512	0.486	2.017	5.97	Seyf 2	1.20	CFHTLS + SWIRE 2
566	0.673	0.524	79.03	Seyf 2	0.00	CFHTLS + SWIRE 2
592	0.755	0.113	98.94	Seyf 2	0.00	CFHTLS + SWIRE 2
628	0.691	0.149	98.03	Seyf 1.8	0.60	CFHTLS + SWIRE 2
769	0.900	1.235	28.44	Seyf 2	0.60	CFHTLS + SWIRE 2
770	0.011	0.526	75.66	QSO–high	0.60	CFHTLS + SWIRE 2

Table 6.2: Photometric redshifts for the 32 AGNs found among the extended sources. The first three objects are of Type–1, the rest of Type–2. Column 1 is the object's identification, taken from Table B.1. The second, third and fourth columns give the photo–z value, its χ^2 and the corresponding probability, respectively. Columns 5 and 6 give the best–fit template and the value for A_V, respectively. Finally, Column 7 informs on the data sets used for calculating the photometric redshift of the object.

6.2.3 Literature Query

We looked at the NASA Extragalactic Database[1] for finding any additional information on the 59 quasars and AGN discovered in our survey.

From the 8 QSOs with 2dF spectra we only managed to find some information regarding the source with identification #14 (Table 3.2). It has been detected in the VIRMOS survey, as it shows radio emission at 1.4 GHz (Bondi et al. 2003). Optical photometry (B,V,R,I, in AB magnitudes), obtained with the CFHT/CFH12K facility, is also available (Ciliegi et al. 2005). However, due to the limited number of photometric points available, there is no further information regarding the QSO nature of object #14. For the remaining 7 quasars with 2dF spectra, no information was found in the NED.

Regarding the objects with no spectra, VIRMOS radio measurements and CFHT photometry is available for the quasar with identification #57 (Table B.1) and the AGN #23, #25 and #117 (Table D.2). Similarly to the object #14, due to the lack of sufficient data, no classification is available for these three objects.

For the remaining objects no information was found in the literature. Consequently, we consider that the remaining 55 QSOs and AGN detected in our survey are new quasar discoveries, with spectroscopic identification for 7 of them.

6.3 Revisiting the KX selection criteria

Since the selection of the AGN and QSOs found in the $Rz'K_s$ catalog has come to its end, we are now in a position to evaluate the efficiency and completeness of the K–excess method and to investigate on any possible selection biases associated with it.

Using the K–excess method we managed to detected 25 quasars (24 of Type–1 and 1 of Type–2). These objects were selected among the sources with point–like morphology, having non–stellar $(R - z')$ and $(z' - K_s)$ colors. From what we have already seen, KX missed only two point–like sources with QSO–like SED. These objects were not in the initial QSO candidate sample, because, contrary to the selection criteria applied, their colors were similar to those of stars. Regarding the 32 AGN found among the objects with extended morphology, (3 of Type–1 and 29 of Type–2), the KX method did not select any of them since their morphological classification was not respecting our selection criteria.

Fig. 6.6 shows the $(R - z')$ vs $(z' - K_s)$ color–color diagram for all 59 quasars and AGN discovered using the steps described in the previous sections. The first conclusion we infer from this plot is that a cut in the vertical axis of $(R - z') \approx 0.9$ separates well the Type–1 from the Type–2 population. Although 24 out of the 25 KX–selected QSOs are found in the blue part of the $(R - z')$ axis, the KX also selected one Type–2 quasar; the

[1] http://nedwww.ipac.caltech.edu/

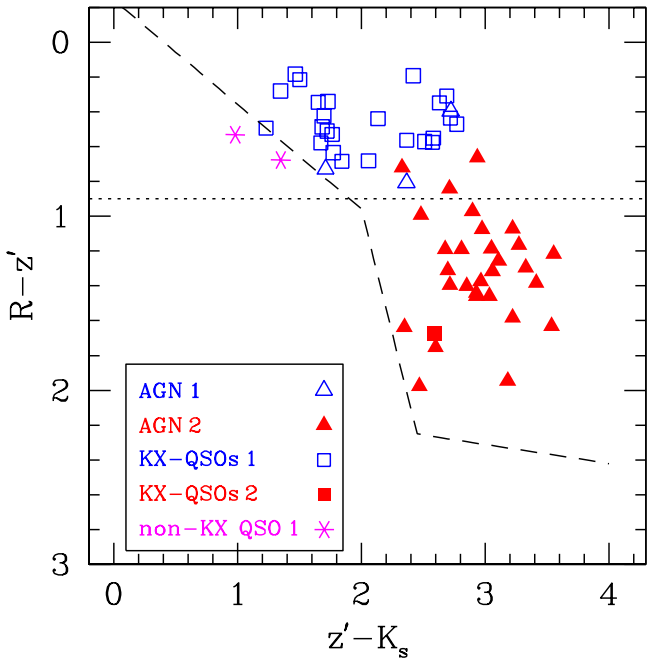

Figure 6.6: Color–color diagram for all QSOs and AGN found in the $Rz'K_s$ catalog. Open squares: KX–selected Type–1 QSOs; open triangles: Type–1 AGN; Filled squares: KX–selected Type–2 QSOs; filled triangles: Type–2 AGN; asterisks: Type–1 QSOs not detected by KX. The dotted line denotes the $(R - z') < 0.9$ color criterion, that separates the Type–1 from the Type–2 population.

reason is that this object has a high stellarity index in both bands. Nevertheless, from the color–color diagram presented in Fig. 6.6 it is clear that the KX favors the selection of "classical" quasars (which emit part of their energy in the UV and optical part of the electromagnetic spectrum) and QSOs undergoing reddening (which have a flat SED from the UV–optical to NIR), but can not detect Type–1 or Type–2 AGN, found at low redshift and which probably have an extended morphology due to the contribution of the host galaxy.

Another interesting conclusion we derive from the same plot concerns the number of QSOs missed by the KX. The two Type–1 QSOs missed lie at the edge between the point–like and extended locus. Should the separating line be slightly shifted toward objects with bluer $(z' - K_s)$ colors, these two sources would have also been included in the quasar candidate sample, with a significant impact, though, on the contamination of the candidate list, by stars. Hence, we should always expect a small number of quasars to be missed, should a cut in the $Rz'K_s$ color–color space be similar to the one we implemented.

Fig. 6.7 shows the morphological properties of the AGN that the KX failed to detect due to the very strict stellarity criteria applied ($\mathrm{stell}_R > 0.95$ and $\mathrm{stell}_{z'} > 0.85$). The

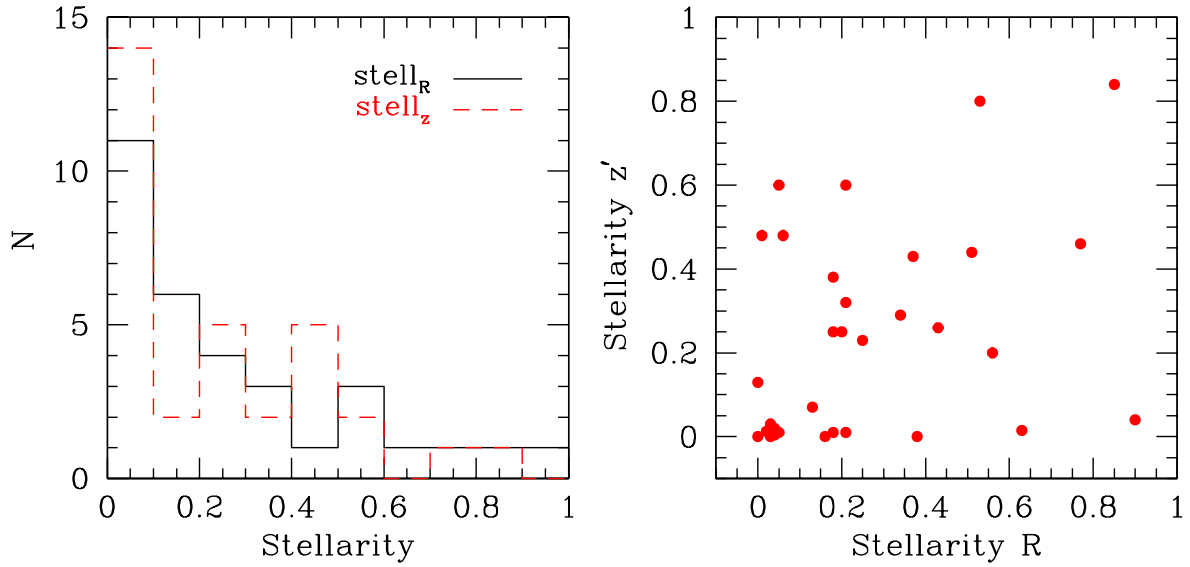

Figure 6.7: *Left panel*: Histogram of the R (solid line) and z'–band (dashed–line) stellarity indices for the 32 AGN not detected by the KX. *Right panel:* Correlation between the two stellarity indices.

left panel shows a histogram of the R, z'–band stellarity indices, while the right panel illustrates how these indices are correlated. From the plots it is obvious that the AGN have, in at least one band, a very low stellarity index, thus the K–excess can not select them.

In section 3.2 we had presented the properties of 9 quasars with spectroscopic identification, found among the $Rz'K_s$ objects. These objects were observed with the 2dF facility, during an observing run with the objective to target sources in the XMM–LSS field showing X–ray emission. These sources were recovered by the K–excess (see Fig. 3.9), since they were complying with our selection criteria. Nevertheless, another 5 spectroscopically confirmed quasars were not included in the initial $Rz'K_s$ catalog, since their detection probability was lower than the probability threshold used (70%) when producing the K_s–band catalog (see Table 3.2).

6.4 X–ray properties of the QSOs and AGN not detected by KX

To complete the analysis of the AGN and QSOs found among the various sub–samples, we studied their properties in the X–ray domain. The identification of the objects, their coordinates and emission in the different bands are presented in Table 6.3.

Fig. 6.8 shows a histogram of the hardness ratio for the Type–1 (short–dashed line)

No	RA (J2000)	Dec (J2000)	B	CD	B+CD	$F_{0.5-2}$	F_{2-10}
	h m s	° ′ ″				erg s^{-1} cm^{-2}	erg s^{-1} cm^{-2}
S1	2 24 24.17	−4 32 29.62	377.3920	78.4135	455.8055	2.43e-14	2.75e-14
S2	2 23 52.20	−4 30 31.54	83.3152	32.1903	115.5055	4.35e-15	8.77e-15
121	2 23 32.17	−4 57 39.38	212.8800	71.2999	284.1799	1.59e-14	2.99e-14
4	2 23 22.02	−4 57 38.09	21.9494	125.2940	147.2434	1.61e-15	5.12e-14
6	2 24 04.07	−4 51 18.54	151.0730	48.6913	199.7643	2.36e-14	2.37e-14
9	2 24 32.72	−4 48 16.63	86.2276	−1.0000	86.2276	9.26e-15	–
16	2 23 07.19	−4 37 02.75	131.2150	39.5536	170.7686	1.14e-14	1.93e-14
23	2 25 22.84	−4 26 47.94	63.0920	99.1653	162.2573	3.10e-15	2.57e-14
27	2 22 50.34	−4 22 53.76	141.9350	49.5883	191.5233	1.11e-14	2.16e-14
28	2 25 38.17	−4 21 08.86	82.4405	28.7362	111.1767	7.63e-15	1.66e-14

Table 6.3: X–ray properties for the AGN and QSOs detected among various sub–samples in the $Rz'K_s$ catalog. Columns 1, 2–3 provide the identification of the objects and their (RA,Dec) coordinates. Columns 4 and 5 present the source counts in the B $(0.5-2\text{keV})$ and CD $(2-10\text{keV})$ energy band, respectively and column 6 gives the sum of counts (B+CD) for each source. Columns 7 and 8 provide the flux in the B and CD energy bands, respectively. The horizontal line separates the Type–1 (QSOs #S1, #S2 and AGN with identification #121) from the Type–2 objects.

and Type–2 (solid line) AGN and quasars. The long–dashed vertical line separates the unabsorbed (HR ≤ -0.2) from the absorbed (HR > -0.2) sources. The HR distribution for the unabsorbed sources agrees well with their SED characterization as Type–1 objects. Regarding the 7 Type–2 AGN, though, there are only two sources (#4 and #23) with a HR corresponding to an absorbed AGN; the rest of the sources are found in the unabsorbed region.

The inconsistency between the optical and X–ray classification of Type–2 AGN has already been reported in the literature (e.g. Szokoly et al. (2004), Silverman et al. (2005), Garcet et al. (2006) in prep.). Since the largest difference between the optical and X–ray coordinates is not more than 2.5″, well below XMM–Newton's point spread function width, the mis–matching hypothesis can be excluded. Therefore, the reason for this disagreement has to be looked among explanations having a physical meaning:

- One possible explanation could be the fact that the hardness ratio limit that separates the absorbed from the unabsorbed sources is model dependent. When modeling X–ray spectra, the -0.2 value corresponds to an AGN with column density of $N_H = 10^{22}\text{cm}^{-2}$ (under certain assumptions, see Silverman et al. (2005) for more details). Higher column densities (which correspond to $HR > -0.2$) will absorb the broad emission–line region. Objects with $N_H > 10^{22}\text{cm}^{-2}$ will be optically classified as Type–2 AGN. Although different parameters would slightly shift the value that separates the absorbed from the unabsorbed sources, it is still difficult to explain the observed discrepancy for the Type–2 sources with a hardness ratio around

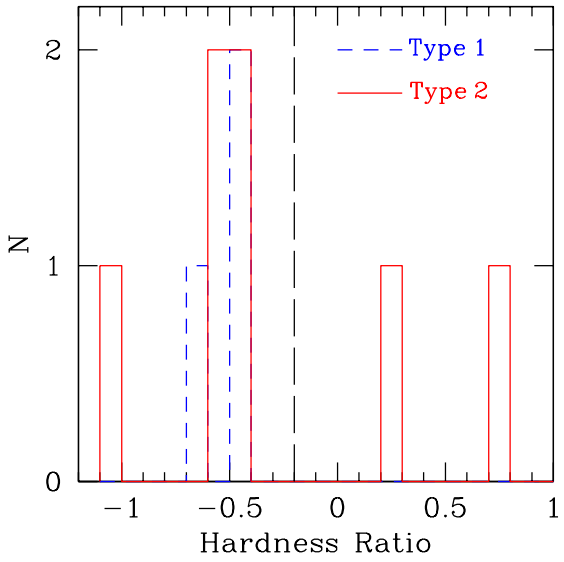

Figure 6.8: Hardness ratio histogram for the additional 3 Type–1 and the 7 Type–2 AGN, found in different sub–samples in the $Rz'K_s$ catalog. The characterization of the objects as Type–1/Type–2 has been based on their spectral energy distribution. For some of the Type–2 sources there is a clear discrepancy between their HR (which corresponds to that of unabsorbed sources, with HR ≤ -0.2) and their SED classification.

$HR \approx -0.45$.

- Another explanation could be the fact that the hardness ratio is sensitive to the object's redshift. From its definition ($HR = (H - S)/(H + S)$), the hardness ratio is computed based on the counts in the "soft" and "hard" energy bands, as they are measured in the observer's frame. Nevertheless, because of the power–law shape of the X–ray spectra ($S_\nu \propto \nu^\alpha$, with $\alpha \sim -0.7$), the HR ratio becomes "softer" as the redshift increases. This can be explained by the fact that the difference between the rest–frame soft and hard counts reduces as they are translated to higher redshift values. In their Fig. 8, Szokoly et al. (2004) present the evolution of the HR as a function of redshift, for different column densities. Based on this graph, we derived that, for a typical column density of $N_H = 10^{22}\mathrm{cm}^{-2}$ for a Type–2 AGN, a hardness ratio of $HR < -0.45$ (see Fig. 6.8) would imply the objects to be at redshift $z \approx 4$. As the maximum redshift of the Type–2 AGN does not exceed the $z = 1.5$ limit, the dependence of the hardness ratio on the redshift seems to be a rather poor explanation for the inconsistency between the optical and X–ray classification.

- Silverman et al. (2005) also propose several scenarios for explaining the lack of broad lines in Type–2 optical spectra of objects classified as unabsorbed, such as the dilution of the AGN emission by the host galaxy, a high dust–to–neutral gas ratio, or even beamed emission from BL Lac objects.

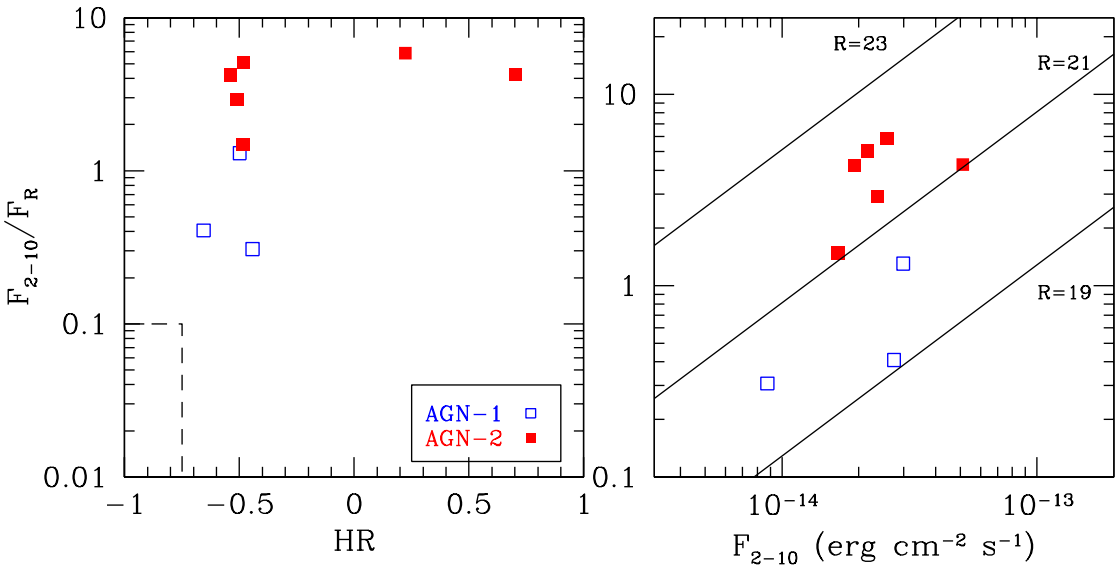

Figure 6.9: X–ray-to-optical flux ratio as a function of hardness ratio (left panel) and flux in the hard $(2-10$ keV) energy band (right panel). The open and filled squares correspond to the Type–1 and Type–2 AGN, respectively.

To conclude, although Fig. 6.8 clearly shows a discrepancy between the X–ray and SED classification, the lack of spectra does not allow us to provide a solid explanation.

Finally, the relation between the hardness ratio, the X–ray-to-optical flux and the flux in the hard energy band for the Type–1 and Type–2 objects is shown in Fig. 6.9. The two diagrams are in full agreement with the ones presented in Fig. 4.7, with the X–ray-to-optical flux ratio having indeed AGN values $(0.1 < \log(F_X/F_R) < 10.0)$. In the left panel, stars with X–ray emission should occupy the area within the dashed–line.

6.5 Conclusions

6.5.1 The final numbers

The KX–method finally detected 25 quasars in an area of $0.68\,\mathrm{deg}^2$, down to $K_s = 18$. These objects were selected among the candidates with point–like morphology and non–stellar colors. For 8 of them we had a 2dF spectroscopic identification (section 3.2) and were classified as Type–1. The remaining 17 were characterized as QSOs from the features identified in their SEDs, with 16 being of Type–1 and 1 of Type–2.

From the 2dF spectroscopic survey we know that there were five additional Type–1 quasars in the $Rz'K_s$ catalog, which were not detected by KX. This was because the

detection probability of the sources in the K_s–band data was lower than the threshold limit used when generating the K_s–band catalog.

When looking at the properties of point–like sources with stellar colors, we identified 2 additional QSOs. They were selected thanks to their colors in the SWIRE color–color diagram and their quasar nature was revealed by visually inspecting their SEDs.

To conclude, **the K–excess method selected 25 quasars and missed 7, giving a success rate of ∼80%**[2]. From the 25 quasars, 24 are found to be Type–1 and 1 to be Type–2, clearly demonstrating the preference of the method in selecting Type–1 QSOs. Among the KX–selected quasars, however, there are at least four objects with flat SEDs, because of reddening (objects #2, #17, #47 and #57 in Fig. 5.4).

By visually inspecting the spectral energy distribution of sources with extended morphology in three sub–samples (see 6.2.2), we **discovered 32 AGN** (3 of Type–1 and 29 of Type–2). All the above numbers are summarized in Table 6.4.

Method	**Found**		**Missed**	
	Type–1	Type–2	Type–1	Type–2
KX	24	1	7	?
non–KX	5	29	?	?

Table 6.4: Logistics regarding the AGN and quasars contained in the $Rz'K_s$ catalog. The first column describes the approach used for the source identification. The non–KX means that (a) a first selection of QSO/AGN candidates was based on the identification of sub–samples, among the point–like or extended sources, having AGN–like colors and (b) that the verification was based on the visual inspection of the SED of the candidates. The second and third column inform about the nature of the detected objects (Type–1 or Type–2), while the fourth and fifth about nature of the objects that were missed. The "?" symbol means that no value could be obtained, as there is no way of knowing how many objects were missed.

At this point it is important to emphasize *the fact that the AGN population not detected by KX should not be considered as a failure of the method.* The KX proposes the use of infrared criteria to separate quasars from stars, the two classes of objects with point–like morphology in a color–color diagram. The identification of lower–luminosity AGN, which do not look point–like (due to the contribution of the host galaxy), can not be performed using the same concept.

From our study we conclude that indeed the QSO population is well concentrated in the bluer part of the $(R - z')$ axis in the color–color diagram. Setting a color limit in the vertical axis, only, will still keep high the contamination by stars, but will also ensure

[2]This percentage should not be confused with the detection probability of our K_s–selected infrared survey, which is characterized by a completeness of 80% at $K_s = 18$ (Kron magnitude).

that all QSOs present in the catalog can be detected by the KX. A visual inspection of the SEDs of the selected candidates will easily allow to separate the stars and galaxies with QSO–like colors from the quasars, hence the final sample will be free from any contaminants. Such an approach will allow to obtain a higher success rate compared to the one when using color–cuts in both vertical and horizontal axes. However, as for every flux–limited sample, one should keep in mind that there will always be a percentage of detections missing, due to the detection cut–off of the catalog used as a reference (in our case, the K_s–band). Simulations are necessary to estimate the percentage of losses (see section 2.4), in order to finally correct for this effect.

6.5.2 Selection effect corrections

In order to compare the AGN and QSOs found in the survey with the model predictions (number of AGN per square degree, see paragraph 6.2), each sub–sample has first to be corrected for its associated selection effects. The selection of AGN and QSOs has been based on the five sub–samples:

- The KX–selected objects (25 QSOs).
- The point–like sources found in the SWIRE AGN locus (2 QSOs / 3 sources).
- The extended sources with X–ray emission (8 AGN / 33 sources).
- The extended sources found in the SWIRE AGN locus (14 AGN / 27 sources).
- The extended sources with emission in the first three IRAC bands (25 AGN / 774 sources).

Let us first deal with the detected QSOs. The K_s–band catalog is flux limited down to K_s= 18, which corresponds to a completeness of 80%. Since no losses due to color–selection were considered, we can first assume that the KX detects indeed all 27 (25 + 2) QSOs present in the $Rz'K_s$ catalog. Since our selection technique and the subsequent visual inspections were conducted in a total area of 0.68 deg^2, we have to correct this number both for the surface coverage and for the completeness. Therefore, the *number of QSOs per square degree* becomes:

$$27 \;/\; 0.68 \;/\; 0.8 \approx 50$$

Regarding the AGN, we have to consider the following: the 8 AGN found in the X–ray sub–sample were also contained in the sample of objects showing emission at $3.6, 4.5, 5.8$ and $8\,\mu$m found in the AGN locus (the "SWIRE–AGN" sample). Additionally, 7 from the objects showing emission in the first three IRAC bands (the "IRAC–3" sub–sample) were also found to be in the SWIRE–AGN sub–sample. Hence, we can simplify the selection process by assuming that the AGN detections are contained in two sub–samples:

- The SWIRE–AGN, which contains 14 objects, and
- The IRAC–3, which contains 18 objects.

At this point we should stress that the emission in the three IRAC bands is not equivalent to an AGN signature in the selected objects. As the position of these objects can not be traced in the 3.6 to 8.0 μm color–color diagram (see Fig. 4.20), the use of the three IRAC band criterion is the second best criterion for studying the evolution of their SED at infrared wavelengths, with a sufficiently good sampling.

The SWIRE coverage of the LCO field is of the order of 90%. Therefore, the above numbers, when correcting for the overlapping of the two surveys and normalizing the LCO area to a square degree, become:

$$\begin{aligned} \mathrm{IRAC-3:} \;\; & 18/0.9/0.68 \;=\; 29 \;\text{objects} \\ \mathrm{SWIRE-AGN:} \;\; & 14/0.9/0.68 \;=\; 22 \;\text{objects} \end{aligned} \tag{6.2}$$

The requirement of three IRAC bands was met by only 37% of the objects in the catalog. This means that the flux limit of 16.4 μJy for the 8.0 μm band produces a selection bias against the number of detections, for which the IRAC–3 AGN sample has to be corrected. Hence, the number of AGN per square degree contained in the IRAC–3 sub–sample becomes:

$$29 \;/\; 0.37 = 79$$

Similarly, only 26% of the extended sources showed emission in all four IRAC bands. Hence, the number of AGN per square degree, selected among the extended sources with IRAC colors complying with the AGN–locus definition (Stern et al. 2005) becomes:

$$22 \;/\; 0.26 = 88$$

As a consequence, the *total number of AGN found among the sources with extended morphology becomes:*

$$79 + 88 = 167$$

bringing the **total number of QSOs and AGN to 217/deg^2**, which is in very good agreement with the number proposed by the models of Xu et al. keeping, of course, in mind that the models are empirical and only reproduce the observations.

The visual inspection of the AGN SEDs revealed only 3 Type–1 objects out of 32 AGN ($\sim$ 9%). Therefore, the number of Type–1 and Type–2 objects found in the AGN sample is expected to be 16 and 151, respectively. Regarding the KX–selected quasars,

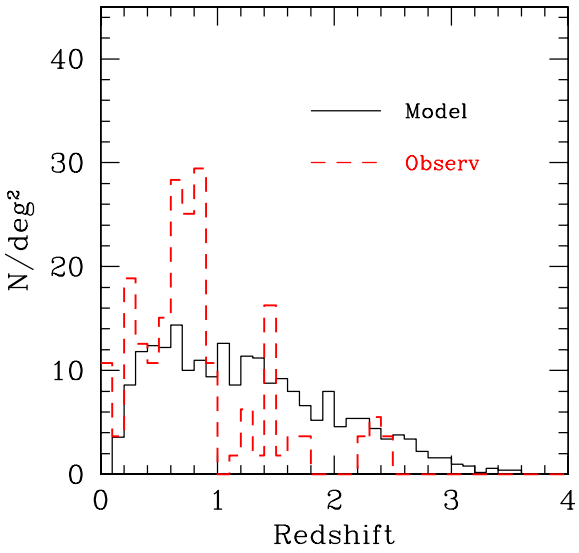

Figure 6.10: Redshift distribution of AGN and QSOs of both types (Type–1, Type–2) predicted by the model of Xu et al. (2006) (solid line). The dashed line corresponds to the photometric redshifts obtained using HyperZ.

only one object out of 25 ended up being Type–2, hence in the K–sample, after correcting for the surface and the completeness we expect to have 48 Type–1 and 2 Type–2 objects. Therefore, the total number of Type–1 and Type–2 AGN and QSOs is expected to be 64 and 153, respectively, giving a Type–2:Type–1 ratio of ≈ 2.4, quite close to the ratio we very often find in the literature (e.g. Lacy et al. (2006)).

The distribution of redshifts, as predicted by the models (solid line) and HyperZ (dashed line), are shown in Fig. 6.10. The observational redshift distribution has been corrected for all selection effects mentioned in the previous paragraphs. From the plot we conclude that the photo–z histogram shows a significant excess in the region $0.6 < z < 1.0$ compared to the model predictions. Our study of the photometric redshifts, presented in paragraphs 5.3.2, 6.2.1 and 6.2.2, did not leave any hints for a bias towards low redshifts. However, thanks to an extended analysis of 14 QSOs with 2dF spectra (see section 5.1.3), where the robustness of the photo–z approach was studied, we concluded that the success of this technique in reproducing the spectroscopic redshifts within a $\Delta z \leq 0.3$ was of the order of $\sim 70\%$. The differences in the two distributions could partially be explained by the limited possibility of the photo–z approach in estimating the redshift of the objects under study.

Although we will make use of the photometric redshifts in the subsequent sections, it is evident that, without spectroscopic redshifts, this study will be subject to uncertainties due to the limitations imposed by the photo–z approach.

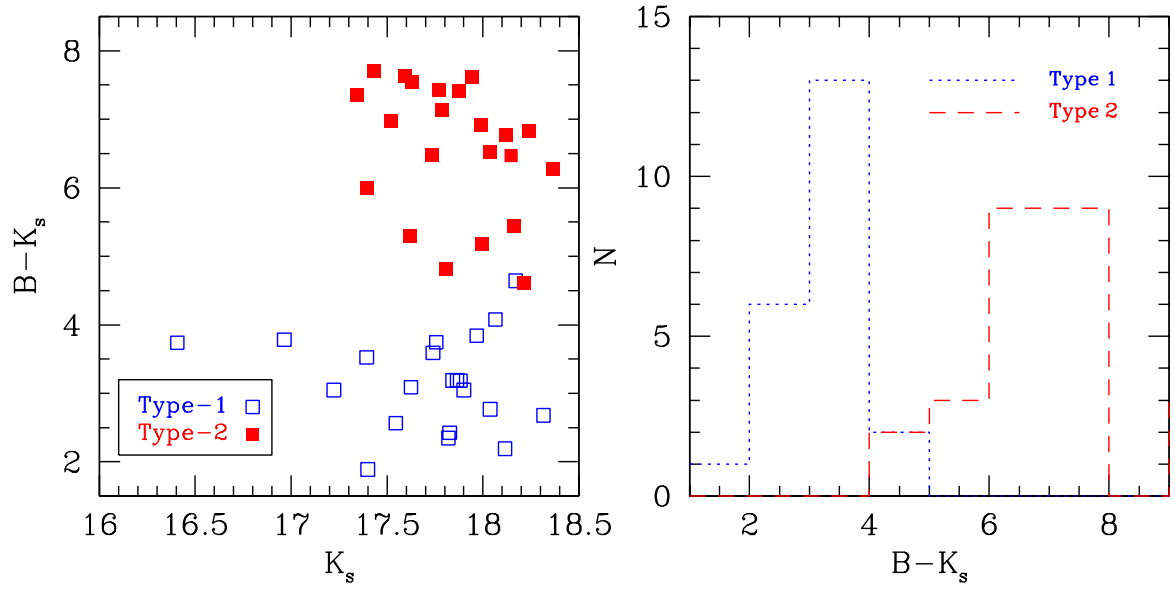

Figure 6.11: *Left panel:* $(B - K_s)$ color index as a function of K_s–band magnitude for the Type–1 (open squares) and Type–2 (filled squares) QSOs and AGN for which it was possible to convert their CFHTLS photometry to B–band magnitudes. *Right panel:* Distribution of the $(B-K_s)$ color indices for the Type–1 (dotted line) and Type–2 (dashed line) AGN and QSOs.

6.5.3 Clear evidence for a red/reddened population

The $(B-K_s)$ color index as a function of K_s–band magnitude, for the QSOs and AGN for which it was possible to convert their CFHTLS photometry to B–band magnitude (see Eq. 4.8), is illustrated in the left panel of Fig. 6.11. Although the $g'-K_s$ color index would also show a similar dependence, since the CFHTLS g'–band filter is close to the Johnson B, we preferred to use the $B-K_s$ color index for reproducing the same color–magnitude diagram as in Webster et al. (1995). This plot is in full agreement with their results, since it clearly demonstrates the presence of a highly reddened population with $(B-K_s > 4.5)$. According to Webster et al. (1995), the reddening could probably be explained by dust in the host galaxy. Nevertheless, the highly reddened population consists of the objects which were classified, on the basis of their spectral energy distribution, as Type–2 objects. A histogram of the $(B - K_s)$ color indices for the Type–1 and Type–2 objects is shown in the right panel of Fig. 6.11.

6.5.4 KX versus optical selection

Webster et al. (1995) were also the first to raise the issue about a missing red/reddened quasar population, due to selection effects. They claimed that "classical" surveys, where

the QSO–selection is based on the UV–excess concept (see section 1.6), could fail to detect up to 80% of quasars due to selection effects. Richards et al. (2003), on a study based on SDSS objects, concluded that the red colors of quasars could be explained by intrinsic processes in the active nucleus and that only $\sim 15\%$ of the objects would be missed due to selection effects. Thanks to the CFHT Legacy Survey covering the XMM–LSS field, which uses the same filters as the SDSS (Fukugita et al. 1996), we can study in more detail how the selection biases affect the "classical" (blue) and red/reddened quasars when using the $u^*g'r'i'z'$ filter combination for detecting quasars.

Fig. 6.12 shows four color–color diagrams, for 46 AGN and quasars with available CFHTLS photometry (open and filled squares for the Type–1 and Type–2 objects, respectively). Overplotted are some 1700 sources with available morphological classification, taken from the CFHTLS data base. The red and green dots represent the point–like sources and the objects with extended morphology, respectively. The first conclusion we draw from these diagrams is that the Type–1 population occupies a well defined region in the bluer part of the plots (in agreement with the classical definition for quasars, which are considered as blue objects). On the contrary, the Type–2 sources are much more dispersed. Even more, it seems to be quite difficult to separate the point–like from the extended sources in a way similar to the $Rz'K_s$ color–color diagram presented in Fig. 3.8. Hence, it is quite improbable for these sources to be selected in a quasar survey using the CFHT photometric bands. The Type–2 population was not selected by the K–excess, neither, clearly showing that each method has its advantages and drawbacks. Therefore, multi–wavelength surveys are essential for minimizing the selection biases from which any survey suffers.

6.6 Summary

Throughout the analysis presented in the precedent sections we performed a full characterization of the selection effects associated to the KX–method. Thanks to the CFHTLS and SWIRE Legacy Surveys associated to the XMM–LSS field, that provided photometric information spanning from $0.37 - 24\,\mu$m, it has been possible to construct the spectral energy distribution of the KX–selected objects. The shape of their SED demonstrated that the K–excess favors the selection of "classical" blue Type–1 quasars, but also QSOs which have redder SED, either because of intrinsic processes or due to dust reddening. Although the KX also detected a quasar with a typical Type–2 shape, the position in the $Rz'K_s$ color–color diagram of the majority of the KX–selected objects (see Fig. 6.6) clearly shows its preference toward Type–1 QSOs.

The initial KX–selected quasar sample, that contained 85 candidates, suffered from a high contamination rate ($\sim 70\%$). This issue was extensively discussed in the current chapter and possible ways of minimizing this effect, without a significant impact on the number of missed quasars, were suggested (see Fig. 6.1). The bottom line is that different

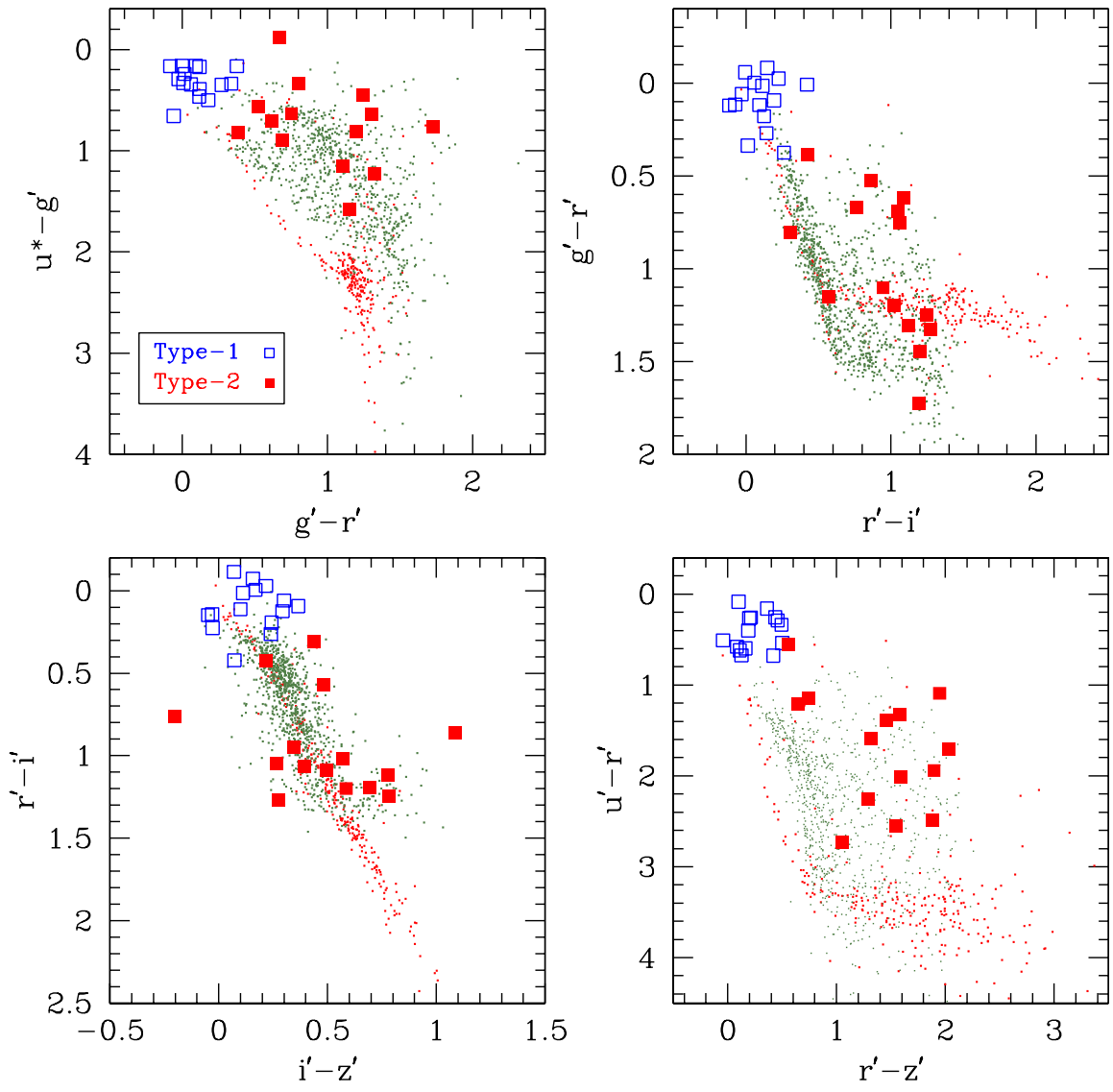

Figure 6.12: Color–color magnitude diagrams for the Type–1 (open squares) and Type–2 (filled squares) quasars and AGN found in the $Rz'K_s$ catalog. Some 1700 objects with morphological classification are plotted as red and green dots (point–like and extended sources, respectively).

color criteria can be implemented, depending on the number of objects one has to deal with (e.g. a relatively shallow or deep survey) and the final use of the selected sample (e.g. a statistical study, spectroscopic observations, etc.). Using stricter color criteria can drastically reduce the number of contaminants, but it is clear that a (probably small) fraction of real quasars will be missed. On the other hand, looser color criteria will result

to a more complete QSO sample, but at the same time a higher degree of contamination. Nevertheless, as we demonstrated in section 5.3.1, the contamination problem can be very well controlled by visualizing the SED of the selected objects, finally allowing to generate a contaminant–free QSO sample.

Some additional QSOs and many AGN, with extended morphology, were found among sub–samples with AGN properties in the X–ray or infrared ($3.6 - 8\,\mu$m) domain. Once more the power of multi–wavelength coverage of the same area sprang out, as this study revealed a few Type–1 and numerous Type–2 AGN that were present in the $Rz'K_s$ catalog, and that the KX–method failed in detecting. As the KX is not designed to detect sources with extended morphology, it is not surprising that none of these objects was included in the initial catalog of QSO candidates.

We have finally succeeded in generating a quite interesting AGN and quasar sample, that contains both Type–1 and Type–2 objects. Their spectral energy distribution spans a wide range of features, from classical quasars with strong emission in the UV–optical to highly obscured ones, where their $0.3 - 1\,\mu$m emission seems to be dramatically attenuated, and from QSOs with a flat–SED to objects with a strong bump when going to IR wavelengths. The objective of the following chapter is to interpret the observed features in terms of physical properties in the interior of the active nucleus and investigate at what extend they can be interpreted in the context of the Unification Scheme.

Chapter 7

IR properties of the AGN sample

The purpose of this chapter is to probe the inner structure of quasars and AGN contained in our sample and more specifically the torus. This will be done by fitting their spectral energy distribution, spanning from UV to mid–IR, with a versatile model that takes under consideration both the diversity in the torus configurations (inner/outer radius, opening angle, viewing angle with respect to the observer) and the complexity of the physical processes taking place in its interior (absorption, scattering and re–emission of the radiation originating from the compact nucleus, type and size of dust grains, etc). Our final objective is to investigate the credibility of the unified scheme, that wants the Type–1 and Type–2 AGN to arise from the same population, with the observed differences being attributed to different angles through which AGN are being viewed.

7.1 Introduction

Although the first study of the infrared properties of NGC 1068 (a prototype Seyfert–2 galaxy) was performed in 1967 (Pacholczyk & Wisniewski 1967), it took astronomers a decade to realize that the $\sim 3.5\,\mu$m excess in the energy distribution of AGN could be attributed to thermal emission due to dust (Rieke (1978), Neugebauer et al. (1979)). Even more, it took another decade before the first models implementing this idea became available (Barvainis 1987).

Modeling the IR properties of AGN is undoubtedly a challenge. At a first level it is essential to fully understand the behavior of dust when found in the vicinity of a radiating source. As dust grains absorb, scatter and re–emit the incoming radiation, and can even become opaque to their proper emission, one needs to solve the non–trivial radiative transfer equation problem. Different computational approaches have intended to tackle this problem (e.g. Pier & Krolik (1992), Granato & Danese (1994)), each one having its own advantages and drawbacks.

Additionally to the complexity of the radiative transfer problem itself, a good knowledge of the dust characteristics is also essential, as the behavior of dust to the incoming radiation depends on the type of grains, their size and their density percentage. Mathis et al. (1977) presented an extended study of the wavelength dependence on the dust extinction, considering that dust principally consists of six materials, among which graphite and silicate. The optical properties of these two main dust components were studied in detail by Draine & Lee (1984).

Through the years, ground–based and space–borne (IRAS, ISO) observations revealed a wealth of information regarding the dust found in astronomical environments, spanning from comets and giant molecular clouds to circumstellar shells and galactic nuclei. Rowan-Robinson (1992), using a graphite–silicate model with dust grain sizes between 0.002–30 μm, managed to provide a successful fit to the observed interstellar extinction curve (0.1−1000 μm), the emissivity of Milky Way at hight galactic latitudes and the 12−100 μm spectra of galaxies observed by IRAS.

The geometry of the torus and the resolution of the grid used for studying the radiative transfer problem are the last two of the main parameters controlling the efficiency of the models in reproducing the infrared emission of active galactic nuclei. See Granato & Danese (1994), Stenholm (1994), Efstathiou & Rowan-Robinson (1995) and Nenkova et al. (2002) for different versions of the proposed models.

7.2 The torus model

An exhaustive description of the torus model is given in Fritz et al. (2006). For a detailed analysis of the fitting method, the definition of the torus properties and results on a sample of spectroscopically confirmed quasars, see Hatziminaoglou et al. (in prep). Here we only briefly outline the most important of the model's characteristics. The torus geometry adopted is that of a *flared disk*, that is a sphere with the polar cones removed. Its size is defined by its outer radius, R_{out}, and the opening angle of the torus itself. The dust components that dominate both the absorption and the emission of radiation are graphite and silicate. The location of the inner radius, R_{in}, depends both on the sublimation temperature of the dust grains and on the strength of the accretion luminosity. A schematic representation of the torus geometry is presented in Fig. 7.1.

In their model, Fritz et al. (2006) considered that the sublimation temperature for the graphite and silicate grains is 1500 and 1000 K, respectively. Although higher values have also been proposed (1700 and 1450 K, Absil et al. (2006)), according to Granato & Danese (1994) a 300 K raise in the sublimation temperature of the grains yields fluxes that are higher by $\sim 30\%$ at 1 μm, whereas at $\lambda \geq 3\,\mu$m the IR continua are practically unchanged.

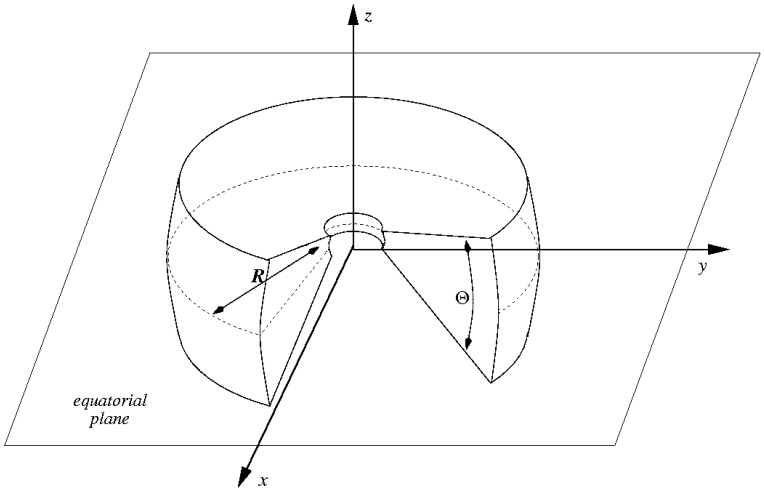

Figure 7.1: Schematic representation of the torus configuration, considered by Fritz et al. (2006). R corresponds to the ratio of the external to internal radius and Θ is the full–opening angle of the torus.

Let us now go through the most substantial physical properties of the torus:

- The central source is assumed to be point–like and its emission isotropic. Its luminosity is considered to be constant, with $L = 10^{46}\,\mathrm{erg\,s^{-1}}$, while its spectral energy distribution is defined by means of a composition of power laws, with different values for the spectral index in the UV and optical (see e.g. Granato & Danese (1994)).

- The main dust components are graphite and silicate grains, with approximately equal fractions. Their sizes are varying, with the minimum being $0.005\,\mu$m and $0.025\,\mu$m for the graphite and silicate, respectively, while the maximum size is $0.25\,\mu$m for both. The silicate grains are responsible for the absorption feature at $\sim 9.7\mu$m in Type–2 objects, while the graphite grains account for the strong decline of the emission at wavelengths shortward of a few μm, which corresponds to a blackbody emitting at the sublimation temperature of these grains ($\sim 1500\,\mathrm{K}$).

- The radiative transfer equation is solved using a numerical approach, the so–called Λ–iteration method.

- A geometrical grid is defined along the three spatial coordinates and the main physical quantities (dust density and temperatures, electromagnetic emission, optical depth, etc.) are computed with respect to the center of the volume elements defined by the grid. The gas density within the torus is modeled in such a way to allow a gradient along both the radial and the angular coordinates.

- When computing the total incoming energy on a given volume element, a very accurate computation of the optical depth is done. Furthermore, since silicate grains

have a lower sublimation temperature with respect to graphite, the fact that the innermost regions of the torus contain only graphite and are silicate free is accounted for. Finally, the global SED is computed for different viewing angles with respect to the torus' equatorial plane, in order to account for the emission from Type–1 and Type–2 objects. It includes three contributions: emission from the AGN, thermal emission and scattering emission by dust in each volume element.

7.3 The 9.7 μm silicate feature

Absorption features at infrared wavelengths result from molecular vibrations (i.e. rotation, stretch or bend) withing the grain material. Although the features observed in infrared spectra are numerous, in this section we will limit ourselves to the properties of the 9.7 μm feature, because of its important role in the Fritz et al. (2006) model. An excellent introduction to the properties of dust and the observed features at infrared wavelengths can be found in Whittet (2003).

The 9.7 μm silicate feature is the strongest and best studied infrared feature in the diffuse Inter–Stellar Medium. The term "silicate" is rather general, covering a wide range of possible combinations of mineral structures, based on SiO_4 units. These units have a tetrahedral shape and negative charge, as shown in Fig. 7.2. The oxygen atoms may serve as a node between adjacent units, or may bond to positively charged cations (usually Mg^{2+} or Fe^{2+}), forming chains of three–dimensional structures. In silicates of the olivine group (e.g. Mg_2SiO_4), the SiO_4 units are isolated from each other by cations, whereas in other groups, such as pyroxenes (e.g. $MgSiO_3$), oxygen atoms act as a bridge between them.

The 9.7 μm silicate feature is generated by stretching modes in the silicate tetrahedra. Depending on its crystallinity, i.e. the degree of long–range order of the tetrahedra, silicate's spectroscopic properties vary: amorphous silicate will produce a smooth, broad feature, whereas crystalline silicate will produce profiles with sharp, narrow structure. The profile of the 9.7 μm feature observed in astronomical spectra is generally broad and smooth, indicating that the silicate is found in an amorphous state, rather than crystalline. An example of the emissivity as a function of wavelength, for three types of silicate found in the disks around Herbig Ae stars, is presented in Fig. 7.3 (van Boekel et al. 2004).

Depending on the conditions of the medium where the silicate grains are found, the 9.7 μm feature can be seen either in emission, or absorption. The former case happens when the medium, the torus in our case, has a low optical depth. In an optically thick torus, two cases have to be distinguished: if the temperature decreases along the line of sight, the 9.7 μm feature appears as an absorption, otherwise it is seen in emission. At an optical depth in between, i.e. $\tau(9.7) \approx 1$, a mixture of both is visible (Schartmann et al. (2005) and references therein). As a result, for Type–2 AGN, we expect to see the 9.7 μm silicate feature in absorption, whilst for Type–1 objects in emission.

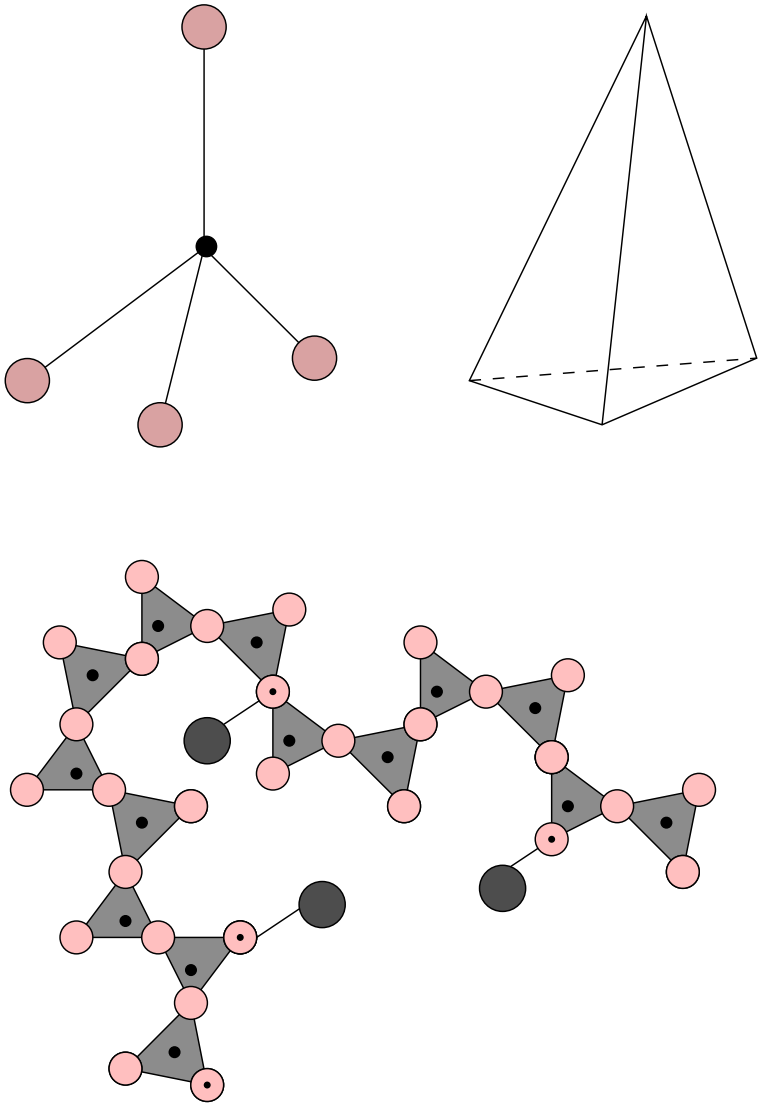

Figure 7.2: Schematic representation of the silicate building blocks. *Top:* The SiO_4 unit (left) and its tetrahedral structure (right). The Si atom (black circle) is located at the tetrahedron and O atoms (light–colored circles) occupy its corners. *Bottom*: Two dimensional representation of a segment of amorphous silicate. The triangles correspond to SiO_4^{2-} anions and the dark–colored large circles to cations (Mg^{2+} or Fe^{2+}). Oxygen atoms (small light–colored circles) serve as nodes between adjacent tetrahedra whilst others are attached to cations (dotted circles). Reproduction of Fig. 1.7 found in Whittet (2003).

The presence of this absorption feature in Seyfert–2 spectra has been indeed used as a proof for the presence of silicate in the AGN environment. At the same time, its absence from spectra of Type–1 Seyfert galaxies has challenged the credibility of the unified scheme model. Nenkova et al. (2002) and Dullemond & van Bemmel (2005) have

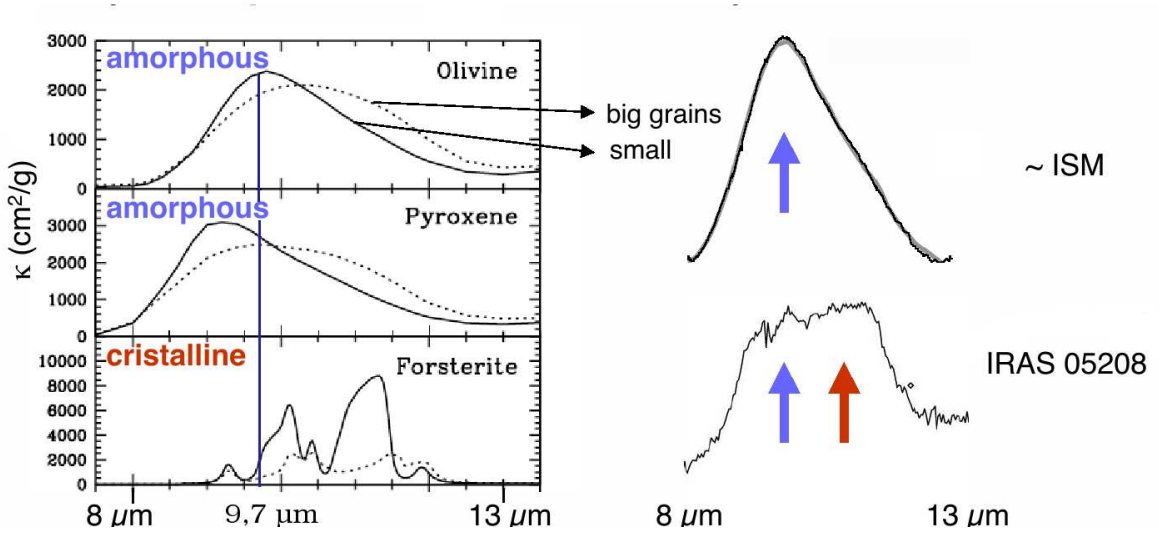

Figure 7.3: *Left panel:* Infrared spectrum of the emissivity ($\mathrm{cm^2\,g^{-1}}$) for amorphous (Olivine, Pyroxene) and crystalline (Forsterite, $MgFe_2SiO_4$) silicate grains, detected in the disk around the Herbig Ae star with identification HD 142527. The striking difference in the emissivity of the amorphous to crystalline silicate permits to identify the nature of the dust grains (van Boekel et al. 2004). *Right panel:* Comparison between the normalized spectrum of the interstellar medium (upper) and that of IRAS 05208−2035 (lower). In the latter spectrum, the contribution of both amorphous (left arrow) and crystalline (right arrow) silicate grains are visible (de Ruyter et al. 2006).

suggested different explanations for the lack of the 9.7 μm absorption feature in Type–1 objects. Only recently Siebenmorgen et al. (2005) and Hao et al. (2005) have reported on the detection of the 9.7 μm feature, in emission, in Type–1 quasars, using infrared spectra taken with the Spitzer satellite.

As the 9.7 μm feature has been so well studied, it can be used to probe the properties of a dusty environment, by correlating the optical depth at 9.7 μm $\tau(9.7)$, to the optical extinction, A_V. For example, for early–type stars, which lie relatively close to the Sun (mostly within $\sim$ 3kpc), there is a clear correlation, consistent with a straight line through the origin (Whittet 2003):

$$\frac{A_V}{\tau(9.7)} = 18.0 \pm 1.0 \tag{7.1}$$

For very small pure silicate particles (of about $\sim 0.01\,\mu$m) this ratio becomes ~ 1, while for classical ($\sim 0.15\,\mu$m) grains it is ~ 5 (Stephens 1980). In the Fritz et al. (2006) model, the adopted values for the specific optical depth are $\tau(9.7) = 0.1, 0.3, 0.6, 1.0, 2.0, 3.0, 6.0$ and 10.0.

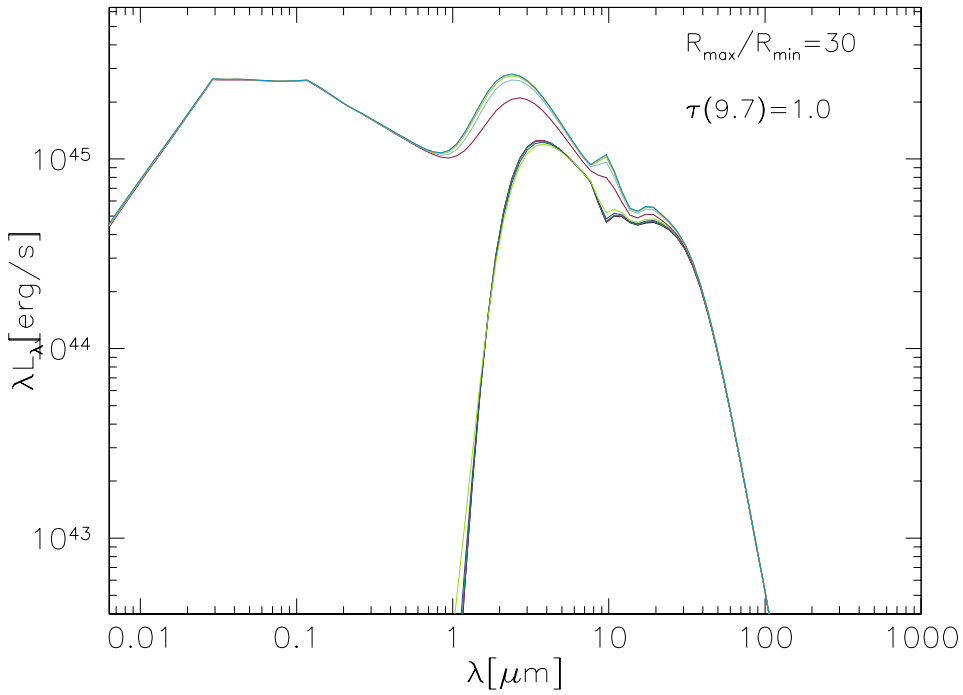

Figure 7.4: Emission spectra obtained for 10 different line–of–sight inclinations from 0° to 90° at regular steps of 10°. Here the ratio between the outer and inner radii is $R_{max}/R_{min} = 30$, the density decreases with the square root of the distance to the center and the equatorial optical depth is $\tau(9.7) = 1.0$ (from J. Fritz, private communication).

7.4 SED fitting results: physical properties of the AGN candidates

An example of an emitted spectrum is illustrated in Fig. 7.4. The model has been computed for a half an opening angle of 40 degrees, an equatorial optical depth at 9.7 μm, $\tau(9.7) = 1.0$ and a dust density decreasing with the square root of the distance to the center. The ratio of the outer–to–inner radii is equal to 30. All lines of sight are seen in the plot, starting with a face–on view (Type–1 object, upper curves) down to an edge–on view (Type–2 objects, lower curves). The abrupt change of shape of the SEDs reflects the interception of the torus by the line of sight. We refer to Fritz et al. (2006) for a detailed description of the model developed to compute the emission of dust in AGN.

Each observed SED is compared to a total of 493 models. For Type–1 objects we use lines of sight with the torus not intercepting, while for Type–2 objects we consider only lines of sight that do not see the central engine directly. For an exhaustive description of the methods, the calculation of the properties and the degeneracy issues see the study of SWIRE–SDSS quasars by Hatziminaoglou et al. (in prep).

7.4.1 Torus density and the 9.7 μm optical depth

In the model by Fritz et al. (2006), the density dependence along the radial and polar direction is described by the following law:

$$\rho(r,\theta) \propto r^{\beta} e^{-\gamma|cos(\theta)|} \tag{7.2}$$

where the parameters β, γ account for the spatial gradient and θ for the polar angle. The parameters β and γ take discrete values, with $\beta = 0.5, 0, -0.5$ and -1.0 and $\gamma = 0$ and 6. For both Type–1 and Type–2 AGN and QSOs the vast majority of the objects was better fit by models with constant ($\beta = 0$) or decreasing ($\beta = -0.5, -1.0$) dust density, in the radial direction. Almost half of them ($\sim 47\%$) were better represented by models with the dust density decreasing with the angle from the equator ($\gamma = 6$) while the remaining objects were better matched by a model where the dust density was independent of the angle ($\gamma = 0$). Some 37% of the observed SEDs agreed with models of low equatorial optical depth at 9.7 μm ($\tau(9.7) \leq 1$), while the rest were better reproduced by models with higher optical depth. All these results are in excellent agreement with the findings of Hatziminaoglou (2006). A histogram of the β, γ factors and the $\tau(9.7)$ optical depth is shown in Fig. 7.5.

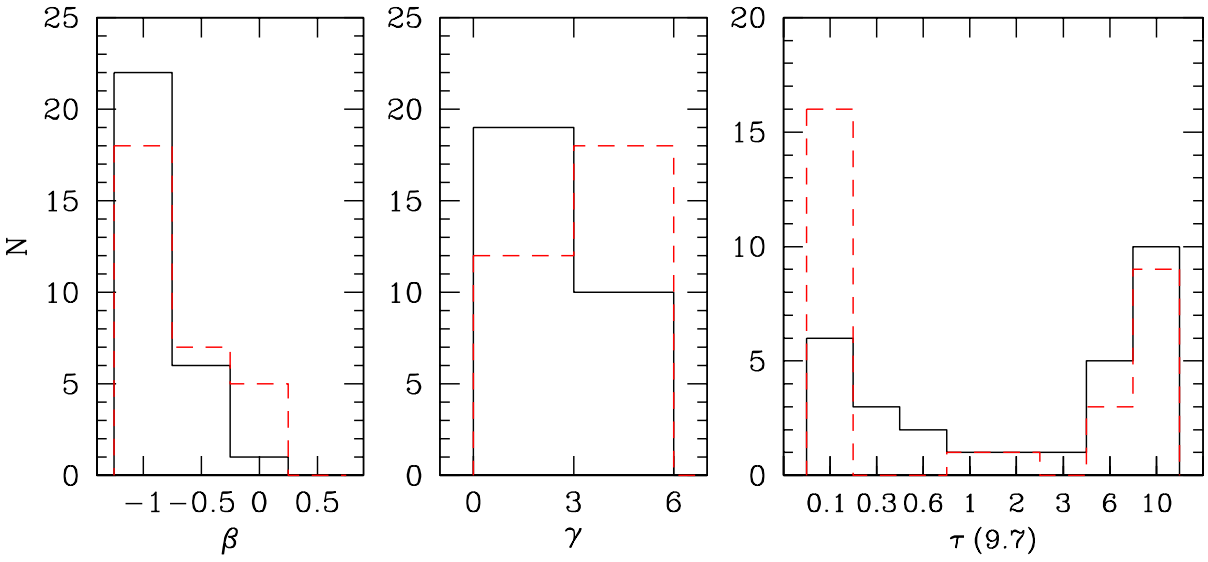

Figure 7.5: *Left Panel:* Histogram of the β factor, which describes the density gradient along the radial direction, for the 59 Type–1 (solid line) and Type–2 (dashed line) AGN and QSOs detected in the Las Campanas Survey. *Middle Panel:* Histogram of the γ factor, which accounts for the density gradient as a function of angle from the equator. *Right Panel:* Histogram of the optical depth at 9.7 μm.

7.4.2 Geometrical Properties

The accretion luminosity, L_{acc}, ranges typically from 10^{45} to 10^{47} erg/sec and both Type–1 and Type–2 objects span the same luminosity range. The inner radii, R_{in}, are of the order of a few pc and both populations take similar values.

The inner radius depends on the sublimation temperature of the dust components, as in very hot environments, where the temperature exceeds their sublimation temperature, the dust grains evaporate. The conditions in the vicinity of the active nucleus depend on the rate at which matter accretes toward the black hole. According to Barvainis (1987), the size of the inner radius is directly proportional to the accretion luminosity, via the relation:

$$R_{\rm in} \simeq 1.3 \times \sqrt{L}_{\rm acc} \times T_{1500}^{-2.8} \quad [\rm pc] \tag{7.3}$$

where the temperature, T, is given in units of 1500 K. In a flux–limited sample, like the one under study, the accretion luminosity L_{acc} (which is directly related to the bolometric luminosity) increases with redshift and therefore so does R_{in}. The above results are schematically summarized in Fig. 7.6.

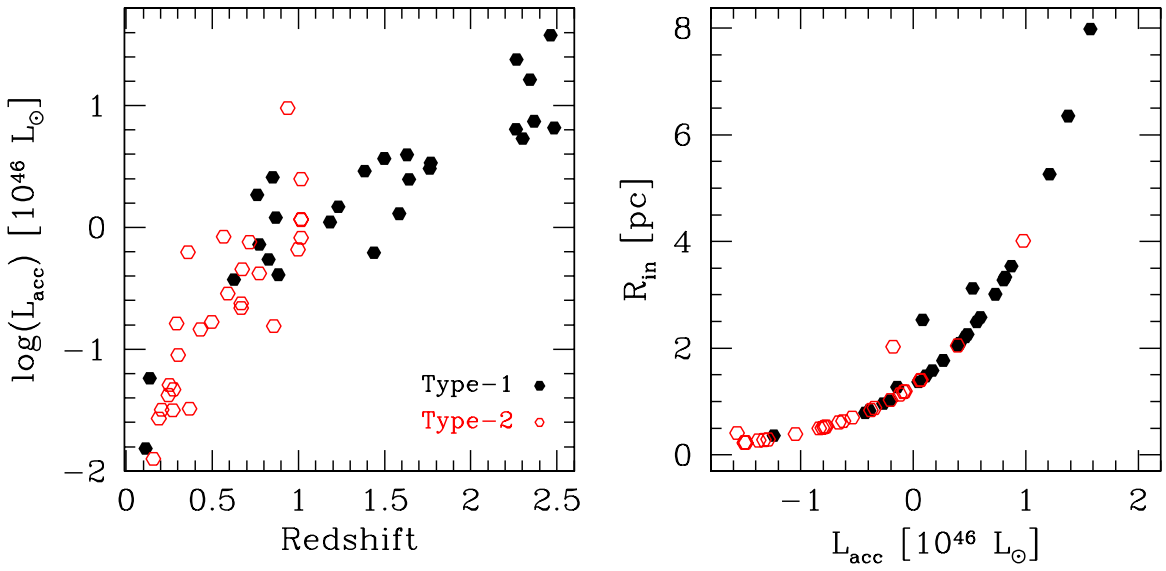

Figure 7.6: *Left panel:* Redshift dependence of the accretion luminosity (log. scale), in units of $10^{46} L_{\odot}$, for the Type–1 (filled circles) and Type–2 (open circles) AGN and QSOs. *Right panel:* Inner radius (R_{in}) dependence on the accretion luminosity.

In the right panel of Fig. 7.6 we clearly see three outliers, with an inner Radius higher than the expected value, for a given accretion luminosity (see also Eq. (7.3)). A possible explanation could be that the computed temperature at R_{in}, when calculated using Eq. 7.3, reaches values higher than the sublimation temperature and in these cases

R_{in} is pushed outwards and the temperature is recomputed. This usually happens to models with high optical depth, where the dust itself becomes an important source of heat. Although a wrong photometric redshift could also explain the outliers in Fig. 7.6, as the accretion luminosity depends on the object's redshift, the fact that for objects with known spectroscopic redshift similar inconsistencies have been observed (Hatziminaoglou et al., in preparation) gives more credit to a physical explanation.

A histogram of the inner (R_{in}) and outer (R_{out}) radius is presented in the upper left and upper right panel of Fig. 7.7, respectively. In the models used, the ratio R_{out}/R_{in} has been set to take two values, 30 and 100. Although Fritz et al. (2006) mention a third value of $R_{out}/R_{in} = 300$, we did not use this ratio when modeling our objects, as in this way we avoid possible model–related degeneracy issues. Very large tori imply even lower temperatures, hence with no coverage in the far–infrared and the submillimeter it is not possible to proper model their spectral energy distribution.

The distribution of the R_{out}/R_{in} is shown in the lower right panel of Fig. 7.7. A histogram of the covering factors (related to the torus opening angle), is presented in the lower left panel. 86% and 83% of the Type–1 and Type–2 objects, respectively, are having a covering factor $\geq$ 50%. The percentage of the entire sample that has a covering factor $\geq$ 90% is $\sim$40%.

7.4.3 Dust mass and column density

The mass of dust spans from some tens solar masses up to $\sim 10^6$ $M_\odot$. Note that this is the mass of the silicate and graphite grains only and not the mass contained in the entire torus. To properly calculate the latter, one has to take into account the mass of gas, typically 100 times larger than that of the dust. The distribution of the dust mass (in logarithmic scale) for the Type–1 and Type–2 AGN is shown in the left panel of Fig. 7.8. From the plot we induce that Type–2 objects seem to be less dusty, with the peak of the mass distribution found at $\sim$ 3.5, while for the Type–1 objects found at about $\sim$ 4.5. If this is a selection bias, then for some reason –which is difficult to identify– we are selecting less luminous (in the IR) Type–2 AGN. If it's not a selection bias, this would imply that Type–2 objects have lower dust masses, which is not in very good agreement with the Unification Scheme. Another possible explanation found in the literature (e.g. Granato & Danese (1994)) could be that the dust density increases proportionally to the distance from the center. Our findings, however, suggest that such models are probably not physical.

The *equatorial* hydrogen column density takes values between $N_H = 10^{21.93}$ and $10^{23.94} cm^{-2}$ and spans the same range for both types of objects, with some 35% of Type–1 objects and $\sim$ 30% of Type–2 being very close to the so–called Compton thick objects ($N_H = 10^{24}$). The hydrogen column density along the line of sight (LoS) was computed only for the Type–2 objects only – since Type–1 are not seen through the torus – taking values between $10^{21.46}$ and 10^{23}. A histogram of the column density for the Type–1 and

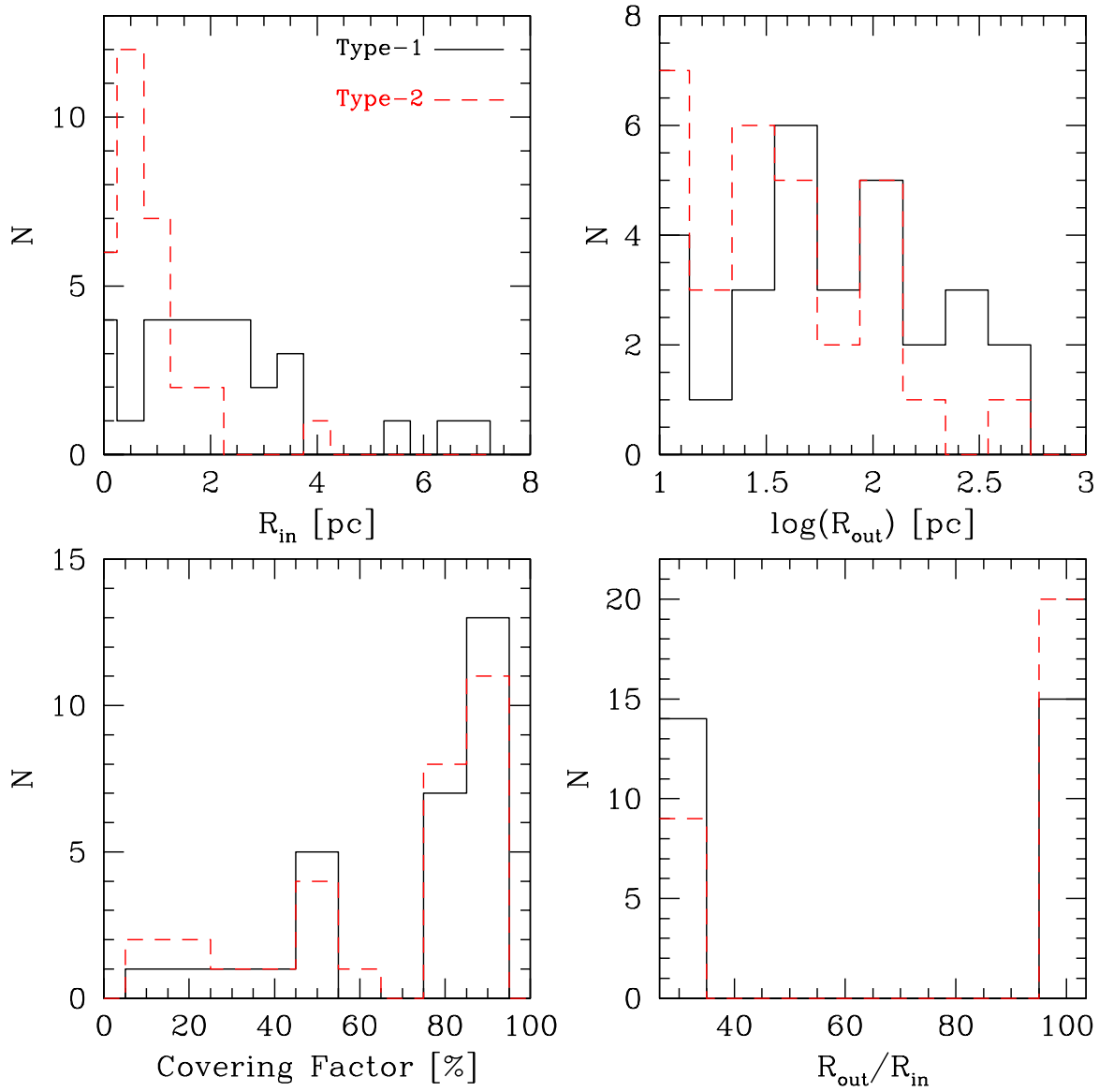

Figure 7.7: *Upper Left:* Histogram of the inner radius R_{in} (in units of pc), for the Type–1 (solid line) and Type–2 (dashed line) AGN and QSOs. *Upper Right:* Histogram of the outer radius R_{out} (in logarithmic scale). *Lower Left :* Histogram of the covering factors. *Lower Right:* Histogram of the ratio R_{out}/R_{in}.

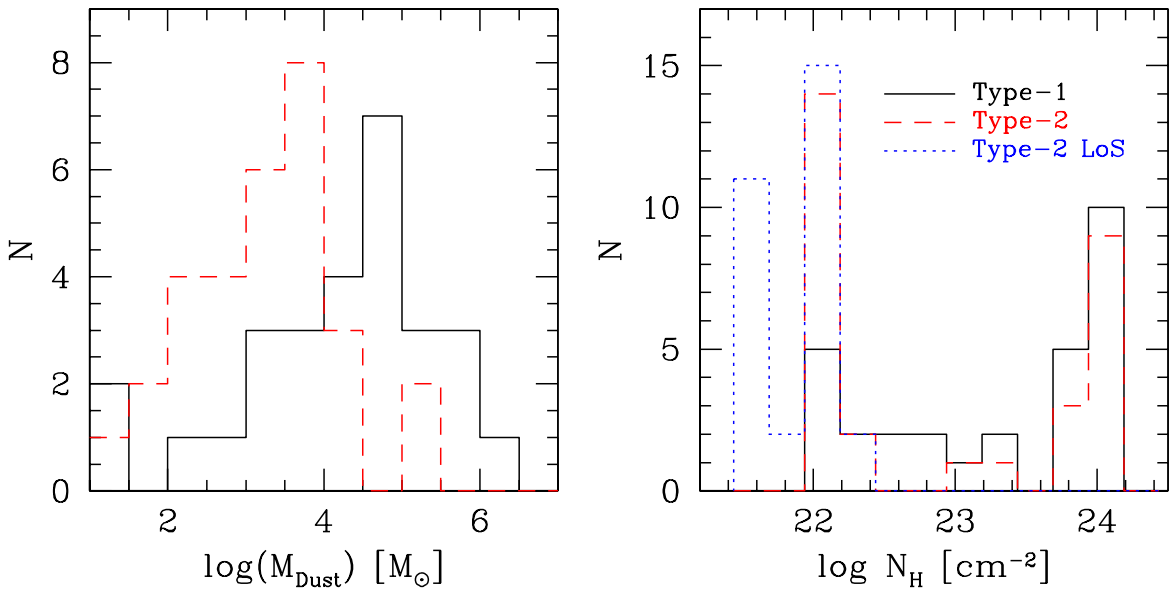

Figure 7.8: *Left:* Histogram of the mass of the silicate and graphite grains for the Type–1 and Type–2 objects, in unit of solar masses (in logarithmic scale). *Right:* Histogram of the equatorial hydrogen column density (N_H) (in logarithmic scale) for the Type–1 (solid line) and Type–2 (dashed line) AGN and QSOs. The dotted line presents the histogram of the line–of–sight (LoS) column density, for the Type–2 objects.

Type–2 objects is presented in the right panel of Fig. 7.8.

7.4.4 The Unified Model

Our analysis has demonstrated that in both Type–1 and Type–2 objects, the tori have the same geometrical (inner, outer radius and opening angle) and physical (density gradient, equatorial column density) properties. This conclusion is in full agreement with the prediction of the Unification Scheme (Antonucci 1993) that wants the two types of AGN to be a result of orientation.

A histogram of the viewing angles through which the different objects are viewed is presented in Fig. 7.9. At this point it is important to emphasize that we have *a priori* constrained the range of viewing angles through which we see Type–1 and Type–2 objects, using the 60° from the equatorial plane as the limit that differentiates the two populations. Such a constraint was necessary, in order to avoid degeneracies in the fitting procedure. Available models in the literature use tori characterized by a high optical depth, hence do not necessary need such a constraint. Nevertheless, as the Fritz et al. (2006) model considers optical depths that span the range $\tau(9.7) = 0.1 - 10$, such an assumption is necessary to avoid models without physical meaning (e.g. tori with a R_{out}/R_{in} ratio of

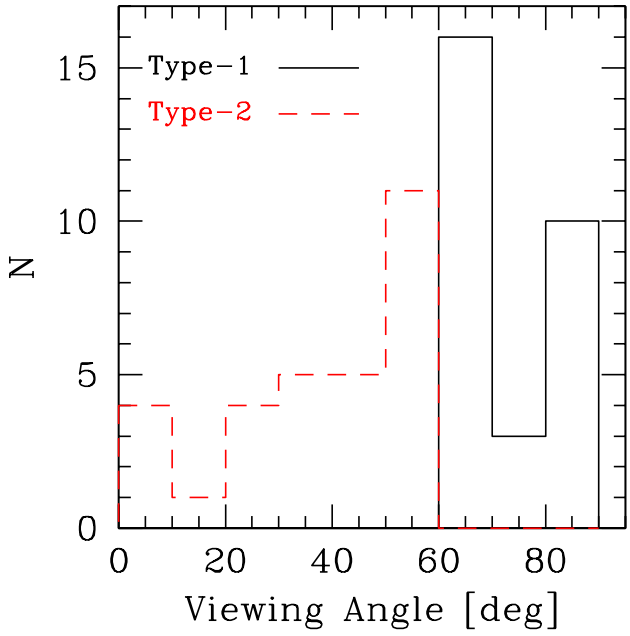

Figure 7.9: Viewing Angle histogram for the Type–1 (solid line) and Type–2 (dashed line) population of the Las Campanas sample.

300 having a low optical depth, that would need information at longer wavelengths for constraining their emission and the contribution from a starburst component). The clear distinction between the two distributions is obvious, suggesting that indeed the significant differences in the spectral energy distribution of the Type–1 and Type–2 objects, presented in Figs 5.5 6.3 and 6.5, can be explained in terms of orientation.

The main properties of the torus models used for fitting the spectral energy distribution of the Type–1 AGN and quasars are summarized in Table 7.1. The best–fit models, together with the observational points used for this purpose, i.e. the CFHT u^*, g'–bands, the R, z' and K_s–band photometry and the five Spitzer bands (spanning from 3.6 to 24 μm), are presented in Fig. 7.10. As $\sim 20\%$ of our objects lacked CFHTLS photometry, we preferred to use the CTIO R and z'–band measurements, instead of the CFHTLS r' and z' ones, when fitting the torus models. When CFHTLS data were available, the u^*, g'–bands were also used. The objects are plotted in the same order as in Table 7.1. Above each cell the object's redshift is also shown.

Table 7.2 presents the results for the Type–2 AGN and single QSO (#33) found in our survey. The corresponding plots are shown in Fig. 7.10 and follow the same order as in Table 7.2.

ID	z	R_{in} (pc)	R_{out} (pc)	Θ (°)	$\tau(9.7)$	$Mass_{dust}$ ($M_\odot$)	A_V (mag)	$\log(N_H)$ cm^{-2}	V.A. (°)
121	0.114	0.16	3.21	140	10.0	1.2×10^2	214.69	23.94	90
1	0.137	0.36	36.04	140	10.0	2.0×10^4	214.05	23.94	70
4	0.110	0.09	8.86	140	0.1	6.9×10^0	2.34	21.99	80
2	0.885	0.83	24.94	140	10.0	3.8×10^4	214.32	23.94	80
3	1.762	2.26	226.46	100	6.0	1.8×10^6	128.56	23.72	70
17	2.265	6.35	126.95	140	3.0	4.1×10^5	65.85	23.43	70
19	2.301	3.01	60.20	100	0.1	3.1×10^3	2.84	22.11	90
20	0.774	1.27	127.49	140	10.0	2.5×10^5	214.05	23.94	90
27	1.439	1.02	20.48	100	0.3	7.4×10^3	7.44	22.50	90
37	0.850	2.09	208.96	140	0.1	2.7×10^4	2.34	21.99	90
45	1.232	1.58	31.63	60	10.0	3.5×10^4	214.50	23.94	80
47	2.367	3.54	354.37	140	0.1	7.8×10^4	2.32	21.99	70
57	0.110	0.15	3.09	140	10.0	8.9×10^2	214.22	23.94	90
78	1.768	3.12	62.44	100	6.0	2.2×10^5	129.32	23.72	70
80	1.641	2.05	40.95	100	0.6	6.9×10^3	14.44	22.79	70
81	0.870	2.53	50.55	100	10.0	2.4×10^5	213.70	23.94	70
83	0.163	0.10	2.04	140	0.1	1.2×10^0	3.36	22.20	70
84	1.584	1.48	44.39	60	0.3	1.8×10^3	7.18	22.48	90
85	0.625	0.79	23.85	140	10.0	4.8×10^4	214.19	23.94	70
735	0.760	1.77	52.98	60	0.1	1.1×10^3	2.32	21.99	70
736	1.629	2.58	258.19	60	0.3	7.2×10^4	6.53	22.43	70
97	0.826	0.96	95.87	60	0.6	2.0×10^4	13.05	22.73	70
100	2.463	7.98	159.67	100	6.0	9.1×10^5	129.74	23.72	90
251	1.497	2.49	49.87	140	10.0	2.9×10^4	214.69	23.94	90
41	2.483	3.33	332.59	100	1.0	8.9×10^4	22.27	22.96	70
11	1.183	1.37	136.81	140	6.0	9.3×10^5	128.49	23.72	70
14	2.263	3.28	98.53	140	10.0	8.3×10^4	214.48	23.94	90
S1	1.384	2.21	66.39	100	2.0	7.8×10^3	44.26	23.26	70
S2	2.344	5.26	525.62	140	6.0	1.7×10^7	127.84	23.71	70

Table 7.1: Summary of the main properties of the torus models used to fit the spectral energy distribution of the Type–1 AGN and QSOs. Columns 1 and 2 refer to the object's identification and its redshift, respectively. Columns 3 and 4 to the inner and outer radius of the torus, in parsec units. Column 5, 6 and 7 give the full opening angle of the torus (in degrees), its optical depth measured at 9.7 μm and the mass of the dust (in solar mass units), respectively. Columns 8 and 9 give the optical extinction (in magnitude) and the column density, at the equatorial plane. Finally, column 10 gives the viewing angle (in degrees), measured from the equatorial plane, from which the object is seen.

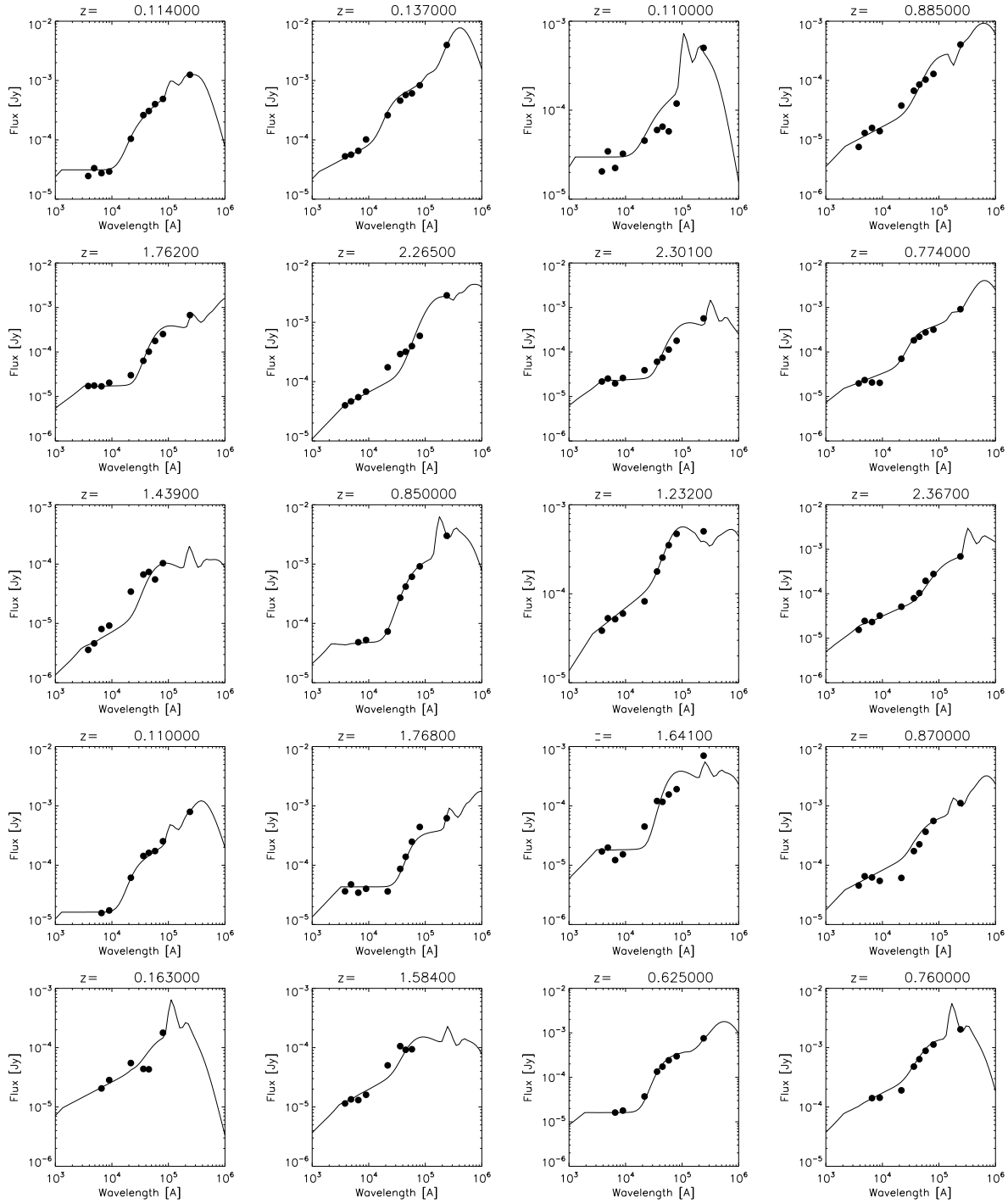

Figure 7.10: Comparison of the observational data (filled circles) and the torus model (solid line) for the Type–1 AGN and QSOs discovered in the Las Campanas survey.

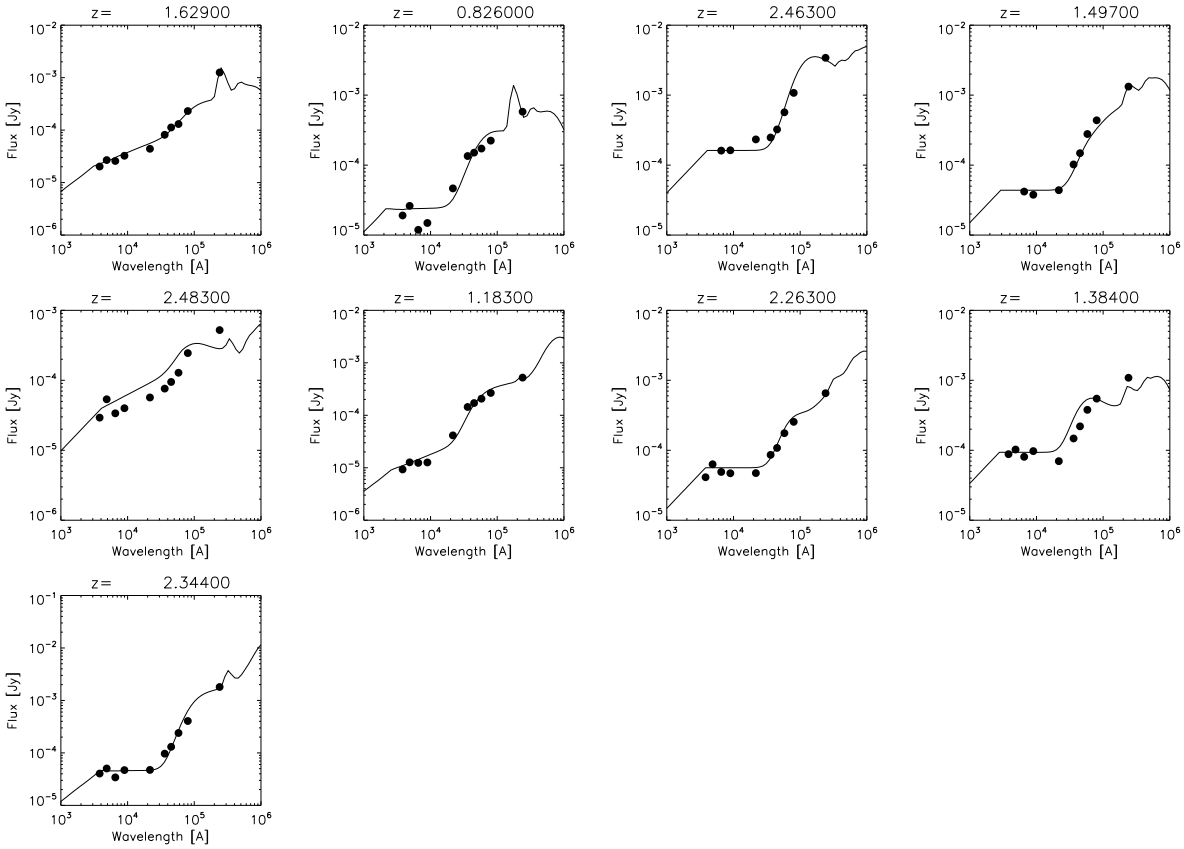

Figure 7.10: Continued

ID	z	R_{in} (pc)	R_{out} (pc)	Θ (°)	$\tau(9.7)$	$Mass_{dust}$ ($M_\odot$)	A_V (mag)	$\log(N_H)$ cm^{-2}	V.A. (°)
86	1.017	1.41	28.11	140	10.0	9.1×10^3	214.69	23.94	60
117	0.591	0.70	13.93	80	10.0	2.3×10^3	214.24	23.94	50
122	0.276	0.28	8.45	140	10.0	6.0×10^2	214.48	23.94	60
123	0.857	0.51	15.35	100	1.0	2.1×10^2	22.64	22.97	30
129	1.017	2.05	205.24	100	0.1	3.2×10^4	2.14	21.94	40
137	0.159	0.19	3.82	140	10.0	1.4×10^3	214.22	23.94	60
206	0.496	0.53	53.14	140	10.0	2.2×10^4	214.34	23.94	60
291	0.668	0.63	12.70	100	0.1	3.3×10^1	3.36	22.20	0
304	0.673	0.88	87.55	100	0.1	5.8×10^3	2.14	21.94	30
415	0.999	2.03	40.63	140	10.0	1.5×10^5	213.60	23.94	60
419	0.271	0.23	23.08	140	10.0	4.1×10^3	214.34	23.94	50
433	0.716	1.13	113.17	140	0.1	1.3×10^4	2.14	21.94	60
435	1.017	1.40	27.92	100	0.1	1.6×10^2	3.36	22.20	0
512	0.367	0.23	23.40	100	2.0	8.7×10^2	43.84	23.25	30
566	0.668	0.61	18.23	140	10.0	2.8×10^3	214.48	23.94	60
592	0.772	0.84	84.12	100	0.1	4.4×10^3	2.18	21.95	40
628	0.114	0.07	2.16	140	6.0	2.4×10^1	129.53	23.72	50
769	1.017	1.18	117.80	140	0.1	1.2×10^3	2.32	21.99	0
770	0.114	0.10	3.01	140	6.0	4.7×10^1	129.53	23.72	50
7	0.208	0.23	23.25	100	0.1	3.3×10^2	2.19	21.95	50
11	0.566	1.19	119.22	100	0.1	6.3×10^3	2.34	21.99	40
4	0.360	1.03	102.82	140	0.1	9.1×10^3	2.18	21.95	60
6	0.190	0.41	8.22	140	10.0	6.3×10^3	213.60	23.94	60
9	0.431	0.50	49.62	100	0.1	1.5×10^3	2.18	21.95	40
16	0.294	0.52	52.41	140	0.1	2.3×10^3	2.18	21.95	60
23	0.302	0.39	38.89	60	0.1	1.2×10^2	2.34	21.99	0
25	0.938	4.01	401.45	100	0.1	1.2×10^5	2.14	21.94	40
27	0.252	0.29	29.39	60	0.1	3.2×10^2	2.18	21.95	20
28	0.245	0.27	8.00	140	6.0	3.3×10^2	129.53	23.72	60
33	0.556	0.79	78.67	100	0.1	4.7×10^3	2.14	21.94	30

Table 7.2: Summary of the main properties of the torus models used to fit the spectral energy distribution of the Type–2 AGN and QSOs (object #33). Same column designation as in Table 7.1

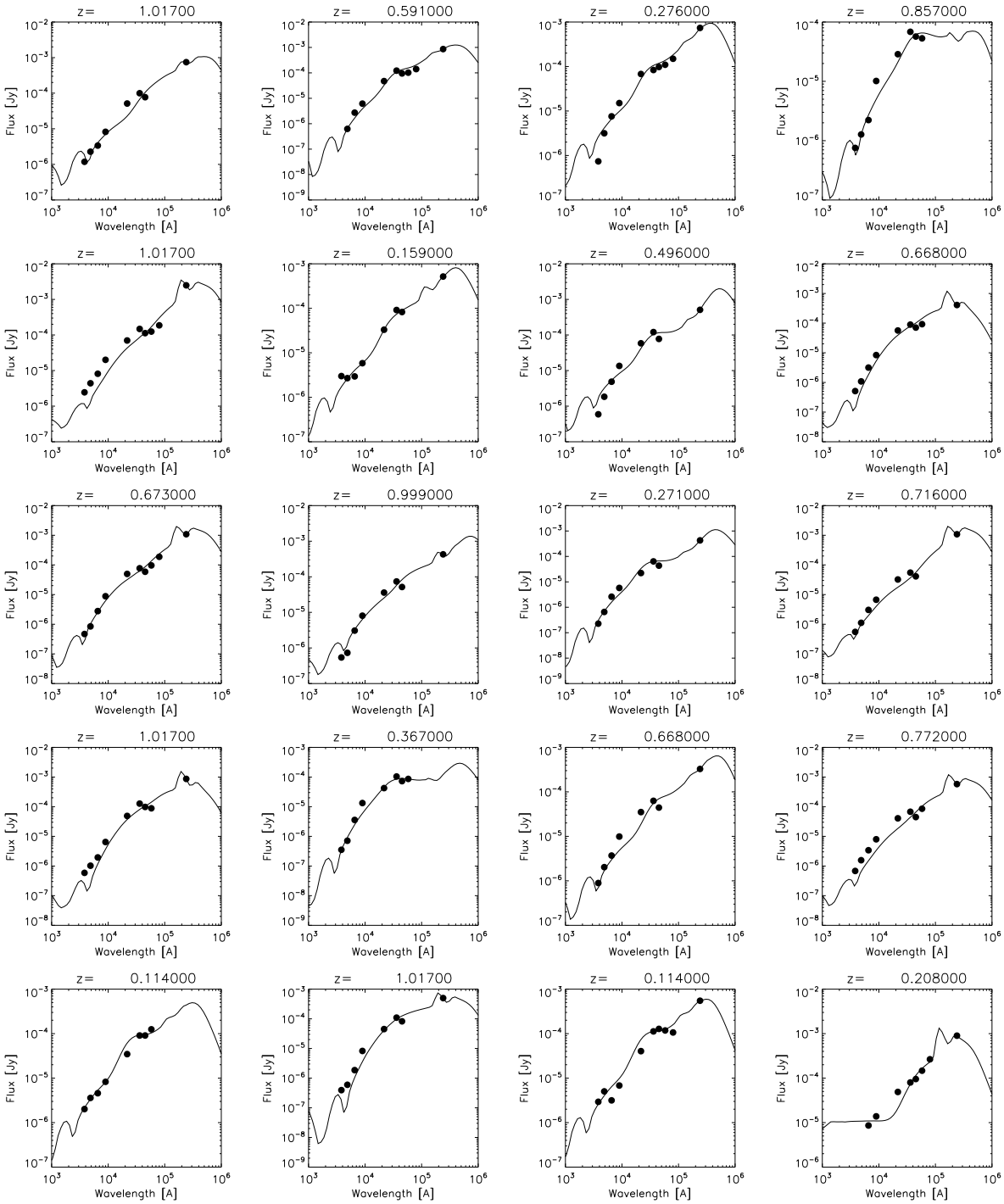

Figure 7.11: Comparison of the observational data and the torus model for the Type–2 AGN and QSOs.

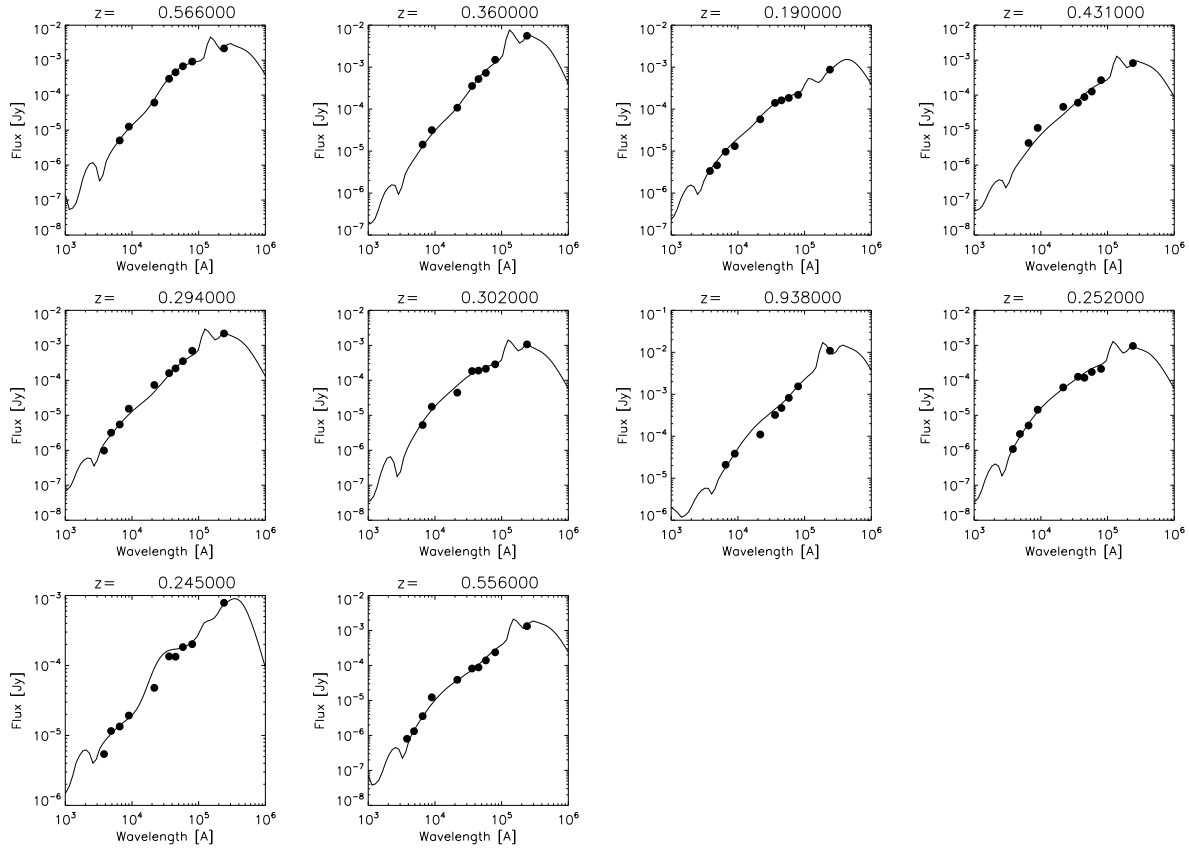
z= 0.566000
z= 0.360000
z= 0.190000
z= 0.431000
z= 0.294000
z= 0.302000
z= 0.938000
z= 0.252000
z= 0.245000
z= 0.556000
Flux [Jy]
Wavelength [A]

Figure 7.11:

Chapter 8

Conclusions

We have presented in this PhD thesis an extended study of the K–excess, a relatively new method meant to select quasars using infrared color and morphological criteria. As all different techniques, the K–excess has its strong and weak points and throughout this work we tried to cover the various aspects. The advantage of selecting objects using IR colors has been, until recently, been constrained by the limited field of view of infrared detectors, the complexity of observations in the infrared and the very demanding data reduction.

In chapter 2 we described in a detailed way all the necessary steps to reduce a set of K_s–band data and demonstrated that a careful, well–designed pipeline can generate properly reduced data, having successfully removed all systematic effects without a major impact on the total error budget.

As we have clearly stated in the introduction, "bias–free surveys do not exist". For this reason it is imperative to study the properties of the selected population, not only for getting an insight on the physics behind it but also for evaluating the complementarity of the method with respect to other selection techniques. The power of combining multi–wavelength information for reaching these goals has been clearly demonstrated in Chapters 5 and 6, where, without any spectroscopic information, we managed to successfully identify some 60 AGN and quasars – the majority of them being new discoveries – and evaluate the efficiency of the KX–method.

Modeling the environment of an active nucleus is still a challenge, not only because of the complexity of the physical processes taking place in its interior, but also due to the diversity in the AGN architecture (luminosity of the central black hole, size of the accretion disk and torus, polarization of jets) needed for explaining the observed features. Under such a complexity, it is inevitable to impose some constraints when modeling the emission in active galaxies. In Chapter 7 we have presented the results of the modeling of the tori of the AGN and quasars in our sample, assuming that the Unification Scheme is valid, i.e. that the differentiation between the Type–1 and Type–2 objects can be

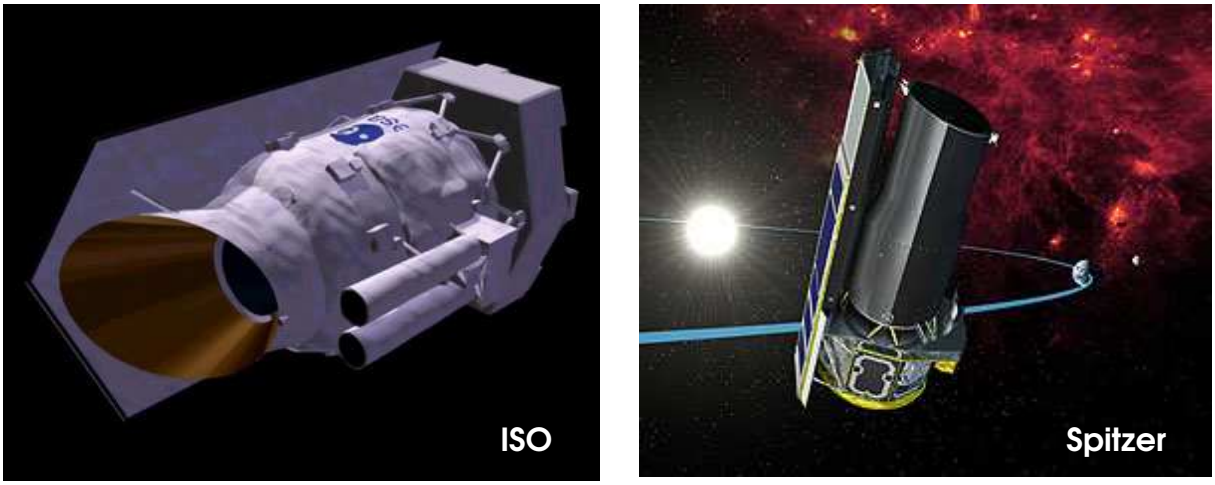

Figure 8.1: The past and the present of infrared space–borne astronomy: an artist's impression of ISO and Spitzer satellites (credits ESA, NASA).

attributed to the different viewing angles through which we look at the active nucleus. Under this assumption, the models used managed to successfully describe the spectral energy distribution of our objects, hence giving credibility to the Unification Scheme. Nevertheless, one should not forget that there are still open issues, as there are cases where the conclusions we induce from observations at different wavelengths do not fit in the Unification Scheme context. Such an example was given in Chapter 6, where the X–ray classification of some AGN was not in agreement with the one based on the shape of their spectral energy distribution.

UKIDSS, a near–infrared survey as ambitious (both in terms of sky coverage and limiting magnitude) as the SDSS in the optical, will undoubtedly complement our knowledge on AGN and QSOs, filling in the gap between the last of the SDSS filters, at $\sim 1\,\mu$m and the $3.6\,\mu$m, that corresponds to the first of the IRAC bands on–board Spitzer.

Space–borne observatories have played a major role in the quest for the AGN "Holy Grail", mainly with observations in the X–rays and infrared wavelengths. There is no doubt that the most efficient technique for detecting AGN is observations in the X–ray domain, as there are only few types of sources in the universe, other than AGN, that emit as much energy in the X–rays as AGN do. Additionally, observations at these wavelengths (or, to be X–ray "compliant", to these energies) is the only way to probe the processes that take place so close to the heart of the active galaxy, the black hole.

Going from the hot environment, in the vicinity of the back hole, to some 10^6 orders of magnitude lower temperature, in the cold torus, infrared satellites are complementing our effort to fill in the AGN puzzle. The potential of using infrared criteria, exclusively, to detect AGN, has already been demonstrated thanks to the color–color diagrams constructed using Spitzer photometry (namely the four IRAC bands, spanning from 3.6 to $8\,\mu$m). An extended study on this method has been presented in Chapter 4. Using this approach,

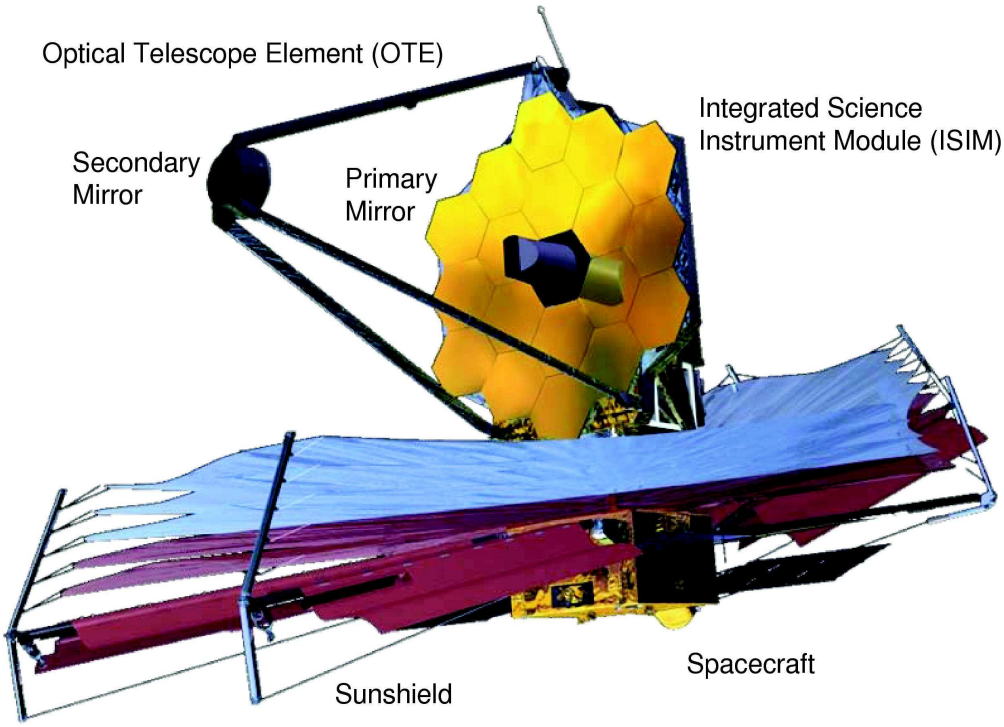

Figure 8.2: The future of infrared astronomy in space: the James Webb Space Telescope (taken from Gardner et al. (2006)).

we have managed to complement our sample of KX–selected quasars with the detection of some $\sim$ 35 additional AGN and quasars. Our analysis has demonstrated that these objects, the majority of which is Type–2 AGN, would be missed by any survey which *only partially* uses infrared color criteria. Therefore, the issue of observational biases from which surveys suffer comes once more into the discussion.

The James Webb Space Telescope (JWST), HST's successor, whose launch is foreseen in 2013, will significantly contribute to our current knowledge of the "dusty" Universe, with observations in the range $5 - 28\,\mu$m. With its 6.5 m primary mirror (we recall that HST's is 2.4 m and Spitzer's 0.85 cm), JWST 's advance will be close to 100–fold, in sensitivity, over Spitzer and 1000–fold over the largest ground–based telescopes (limited by the huge background noise due to Earth's atmosphere). The open issues in AGN research are multiple and JWST will be used to shed light to many of them, both at microscopic (i.e. parsec) and macroscopic (i.e. Giga–parsec) scales. Some of these issues are itemized below:

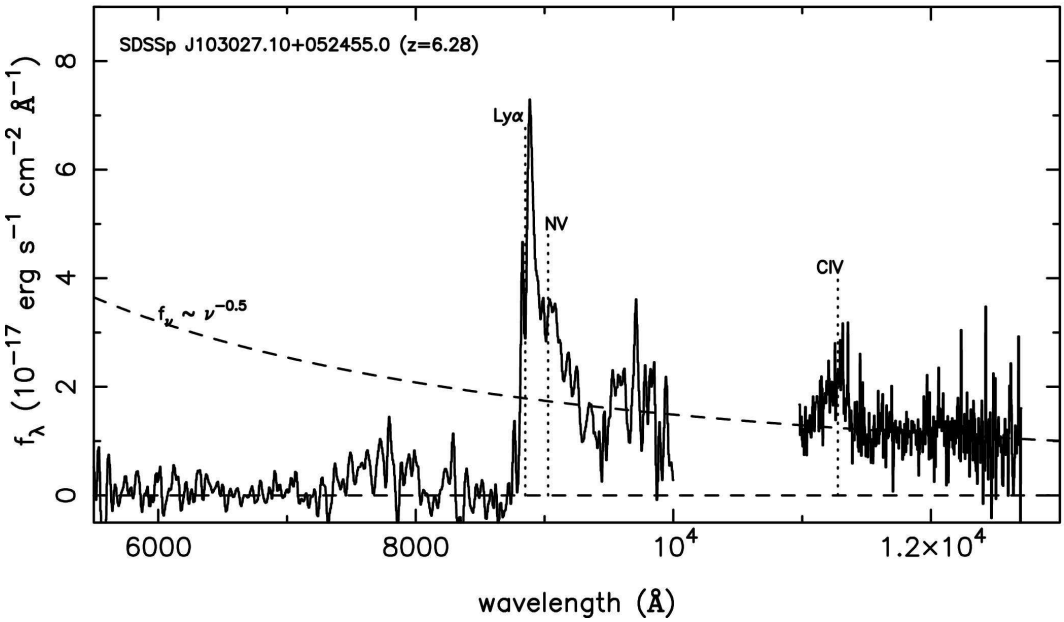

Figure 8.3: The first observational signature of the Gunn–Peterson trough effect, predicted in 1965 (Gunn & Peterson 1965), only took place 26 years later, with spectra of a SDSS quasar at a redshift $z = 6.28$ (Fan et al. 2001).

- The reionization epoch: when and how did it occur? Gunn & Peterson (1965) predicted that hydrogen in neutral state should depress the QSO continuum at wavelengths shortward of the Lyα emission line. This effect, known as the Gunn–Peterson trough, should be easy to discriminate from the sporadic absorption features due to intervening clumpy gas concentrations, known as the Lyα forest (Lynds (1971), Sargent et al. (1980)). Nevertheless, it took 26 years for the premier observation of the Gunn–Peterson trough, with observations of a SDSS quasar found at $z = 6.28$ (Fan et al. (2001), Becker et al. (2001)) Although this result shows that the fraction of neutral hydrogen has increased substantially between $z = 5.7$, where we see no signature of the Gunn–Peterson effect, and $z \approx 6$, where resides the closest quasar for which this effect is detected (Fan et al. (2001), Fan et al. (2003)), it is still unclear when the exact time of the reionization epoch took place, as even fractional abundances of neutral hydrogen would be sufficient to produce this effect (Miralda-Escudé et al. (2000), Gnedin (2000)). Deep infrared spectroscopic observations using NIRSpec, the near–IR ($1 - 5\,\mu$m) spectrograph on–board JWST, will allow to identify high–redshift quasars ($z > 7$) and study the presence of the Gunn–Peterson trough or their damped Lyα features, in order to explore the ionization phase of the cosmic history.

- Understand the mechanism that triggers the AGN activity and its link to the stellar population in the bulge, the latter one being directly related to the black hole

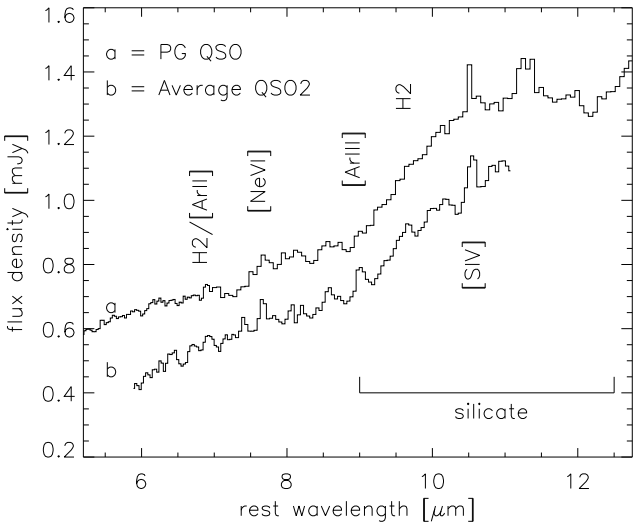

Figure 8.4: Detection of the silicate 9.7 μm feature, in emission, in a Type–2 QSO (Sturm et al. 2006).

residing in the center of the host galaxy (Tremaine et al. 2002). The active phase in galaxies, that has its peak at $z \sim 2.2$ as the QSO luminosity function reveals, seems to be contemporary with a high star–formation rate. Nevertheless, the spectra of the (few) QSOs found at $z > 6$ show minor differences compared to those of quasars residing at much lower redshifts (Fan et al. (2001), Freudling et al. (2003)), implying that some massive black holes and their associated stellar populations formed during the early stages of the universe. The Mid–InfraRed Instrument (MIRI) and NIRSpec will observe active galaxies spanning a wide luminosity range, in order to associate the stellar formation and the black hole to the galaxy's active phase and its relation with cosmic evolution.

- The origin of dust emission in AGN: although current torus models successfully fit the observational data of low–brightness AGN, there is quite some confusion regarding the properties of the high–luminosity members. The well–studied 9.7 μm silicate feature, which, according to the Unification Scheme, should be present in absorption in Type–2 AGN and in emission in Type–1 objects, has only recently been discovered in the latter ones (Siebenmorgen et al. (2005), Hao et al. (2005)). More surprisingly, though, Sturm et al. (2006) reported on the detection of the 9.7 μm feature *also in emission*, in observations of an infrared–selected Type–2 QSO. Additionally, Type–2 quasars are often believed to be related to Ultra Luminous InfraRed Galaxies (ULIRGs, Sanders & Mirabel (1996)). Nevertheless, Sturm et al. (2006) report on the complete absence of PAH (Polycyclic Aromatic Hydrocarbons),

emission features associated to dusty environments. JWST, with its infrared filters spanning from $5-28\,\mu$m will sample dusty Type–2 QSOs and analyze their spectra, with the scope of better studying the origin of dust emission in active galaxies.

I hope that the work presented in this thesis has contributed to our knowledge in the AGN field. What I am sure of, however, is that the future of infrared astronomy looks definitely exciting.

Chapter 9

Erratum

Just before the printing of the manuscript we realized that there was a mis–identification of object S03, the third of the objects with point–like morphology and AGN SWIRE colors (see Fig. 6.3). Object S03 was mis–identified with the isolated object found at the outer right part of the SWIRE color–color diagram. The spectral energy distribution of object S03 also shows a QSO–like shape, as it can be seen in the lower left part of Fig. 9.1. The energy distribution of the isolated object, which is a starburst galaxy, has also been plotted in the lower right panel. Its photometric redshift was estimated to be $z_{phot} = 0.9$ and was fitted by a "QSO low" template.

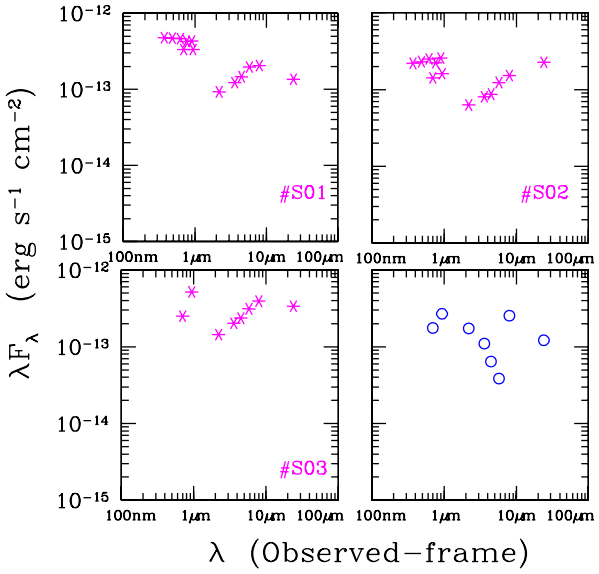

Figure 9.1:

As far as the X–ray properties of object S03 are concerned, it has 68.65 counts in the soft (0.5−2 keV) band and a flux equal to $\mathrm{Flux}_B = 1.49 \times 10^{-14}$ erg s^{-1}, but does not

show any hard X–ray emission.

The identification of object S03 as quasar does not change in any significant way the conclusions we have drawn, as we have generated a statistically large sample that contains 59 QSOs and AGNs. As this QSO also shows X–ray emission and has AGN–like infrared ($3.6-8.0\,\mu$m) color properties, it confirms that X–ray and IR surveys are very efficient in detecting objects containing an active nucleus.

Bibliography

Absil, O., di Folco, E., Mérand, A., et al. 2006, A&A, 452, 237

Andreon, S., Willis, J., Quintana, H., et al. 2004, MNRAS, 353, 353

Antonucci, R. 1993, ARA&A, 31, 473

Bailer–Jones, C. A. L. & Lamm, M. 2003, MNRAS, 339, 477

Baldi, A., Molendi, S., Comastri, A., et al. 2002, ApJ, 564, 190

Barcons, X., Carballo, R., Carrera, F. J., et al. 2003, MNRAS, 343, 137

Barger, A. J., Cowie, L. L., Capak, P., et al. 2003, AJ, 126, 632

Barkhouse, W. A. & Hall, P. B. 2001, AJ, 121, 2843

Barvainis, R. 1987, ApJ, 320, 537

Baum, S. A., O'Dea, C. P., Dallacassa, D., de Bruyn, A. G., & Pedlar, A. 1993, ApJ, 419, 553

Becker, R. H., Fan, X., White, R. L., et al. 2001, AJ, 122, 2850

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559

Benn, C. R., Vigotti, M., Carballo, R., Gonzalez-Serrano, J. I., & Sánchez, S. F. 1998, MNRAS, 295, 451

Berta, S., Fritz, J., Franceschini, A., Bressan, A., & Pernechele, C. 2003, A&A, 403, 119

Bertin, E. & Arnouts, S. 1996, A&A Supl., 117, 393

Bessell, M. S. 1990, PASP, 102, 1181

Bevington, P. & Robinson, K. 2002, Data Reduction and Error Analysis for the Physical Sciences (Mc Graw Hill)

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Bondi, M., Ciliegi, P., Zamorani, G., et al. 2003, A&A, 403, 857

Borguet, B. 2006, Master's thesis, Faculté des Sciences, Département d'Astrophysique, Géophysique et Océanographie, Universite de Liège

Boyle, B. J., Shanks, T., Croom, S. M., et al. 2000, MNRAS, 317, 1014

Brotherton, M. S., Tran, H. D., Becker, R. H., et al. 2001, ApJ, 546, 775

Buser, R. & Kurucz, R. L. 1992, A&A, 264, 557

Calabretta, M. & Greisen, E. 2002, A&A, 395, 1077

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Carilli, C. L., Harris, D. E., Pentericci, L., et al. 2002, ApJ, 567, 781

Charmandaris, V. 2006, arXiv:astro-ph/0606143

Chiappetti, L., Tajer, M., Trinchieri, G., et al. 2005, A&A, 439, 413

Ciliegi, P., Zamorani, G., Bondi, M., et al. 2005, A&A, 441, 879

Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393

Croom, S. M., Warren, S. J., & Glazebrook, K. 2001, MNRAS, 328, 150

Dahn, C. C., Harris, H. C., Vrba, F. J., et al. 2002, AJ, 124, 1170

de Ruyter, S., van Winckel, H., Maas, T., et al. 2006, A&A, 448, 641

Della Ceca, R., Maccacaro, T., Caccianiga, A., et al. 2004, A&A, 428, 383

Draine, B. T. & Lee, H. M. 1984, ApJ, 285, 89

Dullemond, C. P. & van Bemmel, I. M. 2005, A&A, 436, 47

Efstathiou, A. & Rowan-Robinson, M. 1995, MNRAS, 273, 649

Ehle, M., Breitfellner, M., González Riestra, R., et al. 2005, XMM–Newton User's Handbook, ESA

Eisenhardt, P. R., Stern, D., Brodwin, M., et al. 2004, ApJS, 154, 48

Ellison, S. L., Hall, P. B., & Lira, P. 2005, AJ, 130, 1345

Elvis, M., Wilkes, B. J., McDowell, J. C., et al. 1994, ApJS, 95, 1

Falco, E. E., Impey, C. D., Kochanek, C. S., et al. 1999, ApJ, 523, 617

Fan, J. H., Xie, G. Z., Pecontal, E., Pecontal, A., & Copin, Y. 1998, ApJ, 507, 173

Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, AJ, 122, 2833

Fan, X., Strauss, M. A., Schneider, D. P., et al. 2003, AJ, 125, 1649

Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10

Fiore, F., Brusa, M., Cocchia, F., et al. 2003, A&A, 409, 79

Francis, P. J., Hewett, P. C., Foltz, C. B., et al. 1991, ApJ, 373, 465

Francis, P. J., Whiting, M. T., & Webster, R. L. 2000, Pub. Ast. Soc. Austr., 17, 56

Freudling, W., Corbin, M. R., & Korista, K. T. 2003, ApJ Letters, 587, L67

Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, ArXiv Astrophysics e-prints

Gnedin, N. Y. 2000, ApJ, 535, 530

Granato, G. L. & Danese, L. 1994, MNRAS, 268, 235

Gras–Velázquez, À. & Ray, T. P. 2005, A&A, 443, 541

Gregg, M. D., Lacy, M., White, R. L., et al. 2002, ApJ, 564, 133

Greisen, E. & Calabretta, M. 2002, A&A, 395, 1061

Gunn, J. E. & Peterson, B. A. 1965, ApJ, 142, 1633

Hall, P., Green, R., & Cohen, M. 1998, ApJS, 119, 999

Hao, L., Spoon, H. W. W., Sloan, G. C., et al. 2005, ApJ Letters, 625, L75

Hatziminaoglou, E., Pérez-Fournon, I., Polletta, M., et al. 2005, AJ, 129, 1198

Hewett, P. C., Warren, S. J., Leggett, S. K., & Hodgkin, S. T. 2006, MNRAS, 367, 454

Joyce, R. 1992, in Astronomical CCD Observing and Reduction Techniques, ed. S. B. Howell, Vol. 23, ASP Conference Series, 258

Katgert, P., Merkelijn, J. K., Le Poole, R. S., & van der Laan, H. 1973, A&A, 23, 171

Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, AJ, 98, 1195

Kenyon, S. J. & Hartmann, L. 1987, ApJ, 323, 714

Kharb, P., O'Dea, C. P., Baum, S. A., Colbert, E. J. M., & Xu, C. 2006, ArXiv Astrophysics e-prints

Kleinmann, S. G., Lysaght, M. G., Pughe, W. L., et al. 1994, Astroph. & Space Science, 217, 11

Krolik, J. H. 1999, Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment (Princeton University Press)

Kukula, M. J., Holloway, A. J., Pedlar, A., et al. 1996, MNRAS, 280, 1283

Lacy, M., Petric, A., Sajina, A., et al. 2006, ArXiv Astrophysics e-prints

Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166

Lal, D. V., Shastri, P., & Gabuzda, D. C. 2004, A&A, 425, 99

Landolt, A. U. 1992, AJ, 104, 340

Lawrence, A., Warren, S. J., Almaini, O., et al. 2006, ArXiv Astrophysics e-prints

Lehmann, I., Hasinger, G., Schmidt, M., et al. 2001, A&A, 371, 833

Lewis, I.J., e. a. 2002, MNRAS, 333, 279

Lobanov, A. P. & Zensus, J. A. 2006, arXiv:astro-ph/0606143

Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897

Lynds, R. 1971, ApJ Letters, 164, L73+

Madau, P. 1995, ApJ, 441, 18

Maiolino, R., Salvati, M., Bassani, L., et al. 1998, A&A, 338, 781

Manners, J. 2002, PhD thesis, The University of Edinburgh

Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425

Matt, G. 2002, Royal Society of London Philosophical Transactions Series A, 360, 2045

McCarthy, P. J. 1993, ARA&A, 31, 639

McMahon, R. G. & Irwin, M. J. 1992, in ASSL Vol. 174: Digitised Optical Sky Surveys, ed. H. T. MacGillivray & E. B. Thomson, 417+–

Miralda-Escudé, J., Haehnelt, M., & Rees, M. J. 2000, ApJ, 530, 1

Monet, D., Levine, S., & Canzian, B. 2003, AJ, 125, 984

Nenkova, M., Ivezić, Ž., & Elitzur, M. 2002, ApJ Letters, 570, L9

Neugebauer, G., Green, R. F., Matthews, K., et al. 1987, ApJS, 63, 615

Neugebauer, G., Oke, J. B., Becklin, E. E., & Matthews, K. 1979, ApJ, 230, 79

Oke, J. B. & Korycansky, D. G. 1982, ApJ, 255, 11

Osterbrock, D. E. 1981, ApJ, 249, 462

Osterbrock, D. E. & Koski, A. T. 1976, MNRAS, 176, 61P

Pacholczyk, A. G. & Wisniewski, W. Z. 1967, ApJ, 147, 394

Persson, S., Murphy, D., Gunnels, S., et al. 2002, AJ, 124, 619

Persson, S., Murphy, D., Krzeminski, W., Roth, M., & Rieke, M. 1998, AJ, 116, 2745

Peterson, B. M. 2001, An introduction to Active Galactic Nuclei (Cambridge University Press)

Pickles, A. 1998, PASP, 110, 863

Pier, E. A. & Krolik, J. H. 1992, ApJ, 401, 99

Pierre, M., Valtchanov, I., Altieri, B., et al. 2004, Journal of Cosmology and Astro-Particle Physics, 9, 11

Polletta, M. d. C., Wilkes, B. J., Siana, B., et al. 2006, ApJ, 642, 673

Refsdal, S. 1964, MNRAS, 128, 295

Refsdal, S. & Surdej, J. 1994, Reports of Progress in Physics, 57, 117

Rengelink, R., Nonino, M., da Costa, L., et al. 1998, ArXiv Astrophysics e-prints

Riaud, P., Mawet, D., Absil, O., et al. 2006, A&A, in press

Richards, G. T., Hall, P. B., Vanden Berk, D. E., et al. 2003, AJ, 126, 1131

Rieke, G. H. 1978, ApJ, 226, 550

Rieke, G. H., Young, E. T., Engelbracht, C. W., et al. 2004, ApJS, 154, 25

Risaliti, G. & Elvis, M. 2004, A Panchromatic View of AGN (ASSL Vol. 308: Supermassive Black Holes in the Distant Universe), 187

Rowan-Robinson, M. 1992, MNRAS, 258, 787

Sandage, A. 1965, ApJ, 141, 1560

Sanders, D. B. & Mirabel, I. F. 1996, ARA&A, 34, 749

Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D. 1980, ApJS, 42, 41

Schartmann, M., Meisenheimer, K., Camenzind, M., Wolf, S., & Henning, T. 2005, A&A, 437, 861

Schmidt, M. 1963, Nature, 197, 1040

Schmidt, M. & Green, R. F. 1983, ApJ, 269, 352

Schneider, D. P., Fan, X., Hall, P. B., et al. 2003, AJ, 126, 2579

Seyfert, C. K. 1943, ApJ, 97, 28

Sharp, R. G., Sabbey, C. N., Vivas, A. K., et al. 2002, MNRAS, 337, 1153

Siebenmorgen, R., Haas, M., Krügel, E., & Schulz, B. 2005, A&A, 436, L5

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103

Silverman, J. D., Green, P. J., Barkhouse, W. A., et al. 2005, ApJ, 618, 123

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Skrutskie, M. F., Schneider, S. E., Stiening, R., et al. 1997, in ASSL Vol. 210: The Impact of Large Scale Near-IR Sky Surveys, ed. F. Garzon, N. Epchtein, A. Omont, B. Burton, & P. Persi, 25

Smith, J. A., Tucker, D. L., Kent, S., et al. 2002, AJ, 123, 2121

Stenholm, L. 1994, A&A, 290, 393

Stephens, J. R. 1980, ApJ, 237, 450

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163

Strüder, L., Briel, U., Dennerl, K., et al. 2001, A&A, 365, L18

Sturm, E., Hasinger, G., Lehmann, I., et al. 2006, ApJ, 642, 81

Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, ApJS, 155, 271

Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, ApJ, 574, 740

Turner, M. J. L., Abbey, A., Arnaud, M., et al. 2001, A&A, 365, L27

Turnshek, D. A. 1984, ApJ, 280, 51

Urry, C. M. & Padovani, P. 1995, PASP, 107, 803

van Boekel, R., Min, M., Leinert, C., et al. 2004, Nature, 432, 479

Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549

Véron-Cetty, M. P. & Véron, P. 2000, VizieR Online Data Catalog, 7215, 0

Warren, S. & Hewett, P. 2002, in ASP Conf. Ser. 283: A New Era in Cosmology, ed. N. Metcalfe & T. Shanks, 369–+

Warren, S. J., Hewett, P. C., & Foltz, C. B. 2000, MNRAS, 312, 827

Webster, R. L., Francis, P. J., Peterson, B. A., Drinkwater, M. J., & Masci, F. J. 1995, Nature, 375, 469

Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, ApJ, 373, 23

Whittet, D. 2003, Dust in the galactic environment /2nd edition/ (Bristol & Philadelpia, Institue of Physics Publishing, 2003. 390 p.)

Wills, B. J., Netzer, H., & Wills, D. 1985, ApJ, 288, 94

Wills, B. J., Wills, D., Evans, N. J., et al. 1992, ApJ, 400, 96

Winn, J. N., Lovell, J. E. J., Chen, H.-W., et al. 2002, ApJ, 564, 143

Wright, A. E. & Otrupcek, R. 1990, in PKSCAT90: Radio Source Catalogue and Sky Atlas, Australia Telescope National Facility

Wyithe, J. S. B. & Padmanabhan, T. 2006, MNRAS, 366, 1029

Xu, C. K., Lonsdale, C. J., Shupe, D. L., et al. 2003, ApJ, 587, 90

Zakamska, N. L., Strauss, M. A., Krolik, J. H., et al. 2003, AJ, 126, 2125

Zheng, W., Kriss, G. A., Davidsen, A. F., et al. 1995, ApJ, 444, 632

Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, ApJ, 475, 469

Appendix A

SExtractor

A.1 Introduction

SExtractor (Source Extractor (Bertin & Arnouts 1996)) is a program that builds a catalog of objects from an astronomical image. The purpose of this Appendix is to provide the reader with the basic information, that will allow to understand the steps of the reduction involving SExtractor. For a detailed description, however, we propose to read the "official" user's guide, written by E. Bertin. Another manual with a lot of useful information and tips on how to run the program is "Source Extractor for Dummies", by Benne W. Holwerda (Space Telescope Science Institue) that can be found on the WWW.

SExtractor is a freely distributed program that can be installed on any UNIX machine. It is run from the shell with the following syntax:

% sex image -c configuration-file

where *image* is a FITS file and *configuration-file* an ASCII file containing the parameters controlling SExtractor. In addition to the configuration file, SExtractor needs a file containing the list of parameters that will be listed in the output catalog for every detection. This allows the software to compute only catalog parameters that are needed. The catalog parameter file is also an ASCII file that can be edited by the user. Apart from the output catalog with the photometry, positions and other information on the detections, SExtractor can produce a series of FITS files, that serve as "check" images. These files contain information on the sky background and the object detections (what is called "object mask" image).

The series of steps followed by SExtractor when operating on an astronomical image are listed below:

- The program scans the whole image and determines the mean background level and

its RMS.

- Subtracts the background.
- Filters the background–subtracted image, i.e. convolves the image with a specified profile.
- Finds objects, according to the user–defined threshold parameters.
- Deblends detections, i.e. breaks up detections into different objects.
- Measures shapes and positions.
- Cleans the detections, i.e. reconsiders the detections, accounting for contributions from neighboring objects.
- Performs photometry.
- Provides a stellarity index for each detection, i.e. the probability that the detection is a point–like (stellarity = 1) or an extended source (stellarity = 0).
- Produces output catalogs and "check" images, according to the contents of the configuration and catalog parameter files.

A.2 Detection parameters

The main parameters of the configuration file related to the detection of objects are the following:

- BACK_SIZE regulates the estimation of the sky background. In an area of BACK_SIZE the mean and the σ of the distribution of the pixel values is computed. The most deviant values are discarded and the median and the standard deviation σ are re-computed. This is repeated until all the remaining pixel values are within mean $\pm 3\sigma$. The size of BACK_SIZE is very important for the proper estimation of the sky background, since too small values will be influenced by the presence of very bright objects in the fields, too high values will not be able to account for small scale variations. The value set for our data was 70 pixels (14 arcseconds).
- FILTER_NAME controls the type of filter to be used for smoothing the sky–subtracted image. There are predefined filters provided by SExtractor. The choise of the filter mainly depends on two factors:
  - The width of the filter has to be of the same size as the FWHM of the point–like sources.

  - The type of sources we want to detect. A "Mexican Hat" is a very sharp filter, that makes the intensity at the wings of an object's profile drop very fast. It favors crowded fields with bright objects, but it is unrecommended for fields with faint objects, since they will be smoothed at the level of the background. For the latter case, a Gaussian filter is recommended.

- There are two threshold parameters, which indicate the level above which SExtractor should start treating *adjacent* pixels as if they were part of objects.

  - DETECT_THRESH: all the pixels values have to be above this threshold.
  - At least DETECT_MINAREA pixels must be above the DETECT_THRESH.

  Both parameters can be specified as relative to the background RMS or in absolute values, i.e. ADUs or surface brightness.

- WEIGHT_TYPE is a parameter that allows to use a weight image (also called "weight maps") for the detection of objects. This option is very useful in the case where the noise properties are spatially varying and global detection thresholds can not be used over the whole frame. For the Las Campanas mosaic it was essential to use this option, due to the inhomogenuity of the data.

- ASSOC_NAME: it is possible to force SExtractor to look for objects within a radius ASSOC_RADIUS from a list of (x,y) positions. These positions are stored in an ASCII file, whose name is the value for ASSOC_NAME. SExtractor will consider that at the indicated positions there is a detection only if the objects respect the threshold criteria.

A.3 Photometry parameters

SExtractor measures the fluxes of the detected objects. The output of the photometry can either be fluxes or magnitudes, depending on the user's preference. SExtractor, however, has different ways for measuring fluxes and there are five parameters controlling the photometry:

- ISO: the pixels above the detection threshold constitute an isophotal area. The flux (counts in pixels above threshold minus the background) is the isophot magnitude. There are no parameters for the user to adjust.

- ISOCOR: considering that the pixels above the detection threshold make part of an analytical two–dimensional profile, the ISO flux corrected for deformations from a Gaussian profile is called ISOCOR. There are no parameters for the user to adjust.

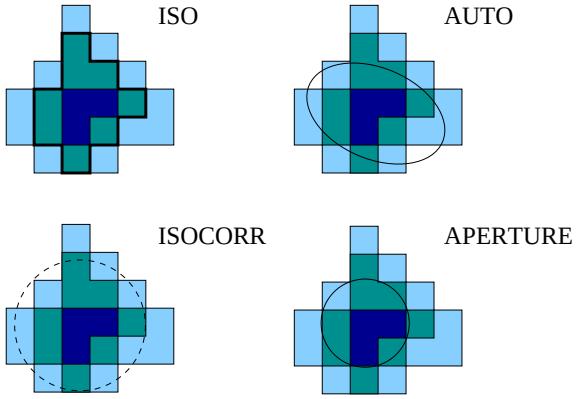

Figure A.1: Schematic representation of the different types of photometry that can be performed with SExtractor.

- AUTO: SExtractor uses a flexible elliptical aperture around every detected object and measures all the flux inside that. The default parameters in the configuration file, controling AUTO's photometry seem to be the best to use.
- BEST A choise between AUTO and ISOCORR. In a crowded field, the photometry can be influenced by the presence of nearby objects. If the flux contribution from pixels adjacent to a given detection is more than 10%, then BEST = ISOCOR. If not, BEST = AUTO.
- APER Photometry measured within a fixed aperture centered on each detection. The user has to define the aperture's diameter.

A.4 Object classification

SExtractor distinguishes between point–like objects and galaxies by assigning a stellarity index to each detection, one for "stars" and zero for galaxies. The stellarity is determined using a neural network, which learns how to classify the detections based on high signal–to–noise point–like objects in the field. The classification is seeing–dependent, i.e. the way the objects are classified depends on the shape of the objects used for teaching the neural network to perform the classification. As a result, for a group of frames obtained under different seeing conditions, a certain stellarity index might not correspond to the same type of objects. The classification will be less accurate for faint stars than faint galaxies, because of crowing effect: faint stars have a higher probability of having their profiles deformed because of a background galaxy in the same region. In such a case, it would be impossible for SExtractor to correctly identify the star. As a result, a combination of a magnitude and stellarity cut might be a better criterion for classifying objects, instead of a simple stellarity limit.

Appendix B

List of QSO candidates

Table B.1 presents the photometry and astrometry for the 85 QSO candidates. The first column provides the identification number of each detection. Their celestial coordinates, presented in columns two and three were calculated relative to the USNO-2 catalog (Monet et al. 2003). The units of the right ascension are hours, minutes and seconds and the units of the declination are degrees, arcminutes and arcseconds; the equinox is J2000. Columns 4, 6 and 8 show the $3''$ aperture photometry for the R, z' and K_s–band, respectively. The R, K_s magnitudes are given in the Vega photometric system, while the z'–band photometry is given in AB magnitudes. To convert the AB magnitudes to Vega, use the formula $z'_{Vega} = z'_{AB} - 0.549$. The 1σ magnitude uncertainties are presented in columns 5, 7 and 9. Column 10 shows the detection probability for each source, as computed from simulations (see section 2.4).

No	RA (J2000) h m s	Dec (J2000) ° ′ ″	R	σ_R	z'	$\sigma_{z'}$	K_s	σ_{K_s}	DP (%)
1	2 22 42.24	−4 30 44.21	20.858	0.012	20.376	0.017	17.538	0.171	94.4
2	2 22 42.60	−4 30 18.14	20.677	0.010	21.033	0.030	18.065	0.195	75.0
3	2 22 42.96	−4 33 14.54	20.610	0.010	20.629	0.021	18.316	0.195	83.3
4	2 22 44.04	−4 20 13.99	21.316	0.018	20.564	0.020	17.330	0.116	85.7
5	2 22 47.28	−4 35 31.45	20.906	0.013	20.476	0.018	17.667	0.099	80.6
6	2 22 47.64	−4 25 47.50	20.565	0.009	20.009	0.012	16.961	0.047	91.4
7	2 22 48.36	−4 17 37.50	21.449	0.021	20.909	0.028	17.567	0.110	91.4
8	2 22 48.36	−4 34 06.46	20.586	0.010	20.248	0.015	17.411	0.079	88.9
9	2 22 50.16	−4 28 41.48	22.664	0.063	21.344	0.040	17.979	0.122	88.9
10	2 22 50.88	−4 20 00.38	21.003	0.014	19.902	0.011	16.710	0.047	00.0
11	2 22 51.96	−4 25 07.97	19.213	0.003	18.847	0.004	16.091	0.028	00.0
12	2 22 52.68	−4 34 57.65	22.288	0.046	21.113	0.033	17.930	0.107	80.6
13	2 22 54.48	−4 15 59.62	21.489	0.022	20.645	0.022	17.624	0.093	88.9
14	2 22 55.92	−5 00 40.64	18.921	0.002	18.744	0.004	16.529	0.047	88.6
15	2 22 58.44	−4 27 18.65	22.301	0.045	20.908	0.027	17.876	0.110	86.1
16	2 22 58.80	−4 25 55.42	22.850	0.075	21.417	0.043	17.787	0.104	88.6
17	2 22 58.80	−4 58 52.14	19.337	0.003	19.323	0.006	16.407	0.029	82.9
18	2 23 02.04	−4 22 36.55	21.057	0.015	20.680	0.022	18.155	0.110	88.9
19	2 23 04.20	−4 44 35.30	20.446	0.008	20.360	0.017	18.037	0.109	77.8
20	2 23 06.00	−4 33 23.62	20.394	0.008	20.633	0.021	17.394	0.073	82.9
21	2 23 06.72	−4 17 20.94	22.788	0.071	21.424	0.044	17.911	0.113	86.1
22	2 23 07.44	−4 31 33.13	22.554	0.058	21.174	0.034	17.893	0.110	88.9
23	2 23 08.16	−5 00 28.04	18.674	0.002	18.898	0.004	17.322	0.081	94.4
24	2 23 09.24	−4 31 38.60	22.473	0.053	21.271	0.037	17.992	0.095	80.6
25	2 23 13.56	−4 13 59.41	20.372	0.008	20.395	0.017	18.307	0.180	86.1
26	2 23 18.24	−4 24 37.01	21.598	0.024	21.157	0.034	17.994	0.128	83.3
27	2 23 25.80	−4 22 54.05	21.415	0.020	21.492	0.046	18.171	0.217	80.6
28	2 23 25.80	−4 33 30.35	21.919	0.032	21.025	0.030	17.573	0.087	86.1
29	2 23 28.68	−4 33 08.64	21.555	0.023	20.636	0.022	17.414	0.074	94.4
30	2 23 32.64	−4 36 03.20	21.613	0.025	20.664	0.022	17.510	0.069	88.9
31	2 23 34.44	−4 55 14.70	19.238	0.003	19.488	0.007	17.798	0.088	77.8
32	2 23 36.60	−4 53 36.92	20.523	0.009	20.141	0.013	17.587	0.091	91.4
33	2 23 36.96	−4 30 51.52	22.305	0.046	21.179	0.036	18.035	0.150	80.6
34	2 23 40.20	−4 57 26.39	18.485	0.002	18.695	0.004	16.986	0.049	91.4
35	2 23 40.56	−4 38 34.15	19.512	0.004	19.442	0.007	17.243	0.088	97.1
36	2 23 40.56	−4 49 00.08	18.243	0.001	18.491	0.003	16.992	0.048	94.3
37	2 23 40.92	−4 22 55.27	19.473	0.004	19.598	0.008	17.351	0.065	91.4
38	2 23 40.92	−4 49 38.39	18.492	0.002	18.538	0.003	16.600	0.048	97.1
39	2 23 42.36	−4 31 06.17	21.548	0.023	20.783	0.025	18.004	0.119	86.1
40	2 23 42.36	−4 52 41.12	17.838	0.001	17.995	0.002	16.266	0.026	88.6
41	2 23 47.04	−4 24 40.57	21.711	0.026	20.738	0.022	17.691	0.078	97.1
42	2 23 47.40	−4 28 29.35	21.443	0.021	20.842	0.025	17.497	0.068	94.3
43	2 23 47.76	−4 14 21.16	21.473	0.022	21.123	0.032	18.004	0.114	88.9
44	2 23 49.56	−4 17 50.75	19.366	0.003	19.147	0.005	16.744	0.042	94.3
45	2 23 50.64	−4 31 57.94	19.394	0.003	19.457	0.007	17.223	0.057	91.4

Table B.1: QSO candidate list.

No	RA (J2000) h m s	Dec (J2000) ° ′ ″	R	σ_R	z'	$\sigma_{z'}$	K_s	σ_{K_s}	DP (%)
46	2 23 51.36	−4 33 04.97	21.072	0.015	20.377	0.017	17.503	0.076	97.2
47	2 23 52.44	−4 18 21.35	20.267	0.007	20.131	0.013	17.739	0.090	88.9
48	2 23 57.12	−4 19 57.54	20.730	0.011	20.285	0.015	17.426	0.070	88.6
49	2 23 59.28	−4 23 07.98	21.251	0.017	20.708	0.022	17.890	0.117	77.8
50	2 24 01.80	−4 24 56.74	20.783	0.011	20.224	0.014	17.460	0.071	75.0
51	2 24 06.12	−4 59 50.96	20.975	0.013	20.679	0.022	17.664	0.091	74.3
52	2 24 06.84	−4 25 43.54	20.870	0.012	20.265	0.015	17.299	0.050	72.2
53	2 24 07.20	−4 54 18.83	19.581	0.004	19.865	0.010	18.119	0.115	83.3
54	2 24 09.00	−4 18 20.16	20.617	0.010	20.031	0.012	17.077	0.040	94.3
55	2 24 14.04	−4 56 17.34	18.659	0.002	18.939	0.004	17.492	0.060	82.9
56	2 24 14.76	−4 23 13.27	21.334	0.019	20.470	0.018	17.271	0.064	91.4
57	2 24 16.56	−4 56 42.90	20.696	0.010	20.808	0.024	17.539	0.082	94.4
58	2 24 18.36	−4 17 27.92	20.759	0.011	20.768	0.023	18.508	0.140	86.1
59	2 24 19.08	−4 37 26.44	18.875	0.002	18.901	0.004	16.959	0.037	86.1
60	2 24 19.44	−4 35 20.65	21.091	0.015	20.480	0.019	17.543	0.067	94.3
61	2 24 19.80	−4 56 02.00	20.786	0.011	20.356	0.016	17.354	0.054	88.9
62	2 24 24.84	−4 18 27.58	21.366	0.020	20.798	0.024	17.228	0.109	82.9
63	2 24 31.68	−4 58 34.64	17.839	0.000	17.874	0.001	16.012	0.030	88.2
64	2 24 32.04	−4 39 06.91	18.456	0.001	18.782	0.004	17.406	0.173	82.6
65	2 24 32.04	−4 41 52.55	18.657	0.002	18.815	0.004	17.117	0.061	94.1
66	2 24 32.76	−4 40 42.96	20.438	0.008	20.424	0.018	17.893	0.183	94.4
67	2 24 38.16	−4 59 14.60	17.948	0.001	18.070	0.002	16.381	0.033	94.3
68	2 24 46.08	−4 26 39.44	20.832	0.013	20.283	0.016	16.841	0.089	94.1
69	2 24 46.80	−4 38 57.34	20.671	0.011	20.450	0.022	17.869	0.186	72.2
70	2 24 51.12	−4 32 58.42	21.377	0.021	20.605	0.021	17.854	0.248	74.3
71	2 24 51.48	−4 32 32.78	21.924	0.034	20.631	0.022	17.470	0.152	77.8
72	2 24 55.80	−4 46 56.14	21.345	0.020	20.709	0.027	17.664	0.086	88.9
73	2 24 57.96	−4 42 28.94	20.719	0.012	20.436	0.021	17.546	0.098	97.1
74	2 25 00.84	−4 34 37.13	18.907	0.002	18.929	0.005	16.830	0.058	70.6
75	2 25 07.32	−4 25 29.75	21.227	0.018	20.829	0.027	17.978	0.138	86.1
76	2 25 07.68	−4 20 36.02	19.624	0.004	19.867	0.011	18.388	0.176	86.1
77	2 25 09.12	−4 53 16.80	21.360	0.020	20.927	0.033	17.993	0.121	77.8
78	2 25 15.24	−4 40 08.69	19.839	0.005	19.894	0.013	18.115	0.212	77.8
79	2 25 15.60	−5 01 01.52	21.390	0.021	20.610	0.025	17.479	0.070	85.7
80	2 25 34.68	−4 24 01.40	20.965	0.014	20.943	0.029	17.882	0.100	77.8
81	2 25 37.20	−4 21 32.69	19.196	0.003	19.561	0.008	17.546	0.138	88.6
82	2 25 37.56	−4 20 50.21	20.887	0.013	20.542	0.020	17.736	0.098	77.1
83	2 25 37.56	−4 54 40.21	20.399	0.009	20.266	0.018	17.658	0.089	94.4
84	2 25 39.36	−4 22 27.77	20.889	0.013	20.887	0.028	17.756	0.136	71.4
85	2 25 55.56	−4 39 18.07	20.660	0.010	20.770	0.028	18.086	0.158	80.6

Table B.1: Continued.

Appendix C

CTIO spectroscopic proposal

This appendix describes a proposal submitted in March 2006, for a spectroscopic verification of the QSO candidates using Hydra, a multi–object spectrograph mounted on the 4 m CTIO Blanco telescope. The Time Allocation Committee unfortunately rejected the proposal, with the following justification:
The panel found this to be a solid, but not compelling proposal. There were other proposals for similar science that were ranked higher.

C.1 Scientific Objective

In order to study the properties of the quasar candidates, a spectroscopic proposal has been submitted in March 2006. The relatively high object density, 85 candidates spread over $\sim$0.8 degrees^2, limits significantly the number of spectrographs adequate for our scientific needs. For the typical field of view of most spectrographs, of the order of 7×7 arcmin^2, this would correspond to $\sim$1 target per pointing. Such an object density hampers considerably the feasibility of our proposal, due to the total time needed for observing the 85 candidates. One of the few large FoV, multi–object spectrographs available is Hydra, mounted on the 4 m Blanco telescope, at CTIO observatory. Hydra has a field of view of $40'$ in diameter and the possibility of simultaneously observing up to $\sim$140 objects.

The purpose of the proposal is two–fold:

- identify the real quasars in the sample, together with the contaminants. Although the quasar selection process has been carefully designed and implemented, it is quite probable that the candidate sample suffers from contamination effects, due to the presence of:
  - low temperature stars with colors similar to the quasar colors and, or,
  - bulge–dominated galaxies, with galaxy–like colors but point–like morphology.

Redshift	$\lambda_{min} - \lambda_{max}$ (Rest–frame)	Em. Lines
0.5	2670 – 5300 Å	Mg II, [O III]
1.5	1600 – 3200 Å	C III], Mg II
2.5	1140 – 2285 Å	Lyα, N V, Si V, C IV, C III]
3.5	888 – 1777 Å	O VI, Lyα, N V, Si V, C IV

Table C.1: Typical emission lines falling in the wavelength window of the KPGL2 grating for three redshift bins, shifted to the observer's rest–frame.

Spectroscopic observations will not only allow to identify the quasars, but also reveal the spectral type of the contaminants. This information will allow to evaluate the effectiveness of the KX–selection, the filter combination and study possible caveats.

- measure the corresponding redshifts. Thanks to this information it will be possible to shift in the rest–frame the quasar SEDs, in order to model the effect of dust on red and reddened quasars.

C.2 Technical Description

For the spectroscopic verification of the QSO candidates, we proposed to use the KPGL2 grating combined with the 200μm slit plate. This configuration has a high efficiency between $4000 - 8000$Å and a resolution of $R \sim 1155$ at 5500 Å. Second order contamination, by blue light, in the grating's sensitive wavelength range, will be eliminated using the # GG385 filter. Table 1 shows the typical emission lines (column 3) we intend to use for the quasar identification, for four redshift bins (column 1). Column 2 shows the rest–frame ($\lambda/(1+z)$) wavelength range corresponding to the $4000 - 8000$ Å of the proposed configuration.

The Atmospheric Dispersion Compensator on Hydra makes it possible to observe at airmasses as low as 1.6. Performing the observations in October/November permits to obtain six 1–hour pointings per field, leaving sufficient time for calibration purposes (e.g. to observe spectro–photometric standard stars at the beginning and at the end of the night for measuring the response of Hydra's fibers and for the wavelength calibrations). The limiting magnitude of the quasar candidates in the R, z' and K_s band is 21.5, 21.2 and 18.0, respectively. Assuming the above configuration, a total integration time of six hours will give a S/N ≈ 15 for the continuum emission for the faintest of our sources. Since we expect reddened QSOs to still show emission lines, the spectroscopic completeness of our sample will be limited by the number of emission lines present in the obtained spectra.

Given the field of view of Hydra and the distribution of the quasar candidates on the sky, 75 objects can be covered with two pointings. Another two pointings should have to

be dedicated for observing the remaining 10 objects, we decided to omit these candidates from the target list. As a result, we finally asked for two nights (one pointing per night) in order to carry out our project. Fig. C.1 shows the distribution in the RA–Dec plane of the 85 sources. The two circles indicate the field of view of the two Hydra pointings, centered appropriately in order to observe 75 targets. Based on their R–band photometry, the candidates are plotted as filled squares ($17.8 < R \leq 19.6$), open circles ($19.6 < R \leq 21.5$) and filled triangles ($21.5 < R \leq 22.7$). Although 10 candidates are quite faint, it might be possible to measure their redshifts using some of the emission lines. 14 of the 75 targets are common in both pointings. Therefore, despite their relatively faint magnitude, we hope to be able to measure their redshift without any problem.

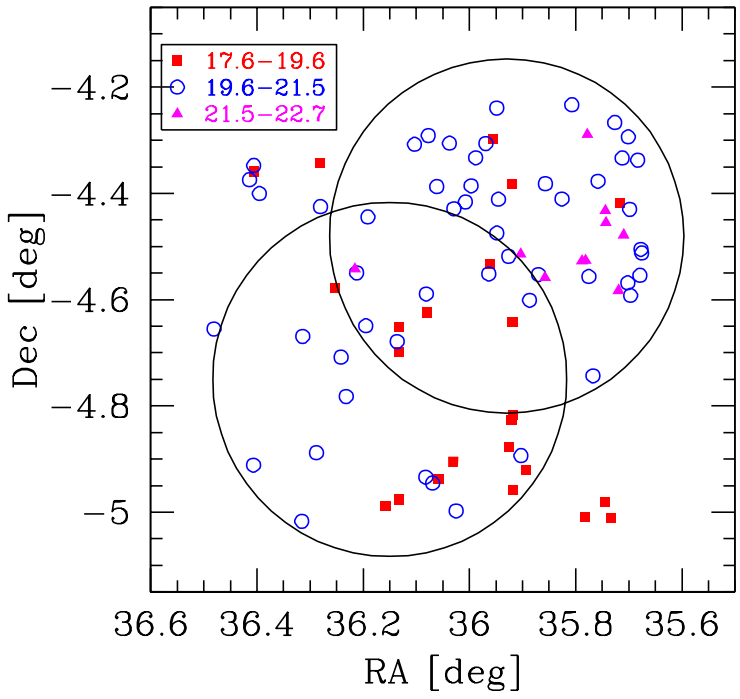

Figure C.1: Distribution in the RA–Dec plane of the 85 QSO candidates. The two circles denote the field of view of the Hydra pointings, centered appropriately in order to observe 75 targets in two nights. According to their R–band photometry, the objects are plotted in different symbols (see text for details).

Appendix D

List of additional AGN & QSO candidates

Table D.1 presents the photometry and astrometry of two additional quasar candidates found in the $Rz'K_s$ catalog. These two objects were not included in the KX–selected sample of QSO candidates.

No	RA (J2000)	Dec (J2000)	R	σ_R	z'	$\sigma_{z'}$	K_s	σ_{K_s}	DP (%)
	h m s	° ′ ″							
S01	2 24 24.16	−4 32 29.62	18.914	0.002	18.931	0.005	17.400	0.146	72.2
S02	2 23 52.19	−4 30 31.50	19.847	0.005	19.719	0.009	17.820	0.099	86.1

Table D.1: Photometry and astrometry for two additional QSOs, not included in the initial KX–selected sample.

.

No	RA (J2000) h m s	Dec (J2000) ° ′ ″	R	σ_R	z'	$\sigma_{z'}$	K_s	σ_{K_s}	DP (%)
1	2 22 53.62	−4 29 28.97	19.144	0.003	18.885	0.004	15.971	0.024	82.4
3	2 23 36.64	−4 17 35.48	20.304	0.008	20.125	0.013	17.865	0.112	80.6
121	2 23 32.17	−4 57 39.38	20.083	0.006	20.236	0.015	16.965	0.045	88.6
4	2 23 22.02	−4 57 38.09	20.791	0.011	20.152	0.014	16.927	0.043	85.7
6	2 24 04.07	−4 51 18.54	21.217	0.017	21.105	0.032	17.618	0.092	94.4
7	2 23 49.60	−4 37 24.82	21.340	0.018	21.048	0.031	17.786	0.067	97.1
9	2 24 32.72	−4 48 16.63	22.098	0.037	21.245	0.036	17.844	0.118	91.4
11	2 25 15.56	−4 56 52.51	21.918	0.034	21.152	0.040	17.545	0.070	94.4
16	2 23 07.19	−4 37 02.75	21.838	0.029	20.928	0.029	17.345	0.059	85.7
23	2 25 22.84	−4 26 47.94	21.878	0.033	20.789	0.026	17.892	0.108	83.3
25	2 25 41.99	−4 24 40.82	20.378	0.008	19.936	0.012	16.904	0.035	96.9
27	2 22 50.34	−4 22 53.76	21.910	0.032	21.000	0.029	17.520	0.093	80.6
28	2 25 38.17	−4 21 08.86	20.861	0.013	20.689	0.023	17.810	0.108	74.3
86	2 23 52.08	−4 58 51.82	22.359	0.046	21.613	0.051	17.735	0.086	88.9
117	2 25 55.02	−4 57 46.91	22.592	0.064	21.926	0.079	17.822	0.085	94.4
122	2 24 13.64	−4 57 23.76	21.481	0.021	20.955	0.028	17.431	0.066	88.6
123	2 25 43.93	−4 57 20.16	22.812	0.078	21.385	0.049	18.366	0.124	91.7
129	2 23 33.68	−4 56 26.99	21.403	0.019	20.643	0.021	17.396	0.063	94.3
137	2 24 02.70	−4 52 44.65	22.508	0.054	21.985	0.071	18.214	0.146	88.9
206	2 23 26.27	−4 49 10.45	21.962	0.032	21.071	0.032	17.593	0.094	86.1
291	2 24 51.77	−4 43 15.28	22.428	0.055	21.593	0.061	17.631	0.102	71.4
304	2 23 53.99	−4 42 32.47	22.573	0.056	21.540	0.048	17.769	0.100	94.3
415	2 22 55.52	−4 37 48.22	22.461	0.051	21.635	0.055	18.121	0.142	77.8
419	2 25 47.03	−4 37 40.62	22.639	0.067	22.000	0.090	18.643	0.173	83.3
433	2 22 56.86	−4 37 05.48	22.474	0.052	21.837	0.066	18.238	0.105	77.8
435	2 22 57.68	−4 36 59.90	22.957	0.084	21.876	0.067	17.789	0.089	75.0
512	2 25 40.76	−4 32 25.98	22.297	0.048	21.093	0.033	17.944	0.117	91.7
566	2 25 15.17	−4 29 36.42	22.260	0.046	21.413	0.044	18.147	0.128	83.3
592	2 23 01.57	−4 25 43.18	22.352	0.048	21.647	0.053	17.989	0.146	77.8
628	2 23 31.09	−4 23 26.63	22.028	0.035	21.607	0.050	18.161	0.108	80.6
769	2 23 55.14	−4 17 05.75	23.005	0.089	21.608	0.050	17.877	0.124	88.9
770	2 22 52.64	−4 16 59.88	22.430	0.051	21.815	0.063	17.994	0.172	83.3

Table D.2: Photometry and astrometry for 32 AGNs found among the sources with extended morphology in the $Rz'K_s$ catalog. From the shape of their spectral energy distribution, the first three objects were characterized as Type–1 AGNs, the rest as Type–2.

.