MASARYKOVA UNIVERZITA
PŘÍRODOVĚDECKÁ FAKULTA
Ústav teoretické fyziky a astrofyziky

BAKALÁŘSKÁ PRÁCE

BRNO 2020 KRISTÍNA KALLOVÁ

MASARYKOVA
UNIVERZITA

PŘÍRODOVĚDECKÁ FAKULTA
Ústav teoretické fyziky a astrofyziky

Hráškové galaxie v rentgenovém oboru

Bakalářská práce

Kristína Kallová

Vedoucí práce:
RNDr. Jiří Svoboda Ph.D.

Brno 2020

Bibliografický záznam

Autor: Kristína Kallová
Přírodovědecká fakulta, Masarykova univerzita
Ústav teoretické fyziky a astrofyziky

Název práce: Hráškové galaxie v rentgenovém oboru

Studijní program: Fyzika

Studijní obor: Astrofyzika

Vedoucí práce: RNDr. Jiří Svoboda Ph.D.

Akademický rok: 2019/2020

Počet stran: viii + 48

Klíčová slova: hvězdotvorná galaxie, trpasličí galaxie, rentgenové dvojvězdy, XMM-Newton, rentgenová spektroskopie

Bibliographic Entry

Author: Kristína Kallová
Faculty of Science, Masaryk University
Department of theoretical physics and astrophysics

Title of Thesis: Green Pea galaxies in X-rays

Degree Programme: Physics

Field of Study: Astrophysics

Supervisor: RNDr. Jiří Svoboda Ph.D.

Academic Year: 2019/2020

Number of Pages: viii + 48

Keywords: starburst galaxy, dwarf galaxy, X-ray binaries,
XMM-Newton, X-ray spectroscopy

Abstrakt

V prvotních galaxiích probíhala velmi silná tvorba hvězd vedoucí ke vzniku početných populací rentgenových dvojhvězd. Je možné, že energetické záření těchto mladých objektů v raných galaxiích mohlo významně přispět k reionizaci vesmíru. Přímé studium těchto vzdálených hvězdotvorných galaxií je však velmi obtížné, protože jsou příliš slabými zdroji. Existují ale i lokální galaxie, jejichž vlastnosti se raným galaxiím podobají. Jejich studium nám může pomoci objasnit, jak význačný byl příspěvek záření související s velmi silnou tvorbou hvězd k zahřátí mezigalaktické látky v éře reionizace. Jedná se o kompaktní galaxie s velkou tvorbou hvězd, známých jako hráškové galaxie.

Nedávno byly publikovány výsledky rentgenových pozorování tří hráškových galaxií pomocí vesmírného rentgenového dalekohledu *XMM-Newton* v práci Svoboda a kol. (2019). Zde autoři naměřili zvýšenou rentgenovou luminozitu u dvou zdrojů, zatímco u třetího zdroje nebyla zaznamenána významná detekce. V této práci provádíme re-analýzu všech tří pozorování a analýzu nově získaného pozorování třetí nedetekované hráškové galaxie. Ani nové delší pozorování nepřineslo prokazatelnou detekci rentgenového záření z této galaxie. Horní limit jejího toku záření vychází o více než řád nižší než bylo naměřeno u dvou detekovaných hráškových galaxií. To naznačuje, že přebytek rentgenového záření v prvních dvou zdrojích může být jedinečný a souviset například s akrecí hmoty na centrální černou díru, ke které nedochází v třetí studované galaxii.

Abstract

Early galaxies were experiencing very strong star-formation leading to the creation of numerous X-ray binary populations. These young objects in the first galaxies may have significantly contribute to the reionization of the Universe by their energetic emission. However, due to their faintness, the direct study of these distant star-forming galaxies is very difficult. Yet, there are also local galaxies whose properties are similar to the early galaxies. Their study may help us understand how significant was the contribution of the emission associated with strong star-formation to the heating of the intergalactic medium in the Epoch of reionization. These local analogues are compact starburst galaxies known as Green Peas.

Recently, the results of X-ray observations of three Green Pea galaxies, studied by X-ray space observatory *XMM-Newton*, were published in the work by Svoboda et al. (2019). The authors measured increased X-ray luminosity at two studied sources, whereas no significant detection was recorded at the third source. In this work, we performed a reanalysis of all three archival X-ray observations of Green Peas and an analysis of the newly acquired observation of the third, undetected Green Pea galaxy. However, even the new longer observation did not provide the detection of the X-ray emission from this galaxy. The upper limit of its X-ray flux is more than an order lower than that measured in the two detected Green Pea galaxies. This suggests, that the X-ray excess in the first two sources may be unique and could be related, for example, to the accretion of matter onto a central black hole, which does not occur in the third studied galaxy.

MASARYKOVA UNIVERZITA
PŘÍRODOVĚDECKÁ FAKULTA
KOTLÁŘSKÁ 2, 611 37 BRNO
IČ: 00216224
DIČ: CZ00216224

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Akademický rok: 2019/2020

Ústav:	Ústav teoretické fyziky a astrofyziky
Studentka:	Kristína Kallová
Program:	Fyzika
Obor:	Astrofyzika

Ředitel *Ústavu teoretické fyziky a astrofyziky* PřF MU Vám ve smyslu Studijního a zkušebního řádu MU určuje bakalářskou práci s názvem:

Název práce:	Hráškové galaxie v rentgenovém oboru
Název práce anglicky:	Green Pea galaxies in X-rays
Jazyk závěrečné práce:	angličtina

Oficiální zadání:
Hráškové galaxie jsou kompaktní trpasličí galaxie s nízkou metalicitou a vysokou tvorbou hvězd. Svými vlastnostmi připomínají rané galaxie, které svou intenzivní tvorbou hvězd mohly přispět k tzv. reionizaci vesmíru. U několika hráškových galaxií se potvrdil únik ionizovaného záření v UV oboru, a proto jsou intenzivně studovanými zdroji v optické a UV oblasti, ale o jejich rtg záření jsme toho zatím moc nevěděli. Nedávná pozorování tří hráškových galaxií v rtg oboru pomocí družice ESA XMM-Newton přinesla zjištění, že u dvou ze tří zdrojů je uvolněného rtg záření asi 6x více, než by se předpokládalo pouze z tvorby hvězd. Naopak u třetího zdroje nebyla detekce v rentgenovém oboru vůbec zaznamenána a v lednu jsme získali nové pozorování s delší expozicí. Studentka se ve své práci naučí redukci dat s XMM-Newton a důkladně zanalyzuje rentgenová pozorování všech tří zdrojů se zaměřením na určení přesnosti naměřeného toku záření s ohledem na možné fluktuace kosmického rentgenového pozadí. Zpracuje nová měření třetího zdroje a zkombinuje nové pozorování s předchozím za účelem určení toku rentgenového záření z tohoto zdroje nebo alespoň zpřesnění jeho horního limitu. V práci také shrne možná fyzikální vysvětlení původu rentgenového záření v těchto galaxiích.
Literatura

- Svoboda J. et al., 2019, submitted to ApJ
- Ward-Thompson D. & Whitworth A.: An Introduction to Star Formation, 2011, Cambridge University Press, ISBN 978-0-521-63030-6
- Basu-Zych A., Lehmer B., Hornschmeier A. et al., 2013, ApJ, 774, 152
- Brorby M., Kaaret P., Prestwich A. & Mirabel F., 2016, MNRAS, 457, 4081
- Cardamone C., Schawinski K., Sarzi M. et al., 2009, MNRAS, 399, 1191
- Izotov Y., Orlitová I., Schaerer D. et al., 2016, Nature, 529, 178
- Mineo S., Gilfanov M. & Sunyaev R., 2012, MNRAS, 419, 2095

Vedoucí práce:	RNDr. Jiří Svoboda, Ph.D.
Datum zadání práce:	13. 3. 2019
V Brně dne:	31. 1. 2020

Souhlasím se zadáním (podpis, datum):

.....	.....	.....
Kristína Kallová **studentka**	RNDr. Jiří Svoboda, Ph.D. **vedoucí práce**	prof. Rikard von Unge, Ph.D. **ředitel Ústavu teoretické fyziky a astrofyziky**

Poďakovanie

Na tomto mieste by som chcela poďakovať v prvom rade mojej rodine, ktorá ma podporovala počas celej doby štúdia ako aj počas písania záverečnej práce a zaháňala moje pochybnosti. Špeciálne ďakujem dedovi, za jeho neutíchajúci záujem o tému mojej práce a ochotu konzultovať jej gramatickú aj obsahovú spránku. Obrovská vďaka patrí tiež Zdendovi, za to, že trpezlivo odpovedal na moje ustavičné otázky najmä ohľadom LaTeXu a pomáhal mi s hľadaním riešení a správnych odpovedí. Za pomoc ohľadom SASu a časté konzultácie vďačím Peterovi Boormanovi a v neposlednej rade patrí moja vďaka samozrejme vedúcemu práce Jiřímu Svobodovi, ktorý pre mňa otvoril nové okno do astronómie vysokých energií.

Prohlášení

Prohlašuji, že jsem svoji bakalářskou práci vypracovala samostatně pod vedením vedoucího práce s využitím informačních zdrojů, které jsou v práci citovány.

Brno 17. srpna 2020 .....

Kristína Kallová

Contents

Introduction

Reionization of the Universe is an important epoch in the history of the Universe, when the neutral matter in the intergalactic medium was heated and ionized by energetic photons. However, the source of this energetic radiation is not yet clearly determined. Early galaxies with massive ongoing star-formation could significantly contribute to the heating of the young Universe as they might have experienced leakage of the ultraviolet Lyman-continuum photons from the young massive stars and hot gas. These distant galaxies with redshift $z > 2$ are usually referred to as Lyman break galaxies or Lyman α emitters, due to the observational technique used for their classification.

To study these high-redshift galaxies a large observational time would be required, yet there would remain an issue with the spatial resolution. To avoid these complications, we can focus our study to local analogues to the distant star-forming galaxies. So-called Lyman break analogues are galaxies in the local Universe ($z < 0.2$) that have similar properties with the hight-redshift galaxies such as their masses, star-formation rates or metallicities. By studying the local analogues to early galaxies we can gain insight into the processes that were behind the leakage of the ionizing radiation.

Recently discovered so-called Green Pea galaxies are possibly a subsample of the Lyman break analogues. They are compact dwarf galaxies with high star-forming rates and powerful [O III] λ 5007 Å optical emission line with large equivalent widths responsible for their specific green colour if originated from redshift range $0.112 \lesssim z \lesssim 0.360$. In this work we study the X-ray properties of three Green Pea galaxies supposedly originating from high-mass X-ray binaries. The high energy photons and particles provide radiative and mechanical feedback resulting in ionization of the gas surrounding the X-ray sources and therefore creating such interstellar medium conditions leading to the escape of Lyman continuum photons. We study the X-ray emission from galaxies with possible Lyman continuum leakage to find out whether their X-ray luminosities correspond to the trends established for other star-forming galaxies, or if additional sources such as active galactic nuclei need to be considered.

1 | Star-forming galaxies

1.1 Star formation

Star formation is an important and complex process in the Universe leading to ignition of new stars. Many studies are dedicated to the topic of this astrophysical process. In this section we briefly summarize the main principles of the formation of new stars following the works by Mikulášek et al. (2005) and Ward-Thompson et al. (2011).

A star formation in a galactic context can be described as a conversion of diffuse interstellar gas clouds into star clusters. A collision of two clouds of interstellar gas is extremely inelastic and dissipative, as the free path between the individual gas particles is very short. In such case, the gas is first heated by the collision and than cooled by emitting radiation. On the other hand, if two star clusters collide, the free path between individual stars is extremely long, their kinetic energy is conserved so the collision is non-dissipative and elastic.

Stars are formed mostly in those regions of interstellar medium, where the gas is predominantly molecular, therefore in giant and relatively cold molecular clouds with typical temperatures of $T \sim 10-40$ K, according to observations. These clouds are formed mostly by molecular hydrogen H_2, also by atomic hydrogen H, helium He and other elements. An important part of molecular clouds are also dust grains, representing about 1% of the total mass of the interstellar media, as the molecular hydrogen forms on their surface. Dust grains also act as a shield on the surface of the clouds, absorbing the high energy photons and radiating the redundant energy in form of heat away, keeping the central parts of the molecular clouds cold. Photons emitted by the grains can have a broad range of wavelengths, however, in range about $10-10^3$ μm dust emission dominates the radiation from the interstellar medium.

If a star is to form in the molecular cloud, it must become gravitationally unstable and consequently collapse. Any fluctuations of density in such cloud may result in condensation due to its self-gravity. On the other hand, its internal pressure based on irregular thermal motion of the molecules forces it to expand back to

the same density as the surrounding medium. Whether the random fluctuation collapses or not depends upon its size and equivalent minimum initial mass. For simplicity, considering a spherical fluctuation formed primarily by hydrogen with uniform density ρ, mass M and temperature T, the self-gravity overcomes internal pressure and the fluctuation condenses out if its minimum initial mass is larger than the Jeans mass $M_{\rm J}$

$$M > M_{\rm J} = \sqrt{\frac{3}{4\pi\rho}\left(\frac{5kT}{G\mu m_{\rm H}}\right)^3}. \tag{1.1}$$

Here k is the Boltzmann's constant, G is the gravitational constant, μ is the mean molecular weight and $m_{\rm H}$ is the hydrogen mass. Resulting from the formula 1.1, the stars are formed only in the coolest and densest regions of the molecular clouds in the interstellar media.

1.1.1 High mass stars

Massive stars with initial mass of about $M_{\rm i} > 11\ M_\odot$ form probably in very dense and massive infrared dark clouds (Ward-Thompson et al., 2011). Their dark molecular cores are cold, with temperatures of $T \sim 10-20$ K as the dust grains shield the center of the clouds and radiate the heat away. Once the process of star-formation begins, the radiation from the contraction of newly forming stars heats the molecular cores up to $T \sim 100-200$ K. When the stars are sufficiently hot and luminous they begin to ionise the matter surrounding them. The energetic ionising photons normally arise from O-type and B-type stars and as the hydrogen is the most abundant element, these regions of photo-ionised gas surrounding the recently formed high-mass stars are known as H II regions.

The lifetime of a newborn star is determined by its initial mass; the more massive the star is, the more rapidly it evolves. For example, the main-sequence lifetime for a star with initial mass of 11 $M_\odot$, eg. time when hydrogen is burning in its core is about 10^7 years. Once the 95% of hydrogen in the core is burned out, the core gets denser and warmer and gradually the conditions for heavier elements burning up to iron are fulfilled. The star passes through the Red Giant Branch and consecutive the Asymptotic Giant Branch in the Hertzsprung-Russell diagram. When the core is made of iron, it collapses into neutron star if its mass is higher than 1.44 $M_\odot$, the Chandrasekhar limit, or black hole, if its heavier than about 3 $M_\odot$.

1.2 Starburst galaxies

Starburst galaxies display excess of star-formation activity in localized regions, typically in and around the centre of the galaxy where stars and star clusters are being produced in very large numbers. The burst of star-formation is driven by high rate of infall of gas towards these regions, triggered possibly by galaxy interactions as galaxy collisions or mergers (Ward-Thompson et al., 2011; Schneider, 2015). These galaxies usually show excess emission at higher energies such as ultraviolet (UV) and soft X-ray radiation, as this emission is typical for very young high-mass stars with short evolutionary time scales. Some starbursts can also be extremely luminous in the far-infrared part of electromagnetic spectrum due to the dust emission. The dust absorbs the energetic UV radiation produced by young massive stars and re-emits this energy in the form of thermal radiation in far-infrared. These objects with highly obscured star-forming regions are presented as Luminous infrared galaxies and may be ones of the most energetic nearby star-forming galaxies (Lehmer et al., 2010). Another significant subsample of the star-forming galaxies are Ultra-luminous compact X-ray sources studied by the *Chandra* X-ray observatory. We discuss the origin of X-ray photons in such galaxies in section 1.4.

To understand how many stars are being produced within a galaxy, a star-formation rate (SFR) is defined as a total mass of stars given in solar masses, formed per year. For star-forming galaxies typical values of SFR are $\sim 10-300\ M_{\odot}/\mathrm{yr}$ (Schneider, 2015).

1.2.1 Similarity with high-redshift galaxies

Local star-forming galaxies with high SFR are often dusty with only a few percent of the UV emission escaping (Basu-Zych; Lehmer; Bret, et al., 2013). However, we are interested in the rare dust-poor starbursts with possible UV radiation leakage into the intergalactic media. The main motivation in study of local dust-poor starbursts is their similarity to high-redshift star-forming galaxies that could significantly contribute to the heating and reionization of the early Universe.

Interstellar and intergalactic medium absorbs ultraviolet radiation at 912 Å, so-called Lyman edge, as this wavelength corresponds to the ionization energy of the most abundant element, hydrogen in its ground state (Schneider, 2015). Observing a distant galaxy emitting photons of Lyman continuum that is not experiencing much reddening by dust shows a break in observed spectrum at this wavelength, as photons with wavelength of 912 Å and shorter are absorbed by neutral hydrogen. This method called the Lyman break technique is very efficient when searching for high-redshift ($z \gtrsim 2.5$) actively star-forming galaxies. Many studies exploring such objects have found that these galaxies, often called Lyman break galaxies, have very high star-formation rates, are dust-poor and metal-deficient (Basu-Zych;

Lehmer; Bret, et al., 2013), but their direct observation is impractical due to large exposure times that would be required. An alternative method in exploring the possible source of ionizing radiation in the early Universe is to study so-called Lyman break analogues, a group of local galaxies ($z < 0.3$) that strongly resemble the high-redshift Lyman break galaxies (Heckman et al., 2005; Hoopes et al., 2007; Overzier et al., 2008; Basu-Zych; Gonçalves, et al., 2009).

Lyman break analogues are galaxies selected to match the far-UV properties with distant Lyman break galaxies (Basu-Zych; Lehmer; Bret, et al., 2013). They are compact, starburst galaxies with low gas-phase metallicities and little dust attenuation (Basu-Zych; Lehmer; Bret, et al., 2013; Basu-Zych; Lehmer; Fragos, et al., 2016). Kunth et al. (2000) suggested, that objects called the Blue compact dwarfs, later studied by Brorby; Kaaret; Prestwich (2014), may be analogues to early galaxies. Study of a large sample of star-forming luminous compact galaxies by Izotov et al. (2011) claims that newly discovered so-called Green Peas (Cardamone et al., 2009) are together with Blue compact dwarf galaxies a subset of luminous compact emission-line galaxies at different redshifts, therefore possibly many of these objects may be found to be Lyman break analogues. Discovery of Green Peas and their properties are discussed is section 1.3.

1.3 Green Peas

Green Pea galaxies were originally discovered by the Galaxy Zoo[1] project, where more than 200 000 volunteers classified optical images from the Sloan Digital Sky Survey (York et al., 2000, hereafter SDSS) main galaxy sample. The SDSS provides photometry of millions of objects from a quarter of the sky in filters *u, g, r* and *i*. In *gri* composite colour images, Green Peas appeared as unresolved greenish point-like sources (see Figure 1.1) with galaxy-type spectral features. Their colour is driven by a very powerful [O III] λ 5007 Å optical emission line with large equivalent widths, that is redshifted into the SDSS *r* band of these *gri* composite images, which is represented by green colour.

Approximately 100 Green Peas identified by the Galaxy Zoo volunteers, taken from the SDSS Data Release 7 spectroscopic sample, were compared to a sample of 10 000 galaxies and 9 500 quasi-stellar objects by Cardamone et al. (2009). The comparison sample was selected to match the Green Pea redshifts in range $0.112 < z < 0.360$. Using colour selection criteria Cardamone et al. (2009) obtained a sample of 251 Green Peas with extreme [O III] equivalent width taken from all SDSS spectroscopic galaxies. Spectral analysis limited the sample to those objects with a signal to noise ratio larger than 3 in the continuum near the

[1] https://www.zooniverse.org/projects/zookeeper/galaxy-zoo

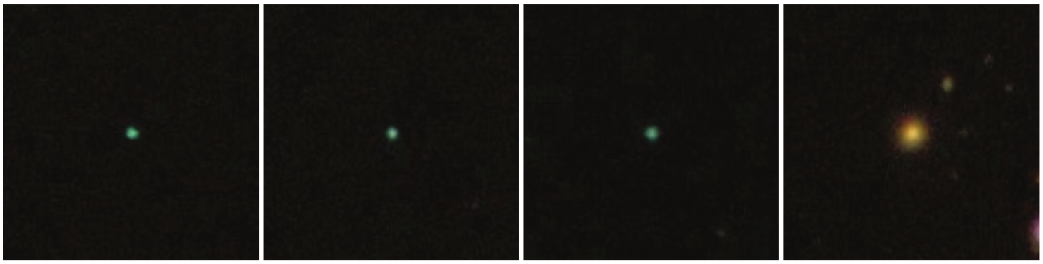


Figure 1.1: Three examples of Green Pea galaxies (three images from left) from *gri* composite colour 50×50 arcsec^2 SDSS images. The green colour represented by *r* filter is clearly dominated. Due to their compact morphology and green appearance Green Peas are easily distinguishable from the classical elliptical galaxy shown in the last image to the right. All objects shown in this figure are at $z \sim 0.2$. Image adapted from Cardamone et al. (2009).

Hα and Hβ regions and in each of the emission lines Hα, Hβ, [O III] λ 5007 Å and [N II] λ 6583 Å.

To separate the purely star-forming objects from active galactic nuclei (AGN) a classical emission line diagnostic devised by Baldwin et al. (1981, hereafter BPT) was applied. A theoretical line of maximum starburst contribution, calculated by Kewley et al. (2001, labelled Kew01 in Fig. 1.2) separates star forming galaxies from AGN. Empirical shift of this line by Kauffmann et al. (2003, labelled Kauf03 in Fig. 1.2) separates purely star-forming objects from composite galaxies. BPT diagram compares ratio of forbidden emission lines, which are excited mainly by AGN, to nearby Balmer lines (Basu-Zych; Lehmer; Hornschemeier, et al., 2012). By comparison of ratios [O III] λ 5007 Å/Hβ to [N II] λ 6583 Å/Hα with model lines of Kewley et al. (2001) and Kauffmann et al. (2003) the remaining 103 narrow-line objects were classified. The majority of the sample turn out to be star-forming (80 objects) with 10 Seyfert 2 galaxies and 13 transition objects lying between the lines Kew01 and Kauf03. The purely star-forming objects turn out to be low-mass galaxies ($M_* \sim 10^{8.5} - 10^{10}\ M_\odot$) with the highest specific star-formation rates (sSFR = SFR/M_*) seen in the local universe (up to $\sim 10^{-8}\,\mathrm{yr}^{-1}$) and are similar in size, mass, luminosity, and metallicity to luminous Blue compact dwarf galaxies and therefore to the high-redshift luminous Lyman break galaxies and Lyman α emitters.

The range of redshifts initially set to $0.112 < z < 0.360$ by Cardamone et al. (2009) was extended to $0.02 < z < 0.63$ by Izotov et al. (2011) who selected luminous compact emission-line galaxies from Data Release 7 of the SDSS on the basis of both their spectroscopic and photometric properties, obtaining a sample of 803 star forming galaxies. They claim that Green Peas are a subsample of

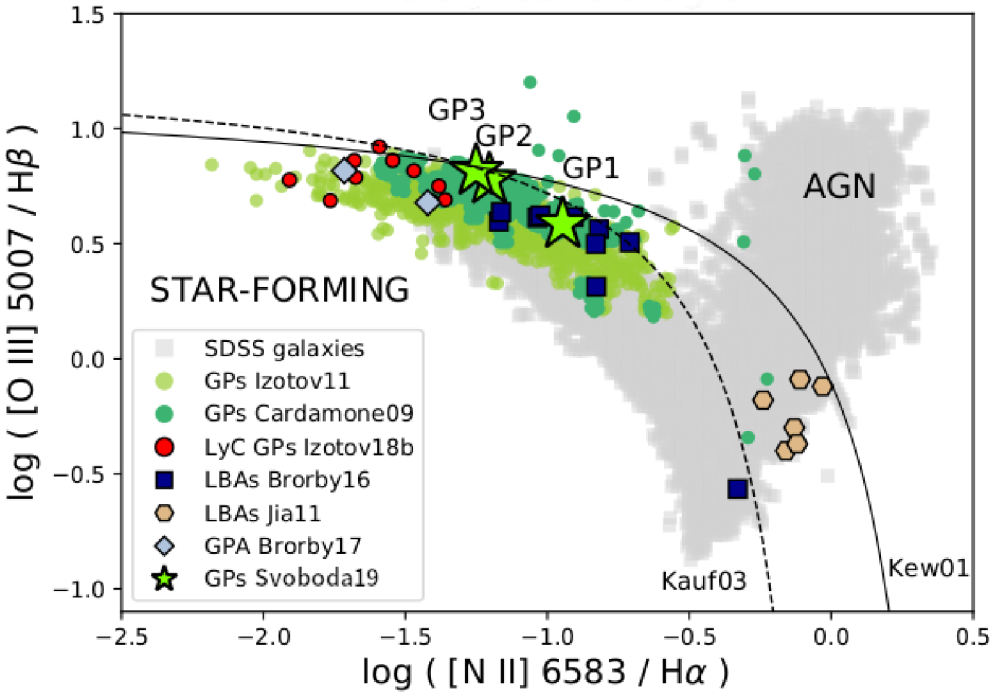


Figure 1.2: BPT diagram for emission-line galaxy classification showing studied Green Peas in comparison with galaxy samples studied in other works. Three Green Peas studied in work by Svoboda et al. (2019) were classified as purely star-forming objects, yet it can be seen that GP2 and GP3 are located closely to the both lines Kauf03 and Kew01. Adapted from Svoboda et al. (2019).

luminous compact galaxies in a determined redshift range, as the colour of an emission-line galaxy on the *gri* SDSS composite images changes with its redshift.

In this work we study three Green Pea galaxies (hereafter GPs) selected by Svoboda et al. (2019) from the parent sample of 80 star-forming galaxies studied by Cardamone et al. (2009). Selected were those objects with the highest star-formation rate among the GPs that were classified as purely star-forming using BPT diagram. The observation in X-rays was performed by *XMM-Newton* satellite that looked at final sample of three targets with expected highest X-ray flux, based on the estimates from the star-formation rate.

The BPT diagnostic diagram shown in Figure 1.2 displays mainly Green Peas galaxies studied by Cardamone et al. (2009) and Izotov et al. (2011) in comparison

with the SDSS galaxies. There are also displayed Lyman break analogues (LBAs) studied by Brorby; Kaaret; Prestwich; Mirabel (2016) and Brorby; Kaaret (2017). The three green star symbols in BPT diagram show position of three Green Peas studied by Svoboda et al. (2019) and hence in this work.

1.4 X-ray emission from star-forming galaxies

The high energy photons such as X-rays are emitted only by very hot and energetic processes in the Universe. Active galactic nuclei are ones of the most luminous X-ray sources, powered by accretion onto compact supermassive black holes. However, in normal galaxies, ergo galaxies without an AGN, there are other sources that can contribute to the X-ray emission, such as X-ray binaries, supernovae and their remnants, hot gas from starburst-driven winds and outflows, or even so-called ultra-luminous X-ray sources, whose nature is still unknown (Basu-Zych; Lehmer; Hornschemeier, et al., 2012). The main contributors to the X-ray flux in normal galaxies are above mentioned X-ray binaries, consisting of a compact remnant such as neutron star or a stellar-mass black hole and a star companion. In star-forming galaxies, the X-ray output is dominated by those binaries, formed by a massive O-type or B-type star accreting onto a compact object, known as high mass X-ray binaries (HMXBs). As these objects are short-lived, their number is expected to scale with SFR (Basu-Zych; Lehmer; Hornschemeier, et al., 2012). On the other hand, the low mass X-ray binaries (LMXBs) trace older stellar population due to their longer evolutionary time scales, so they are expected to dominate the X-ray emission once the young stars are extinct (Basu-Zych; Lehmer; Hornschemeier, et al., 2012). Additionally, hot gas in the state of outflows, driven by the supernovae and winds from massive stars may contribute significantly to the soft X-rays ($E \sim 0.2-2$ keV) of normal star-forming galaxy (Mineo et al., 2012).

1.4.1 High mass X-ray binaries

High mass X-ray binaries are the main contributor to the hard X-rays ($E > 2$ keV) of star-forming galaxies (Basu-Zych; Lehmer; Hornschemeier, et al., 2012). The X-rays are produced in a compact region around the star remnant, the neutron star or a stellar-mass black hole, once the mass transfer begins. Overflow via the inner Lagrangian point occurs if the massive star, called the donor, fills in its Roche lobe and the matter out of external layers of the star falls onto the compact object called the accretor. As the matter carries within most of the angular momentum, the matter does not fall directly onto the accretor, but it forms an accretion disc in the plane of the orbit (Mikulášek et al., 2005). There is another scenario of matter transfer when the donor does not full in the Roche lobe, but when the accreted

matter origins from the stellar winds of the companion star.

Due to friction in the accretion disc the angular momentum is transferred towards the outer parts of the disc allowing the particles to approach the compact star remnant in spiral trajectories, until their trajectories become unstable and they fall directly onto the accretor star. During this process a large amount of energy is released, heating the inner parts of the accretion disc to temperatures up to millions of Kelvins and making the inner part of the disc an powerful X-ray source (Mikulášek et al., 2005).

1.4.2 Dependence on metallicity

The high-mass X-ray binary populations in early star-forming galaxies could play an important role in the reionization of the Universe (Power et al., 2013; Mirabel, I. F. et al., 2011; Jeon et al., 2014; Knevitt et al., 2014). This idea is based on fact, that the high-mass X-ray binary populations are much more numerous and possibly more luminous in low-metallicity environments. The metallicity dependence of the X-ray luminosity of these objects was studied by Douna et al. (2015) on a sample of resolved HMXBs in local galaxies with the high specific star-formation rates and known metallicities and X-ray luminosities. Relation between X-ray luminosity, SFR and metallicity in Lyman break analogues was also studied by Brorby; Kaaret; Prestwich; Mirabel (2016).

These studies proved that HMXB populations are indeed enhanced in low-metallicity galaxies. The HMXB populations were found to be $\sim$10 times larger in low-metallicity ($12 + \log(\mathrm{O/H}) < 8$) galaxies, per unit SFR, than in solar metallicity galaxies (Douna et al., 2015). This result is in agreement with the work by Basu-Zych; Lehmer; Hornschemeier, et al. (2012) and Basu-Zych; Lehmer; Bret, et al. (2013), who studied how the relation between X-ray luminosity and SFR evolves with redshift. The discovery of a mild evolution of the X-ray luminosity per SFR (L_X/SFR increases with redshift) was interpreted as to be driven by the metallicity evolution. At higher redshift, the metallicities are lower and the contribution of HMXBs to X-ray luminosity is higher for a given SFR (Basu-Zych; Lehmer; Hornschemeier, et al., 2012).

There are several arguments suggested to explain the increased ratio of L_X/SFR with decreasing metallicity. Low metallicity massive stars have weaker stellar winds, therefore they loose less mass during their evolution and are more massive on average when they reach core-collapse phase. More massive and numerous black hole companions are produced and since black hole HMXBs can be more luminous than neutron star HMXBs, this results in a more luminous binary population. Wind mass loss also leads to angular momentum loss, leading to an orbit expansion of the binary. Due to weaker wind loss in low-metallicity environments the orbits are tighter on average. This results in more systems that will encounter Roche-lobe

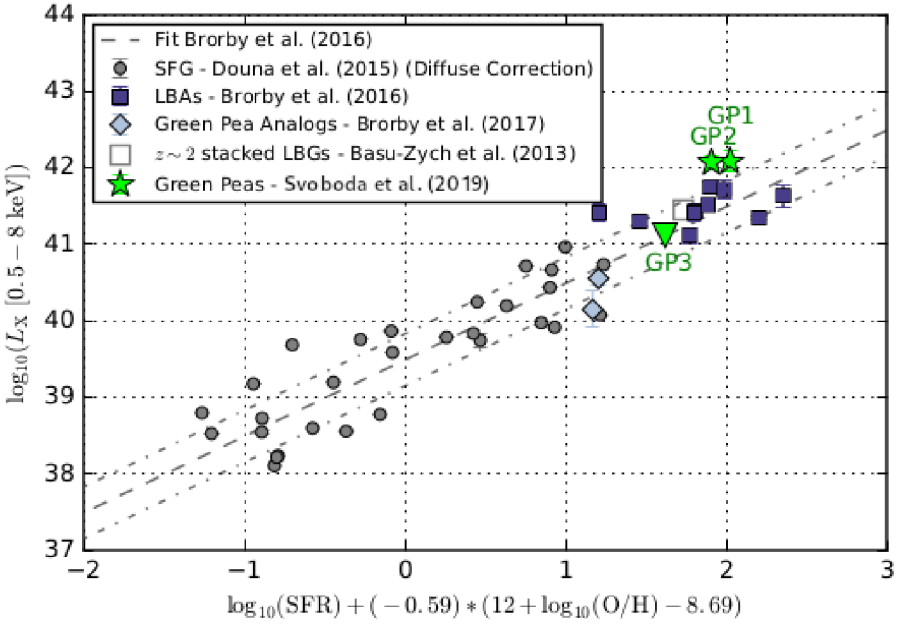


Figure 1.3: Metallicity dependent relation between X-ray luminosity and SFR by Brorby; Kaaret; Prestwich; Mirabel (2016). Figure adapted from Svoboda et al. (2019) shows enhanced X-ray luminosity for two Green Pea galaxies labelled GP1 and GP2 and an upper limit of X-ray luminosity for Green Pea galaxy labelled GP3.

overflow transfer instead of wind accretion and as the Roche-lobe overflow can drive much higher accretion rates, the HMXB population are expected to be more luminous (Basu-Zych; Lehmer; Bret, et al., 2013; Basu-Zych; Lehmer; Fragos, et al., 2016).

Figure 1.3 shows relation between the X-ray luminosity and SFR dependent on the metallicity by Brorby; Kaaret; Prestwich; Mirabel (2016). The diagram shows also the position of three Green Peas studied by Svoboda et al. (2019). The authors found an excess of X-ray flux of $\sim 10^{42}$ erg/s from galaxies labelled Green Pea 1 (GP1) and Green Pea 2 (GP2), but the third studied object labelled Green Pea 3 (GP3) was not detected. They constrained only the upper limit of its X-ray luminosity to 1.3×10^{41} erg/s. This significant difference between the observed galaxies is one of the main motivation for their further study, as the physical properties of all three galaxies are very similar (see Fig. 1.2 where GP2 and GP3 locations in BPT diagram are very similar).

The methodology of our work is described in Chapter 2: the X-ray space observatory and its scientific instruments are described in section 2.1. The methodology of data reduction is described in section 2.2 and the spectral analysis using XSPEC spectral fitting software with model description can be found in section 2.3. The results of our work are given in Chapter 3, where in section 3.1 we present results of reanalysis of the archival observations of Green Pea galaxies and in section 3.2 we present the results from new GP3 observation analysis. We discuss the results in Chapter 4: the different binning approaches are discussed in section 4.1, the dependence on selected background is discussed in section 4.2 and the difference between the studied galaxies is discussed in section 4.3 including possible origin of the enhanced X-ray emission from GP1 and GP2 in subsection 4.3.1.

2 | Methodology

2.1 XMM-Newton observatory

The European Space Agency's X-ray multi-mirror mission was launched by Ariane 504 on December 10th 1999[1]. It is still in operation, the end of the mission was extended until 31st of December 2022. It is a high throughput X-ray spectroscopy mission with great capability to detect X-rays due to its sensitive detectors. The mission name was modified to current *XMM-Newton* later after its launch in honour of Sir Isaac Newton, who is traditionally regarded as the founder of spectroscopy.

The spacecraft consists of several main sections. The 6.80 meters long black telescope tube is made of carbon fibre and is composed of two halves: the upper and lower tubes containing three X-ray telescopes. The schematic of the spacecraft is shown in Figure 2.1.

The mirror support platform within the service module located at the forward broader end of the spacecraft carries the three X-ray mirror modules, the optical monitor and the two star-trackers. There are also two deployable solar arrays with span of 16 meters and the telescope sun shield attached to the the service module.

There are three scientific instruments on board of the spacecraft. The focal plane assembly carries two reflection grating spectrometer (RGS) together with two readout cameras and three European photon imaging cameras (EPIC) at the prime focus of each of the spacecraft's three telescopes. There are two EPIC detectors employing metal oxide semi-conductor (MOS) charge-coupled devices (CCDs) while the third one uses a new type of CCD (PN) built by the Max Planck Institute of Extraterrestrial Physics in Garching and the Astronomical Institute in Tübingen, both in Germany. Each EPIC-MOS detector consists of an array of 7 CCDs which are made up of a matrix of 600×600 pixels, and EPIC-PN detector consists of an array of 12 CCDs made up of a matrix of 64×600 pixels.

Two of the three mirror modules are equipped with reflection grating spectrometer (RGS) array dispersing about 40 per cent of incoming light. With two RGS readout detectors located in secondary focus they form part of the RGS component,

[1] https://sci.esa.int/web/xmm-newton

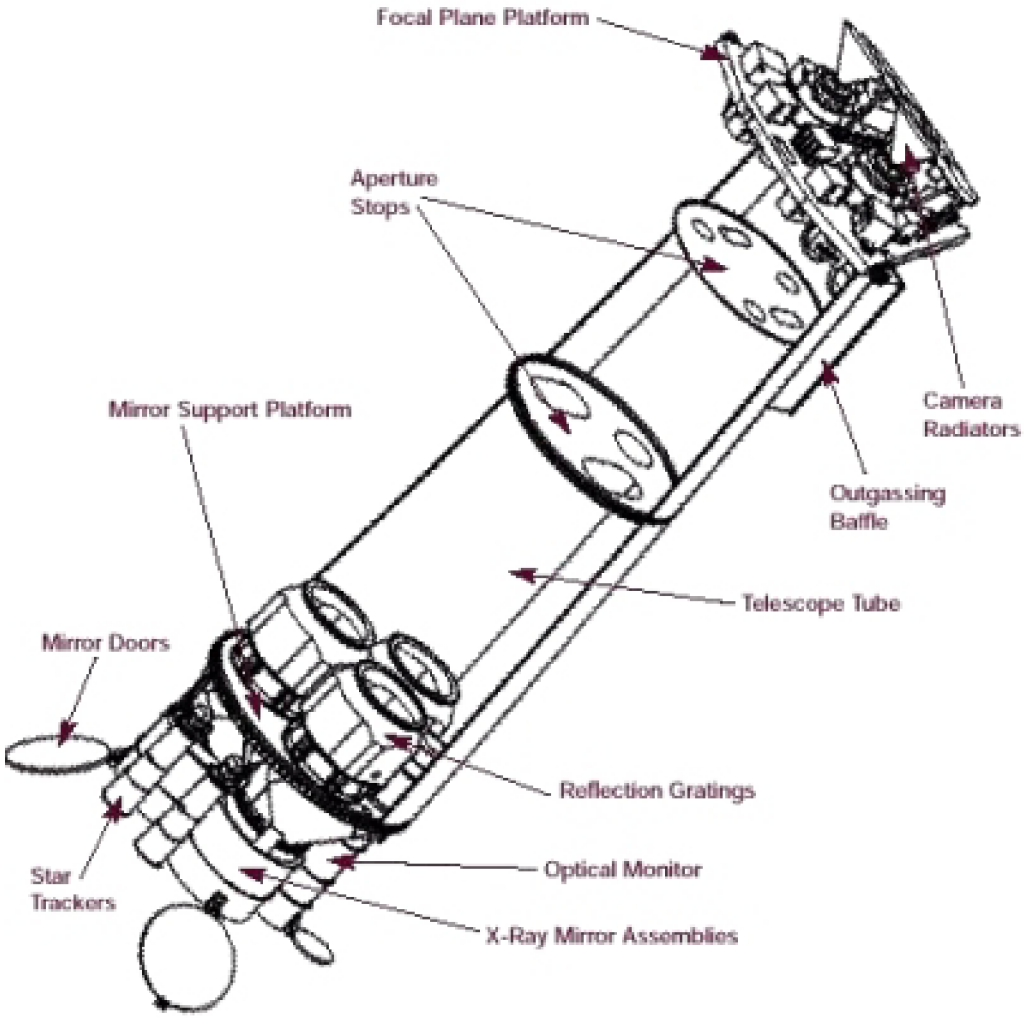


Figure 2.1: *XMM-Newton* schematic adapted from ESA web page[a].

[a] https://sci.esa.int/web/xmm-newton/-/47056-spacecraft

making *XMM-Newton* the very first X-ray telescope to carry such device operating in X-rays.

XMM-Newton is the first X-ray space observatory equipped with very sensitive optical and ultraviolet telescope that gives the mission a multi-wavelength capacity. Optical monitor, the third scientific instrument on board the observatory, can simultaneously observe the same regions as the X-ray telescopes by being co-aligned with them. It is a 30 cm aperture Ritchey-Chrétien telescope with focal length of 3.8 meters, equipped with two CCDs with spectral range of 170 to 650 nanometre.

2.2 Data reduction

The raw data produced by *XMM-Newton* consists of Observation Data Files (ODF) containing raw event science files from all of the scientific instruments on board the spacecraft and the data products generated by the Pipeline Processing System (PPS) at the *XMM-Newton* Science Operations Centre. The access to both sets of ODF and PPS is provided by *XMM-Newton* Science Archive. In order to perform the analysis of the ODF data the set of so called Current Calibration Files (CCF) is necessary and is available for download at the *XMM-Newton* web server [2]. Both ODFs and CCFs are compliant with the Flexible Image Transport standard (FITS).

For data reduction we used the Science Analysis System (Gabriel et al., 2004, hereafter SAS) following the SAS data analysis threads [3]. To process specific ODF files, first we need to generate a Calibration index file (CIF) that points SAS to correct set of CCFs according to a given observation and analysis date. There are three variables that need to be defined while working in SAS: *SAS_CCFPATH*, *SAS_CCF* and *SAS_ODF*. Once we have specified the *SAS_CCFPATH*, which is a path leading to where our CCF files are saved, we used command `cifbuild` to create a CIF file. The path leading to new CIF file was defined in *SAS_CCF* variable. To set the third variable we executed command `odfingest` which takes information from all ODF components and produces a file summarizing all the observational information. The path to this new file was set as the *SAS_ODF* variable.

We proceeded with data processing in SAS by generating the calibrated and concatenated EPIC event lists with `emproc` task for MOS and `epproc` for PN detectors. Event lists can be directly used to generate light curve, images or spectra. We continued with filtering the event lists of time intervals with high background activity. The outcome was a Good time interval (GTI) file containing the definition of GTI for a given observation. After the high energy light curve

[2] https://www.cosmos.esa.int/web/xmm-newton/current-calibration-files
[3] https://www.cosmos.esa.int/web/xmm-newton/sas-threads

($E > 10$ keV for MOS detectors and $10 < E < 12$ keV for PN detector) was produced by `evselect` task, we determined the low and steady background intervals with threshold measured in counts per second. The corresponding GTI file was created by task `gtigen` where the good reference value of the threshold was set to 0.35 counts per second for MOS and 0.4 counts per second for PN instrument as recommended in Smith et al. (2016). For the new observation of GP3 taken in 2019, these values were slightly modified as described in section 3.2. With GTI file the EPIC event lists were easily filtered once again using `evselect` task.

Filtered EPIC event lists were used to produce scientific products. We focused on production of images and extraction of MOS and PN spectra from point-like source. Once the image was produced using `evselect` task, we extracted the source spectrum from the coordinates where our Green Peas are located (see Table 3.2). For GP1 the radius of extraction region was reduced to 24 arcsec due to proximity of another source to avoid contamination of the output spectrum (Svoboda et al., 2019). The source spectra of GP2 and GP3 were extracted from extraction regions with radius of 30 arcsec. The X-ray background spectra were measured in nearby source-free circular regions on the same chip of each detector. We used the same background aperture radius as for the source aperture, 30 arcsec. The details of extraction regions are in Appendix A. For PN detector, the background spectrum extraction region is laying at similar distance from the readout node as the source region, as the instrumental noise rises with the distance (Smith et al., 2016). Finally we created the redistribution matrix using `rmfgen` task and the ancillary file with `arfgen` task. To simplify the process of data reduction we wrote a bash script which can be found in Appendix B.

For futher spectral analysis, we combined the MOS and PN spectra to single EPIC combined spectrum using task `epicspeccombine`. The output is the combined source spectrum and background spectrum, together with redistribution matrix. The `epicspeccombine` command does not produce an ancillary file, as the informations about effective areas for each detector are saved in newly produced response file. The reduced data were further analysed using XSPEC[4] spectral fitting software version 12.10 (K. A. Arnaud, 1996).

2.3 X-ray spectral analysis

2.3.1 C-statistic

We perform the fitting of the model to our data with assumption of Poisson distribution, therefore using C-statistic. The Gaussian distribution ergo the χ^2 statistic is not recommended for our data, as there are limited numbers of counts per bin

[4] https://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/manual/manual.html

in our spectra. There are two operations performed by XSPEC requiring statistics: a parameter estimation for a given model that provides the best fit to the data and estimation of the uncertainties of the parameter. To find the parameter providing the best fit, the maximum likelihood is used. For Poisson distributed data the likelihood L, given by Cash (1979) is

$$L = \prod_{i=1}^{N} (tm_{\mathrm{i}})^{S_{\mathrm{i}}} \mathrm{e}^{-tm_i} / S_{\mathrm{i}}!. \tag{2.1}$$

The maximum likelihood-based statistics (C-statistic) C is

$$C = 2 \sum_{i=1}^{N} (tm_{\mathrm{i}}) - S_{\mathrm{i}} \ln(tm_{\mathrm{i}}) + \ln(S_{\mathrm{i}}!), \tag{2.2}$$

where S_{i} are the observed counts, t is the exposure time and m_{i} are the predicted count rates based on the current model and instrumental response. The formula 2.2 can be overwritten using Stirling's formula for the final term as

$$C = 2 \sum_{i=1}^{N} (tm_{\mathrm{i}}) - S_{\mathrm{i}} + S_{\mathrm{i}} \left(\ln(S_{\mathrm{i}}) - \ln(tm_{\mathrm{i}})\right). \tag{2.3}$$

The uncertainties of the fitted parameter further in Chapter 3 are determined using command `error` estimating 90% confidence ranges for the non-fixed parameter with delta fit statistic of 2.706.

2.3.2 Spectral model

An X-ray spectrum of a star-forming galaxy dominated by HMXBs is expected to have a shape as a power law with possible additional thermal components due to a hot gas (Mineo et al., 2012). As the observed Green Pea galaxies are unresolved and relatively weak sources, we use a simple power law model for the spectral fitting of all our spectra. The power law model named `zpowerlw` in XSPEC,

$$M_{\mathrm{pl}}(E) = K[E(1+z)]^{-\alpha} \tag{2.4}$$

computes the redshifted spectrum, where redshift z is a fixed parameter, α is a dimensionless photon index of the power law and K is a normalization factor measured in photons per keV/cm^2/s at 1 keV.

As the observed targets are extragalactic objects, we correct the intrinsic power law by an X-ray absorption by the interstellar medium (ISM) of our Galaxy. To do so, we use the Tübingen-Boulder ISM absorption model (Wilms et al., 2000) $M_{\mathrm{TB}}(E)$, shown in expression 2.5, which calculates the cross section for X-ray absorption by the gas, gains and molecules in the ISM of our Galaxy.

$$M_{\mathrm{TB}}(E) = \exp\left[-\sigma_{\mathrm{ISM}}(E) N_{\mathrm{H}}\right] \tag{2.5}$$

In our case, we take into account only hydrogen abundances, using command `nh` in XSPEC to calculate average value of the equivalent hydrogen column, corresponding to sky coordinates where the source is located. Parameter N_{H} refers to the total hydrogen column density, therefore to hydrogen in molecular, neutral or ionized form. The parameter σ_{ISM} is energy dependent photoionization cross section of the ISM and can be obtained as sum of the cross sections for X-ray absorption due to the gas-phase ISM, the grain-phase ISM, and the molecules in the ISM.

Our power-law model multiplied by Tübingen-Boulder ISM model is called by command `tbabs*zpowerlw` and calculated by the following formula

$$M(E) = K[E(1+z)]^{-\alpha} \exp\left[-\sigma_{\mathrm{ISM}}(E) N_{\mathrm{H}}\right]. \tag{2.6}$$

3 | Results

3.1 Reanalysis of archival X-ray observations of Green Peas

All of three selected Green Peas were observed in spring 2013 by all three EPIC cameras operating in Full Frame mode with the thin filter during all of the observations. The *XMM-Newton* observational details are reported in Table 3.1, where the values of net exposures are after the subtraction of intervals affected by high background flares.

Table 3.1: Details of archival *XMM-Newton* observations of Green Peas

Source	Observation ID	Observation Date	Net exposure time [ks]
SDSSJ074936.7+333716 (GP1)	0690470101	2013 Mar 25	35.9
SDSSJ082247.6+224144 (GP2)	0690470201	2013 Apr 6	31.5
SDSSJ133928.3+151642 (GP3)	0690470401	2013 Jan 19	27.1

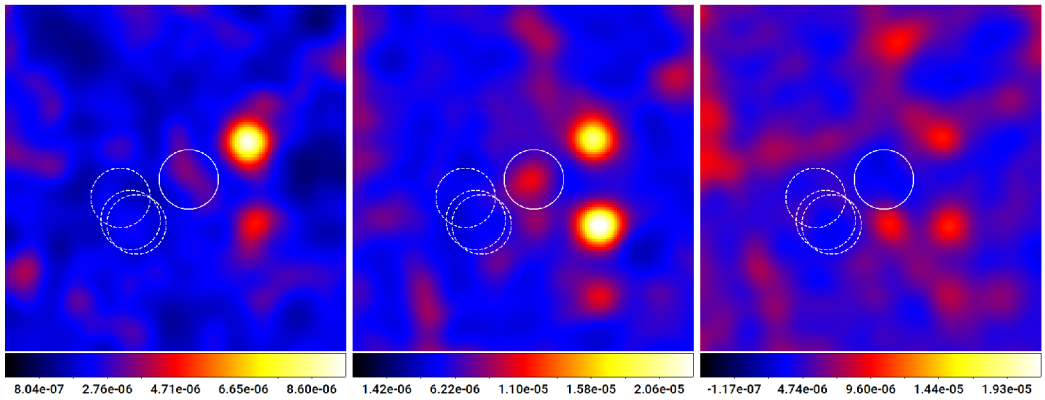
The basic properties of observed Green Peas are in Table 3.2, where the values of SFR for each galaxy are based on Hα measurements from Cardamone et al. (2009), masses are taken from work by Brinchmann et al. (2004) and metallicities were provided by Svoboda et al. (2019).

The X-ray images for all three sources are shown in Figures 3.1, 3.2 and 3.3 respectively, in energy ranges 0.2 – 0.5 keV, 0.5 – 2.0 keV and 2.0 – 8.0 keV, respectively. These science exposure images are combined from all three detectors, EPIC MOS and PN, and are vignetting-corrected and background-subtracted. To generate them, we followed SAS threads "How to Generate Vignetting-Corrected Background-Subtracted Images"[1]. The detection of GP1 and GP2 can be clearly seen in the images, however GP3 remained undetected.

[1] https://www.cosmos.esa.int/web/xmm-newton/sas-thread-images

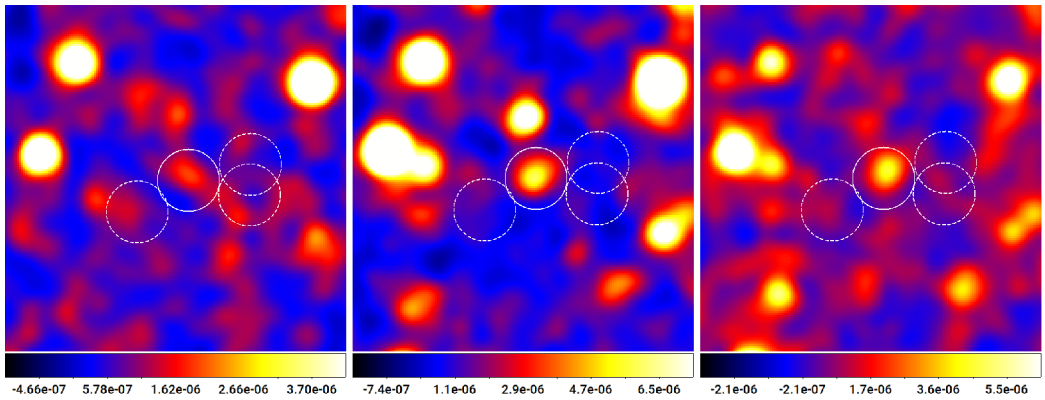
Table 3.2: Properties of Green Peas

Source	R.A. (2000)	Dec. (2000)	Redshift	SFR $[M_\odot\,\mathrm{yr}^{-1}]$	$\log M_*$ $[M_\odot]$	log[O/H] + 12
GP1	117.403215	33.621219	0.2733	58.8	9.8	8.3
GP2	125.698590	22.695578	0.2162	37.4	9.6	8.1
GP3	204.867933	15.278369	0.1920	18.8	9.3	8.1



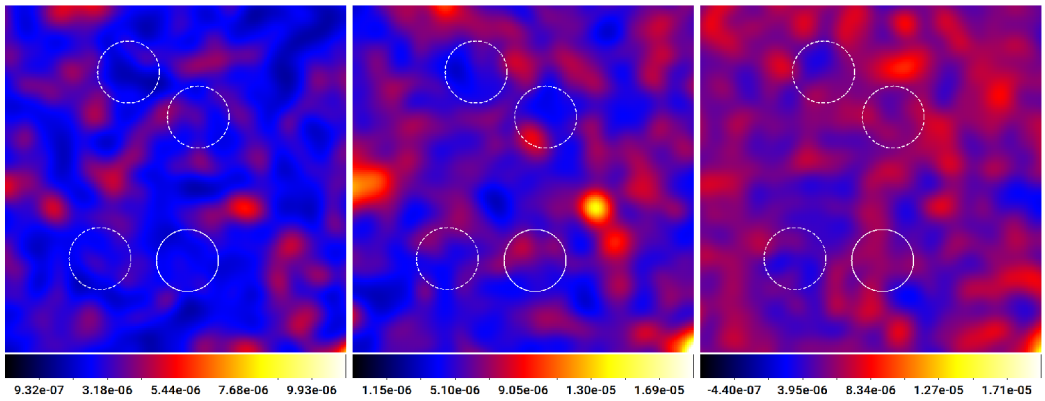
(a) 0.2 to 0.5 keV (b) 0.5 to 2.0 keV (c) 2.0 to 8.0 keV

Figure 3.1: X-ray images of GP1 in three different energy bands. The EPIC combined images from all of the three detectors are vignetting-corrected and background-subtracted. In SAOIMAGE DS9 visualization software we used ZSCALE function to display the images. The solid circle denotes the extraction region of the X-ray source spectrum for GP1 with the radius 24 arcsec. The dashed circles are extraction regions for the background spectra for two MOS and one PN detector also with the radius of 24 arcsec. The intensity is denoted by the color, the width of each sub-figure is 4.53″. The coordinates of extraction regions are in Appendix A.



(a) 0.2 to 0.5 keV (b) 0.5 to 2.0 keV (c) 2.0 to 8.0 keV

Figure 3.2: X-ray images of GP2 in three different energy bands. The extraction region of the X-ray source spectrum denoted by the solid circle and background spectra for three detectors denoted by dashed circles have radius 30 arcsec. The width of each sub-figure is 5.45″.



(a) 0.2 to 0.5 keV (b) 0.5 to 2.0 keV (c) 2.0 to 8.0 keV

Figure 3.3: X-ray images of undetected GP3 in three different energy bands with denoted 30 arcsec extraction regions. The width of each sub-figure is 5.45″.

3.1.1 Spectral analysis

For the spectral fitting, we used C-statisctic; our model is described in section 2.3.2. The spectral fit was performed in a wide energy range 0.3 – 10 keV. Parameter values are in the Table 3.3. The column density of hydrogen N_{H} was obtained using `nh` command in XSPEC calculating the weighted value of N_{H} for determined coordinates. This parameter was fixed by `freeze` command. The value of the photon index of the power law α was taken from Svoboda et al. (2019) for all of the three galaxies and frozen as well. The redshift z was also taken from the work by Svoboda et al. (2019) and by default fixed, so we fitted only the last parameter, which is the normalization factor K. In order to load the spectra processed by SAS to XSPEC, we used `specgroup` tool[2] to specify the background spectrum and response matrix in the header of the source spectrum file. In `specgroup` we defined only the parameter `withoversampling=yes` with `oversample=3` in order not to oversample the intrinsic energy resolution by a factor larger than 3.

Table 3.3: Modelled spectrum of Green Peas

Source	Column density of H $[10^{20}\ \mathrm{cm}^{-2}]$	Photon Index	Normalization factor $[10^{-6}$ phot./keV/cm^2/s$]$
GP1	4.73	3.2	$2.54^{+0.74}_{-0.71}$
GP2	4.27	2.0	$2.77^{+0.58}_{-0.55}$

The values of obtained normalization factor are in case of GP1 and GP2 consistent within the statistical uncertainties with results provided by Svoboda et al. (2019). The corresponding unabsorbed X-ray flux and luminosity are in Table 3.4. The flux was obtained using command `flux` with specifying the energy range corrected by redshift by formula

$$E_{\mathrm{observed}} = \frac{E_{\mathrm{emitted}}}{1+z}. \tag{3.1}$$

The X-ray luminosity was obtained using command `lumin` specifying the observed energy range and redshift for correction. The statistical uncertainties of flux and luminosity correspond to variation of the normalization factor. All measurements of X-ray flux are expressed in unit of ergs per cm^2 per second and X-ray luminosity is expressed in ergs per second. One erg is equal to 10^{-7} Joules, therefore 1 erg/s = 10^{-7} W, solar luminosity in ergs per second is $1\mathrm{L}_{\odot} = 3.826 \times 10^{33}$ ergs/s.

[2] https://xmm-tools.cosmos.esa.int/external/sas/current/doc/specgroup

Table 3.4: Measured X-ray flux and luminosity of Green Peas

Source	Net count rate $[10^{-3}$ cts/s] 0.3 – 10 keV	X-ray Flux $[10^{-15}$ ergs cm^{-2} s$^{-1}]$ 0.5 – 8.0 keV	X-ray Luminosity $[10^{42}$erg/s] 0.5 – 8.0 keV
GP1	3.55 ± 0.85	$4.63^{+1.36}_{-1.29}$	$1.09^{+0.32}_{-0.31}$
GP2	5.48 ± 0.81	$8.31^{+1.73}_{-1.65}$	$1.15^{+0.24}_{-0.23}$

The modelled spectra are shown in Figure 3.4. In order to plot the observed spectra we binned the data of both GP1 and GP2 with minimum signal to noise equal to 1.

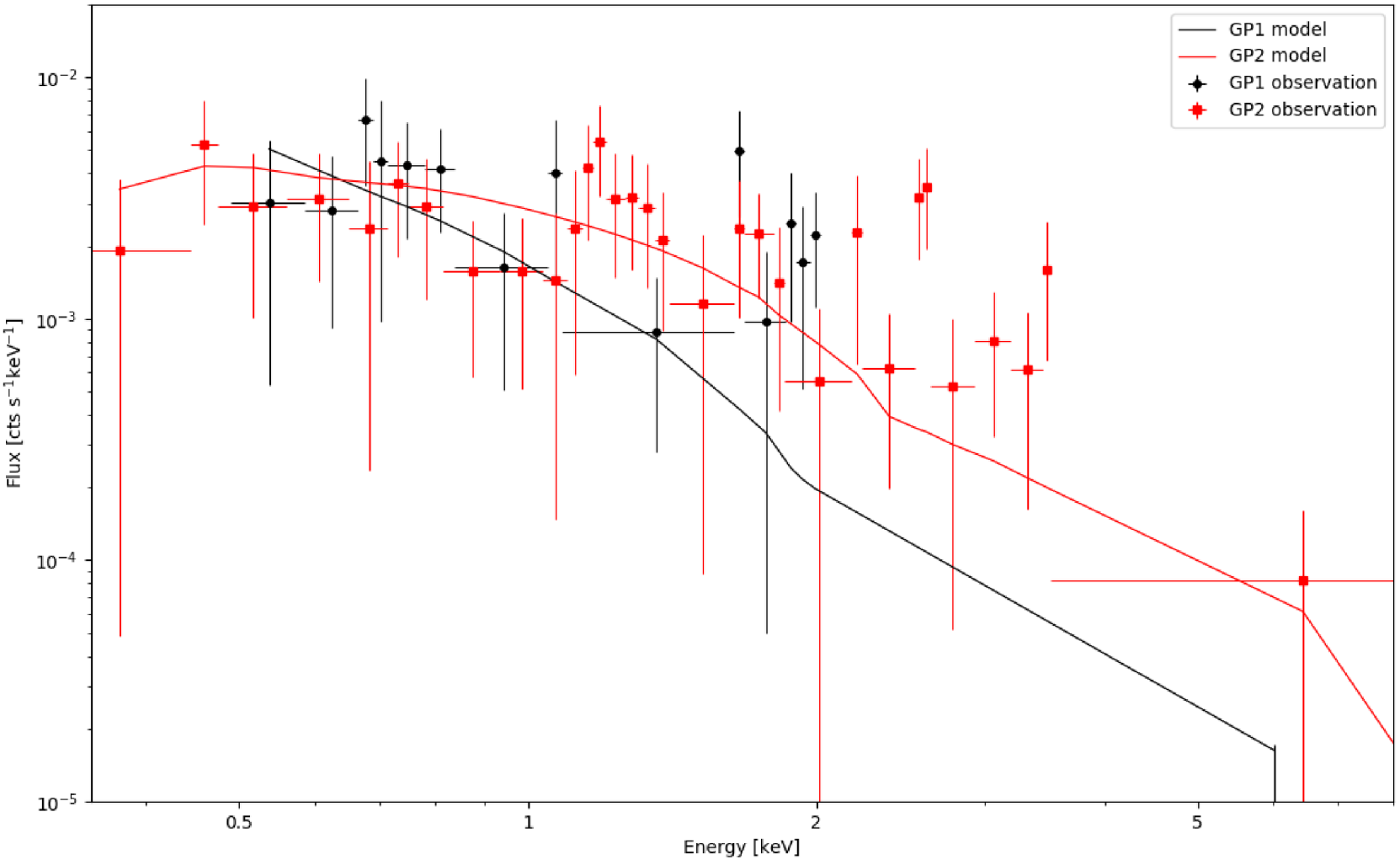


Figure 3.4: Modelled spectra of GP1 and GP2

Using the same methodology of the data preparation for GP3 as for GP1 and GP2 data sets, we obtained following results. The value of the column density obtained by command `nh` for GP3 is $N_{\mathrm{H}} = 1.97 \times 10^{20}\mathrm{cm}^{-2}$ and the photon index taken from Svoboda et al. (2019) is $\alpha = 2.0$. The obtained normalization factor and the corresponding X-ray flux and X-ray luminosity are

$$\begin{aligned} K &= \left(4.02^{+0.93}_{-0.88}\right) \times 10^{-6}\ \mathrm{photons/keV/cm^2/s}, \\ F_{\mathrm{X}} &= \left(12.56^{+2.92}_{-2.75}\right) \times 10^{-15}\ \mathrm{erg/cm^2/s}, \\ L_{\mathrm{X}} &= \left(1.33^{+0.31}_{-0.29}\right) \times 10^{42}\ \mathrm{erg/s}. \end{aligned}$$

As the obtained X-ray luminosity for undetected GP3 galaxy is higher than the output for detected GP1 and GP2 galaxies, it seems not to be reasonable and rises suspicion. To obtain this result we used different method than authors of the work by Svoboda et al. (2019), who obtained only the upper limit of X-ray luminosity for GP3 fitting raw unbinned data. To be able to compare our output with their result of 1.3×10^{41} erg/s we modified the grouping process in `specgroup`: We did not define the `withoversampling` option, so the output product is raw unbinned spectrum with allowed oversampling. From the spectral fitting of raw data we also could estimate only the upper limit of the normalization factor, while the lower limit was pegged to 0. The obtained upper limit of the normalization factor and the corresponding maximum measured X-ray flux and luminosity are

$$\begin{aligned} K &< 2.66 \times 10^{-7}\ \mathrm{photons/keV/cm^2/s}, \\ F_{\mathrm{X}} &< 8.33 \times 10^{-16}\ \mathrm{erg/cm^2/s}, \\ L_{\mathrm{X}} &< 8.83 \times 10^{40}\ \mathrm{erg/s}. \end{aligned}$$

This value is lower than the result provided by Svoboda et al. (2019) and the difference might be explained by usage of different background extraction region. How can different background extraction region affect the results we discuss in section 4.2. Different grouping and binning procedures are described in section 3.2 and the discussion of whether binning methods might lead to artificial detections is in the section 4.1.

3.2 Analysis of new observation of GP3

The net exposure time of previous observation of GP3 turned out to be insufficient for the detection of the source due to the subtraction of intervals affected by high background flares. The upper limit of the 0.5 – 8.0 keV luminosity of the galaxy was set to 1.3×10^{41} erg/s by Svoboda et al. (2019). The new observation taken in January of 2019 should supply sufficient data to put tighter constrains on the upper limit of GP3s X-ray luminosity in case it is not detected. Table 3.5 shows observational details of the new observation.

Table 3.5: Details of new *XMM-Newton* observation

Source	Observation ID	Observation Date	Net exposure time[a] [ks]
GP3	0825100101	2019 Jan 12	49.0

[a]For combined source spectrum from all three detectors.

The data provided by more than 87 kiloseconds of observation were reduced the same way as we describe in section 2.2, except for the GTI specification. Due to elevated count rate of the background in the middle of the observation, we determined the low and steady background intervals manually for MOS cameras by defining regions of good exposure time. Figures 3.5 and 3.6 show the count rate curves for MOS detectors with noticeable peak in the middle of the observation, that would fall under the threshold of 0.35 counts per second if filtered automatically. The Figure 3.7 shows count rate curve for PN detector, more sensitive in high energies than MOS instruments. Here the peak is much higher than the threshold of 0.4 counts per second, however, we modified the threshold to 0.5 counts per second so the low and steady background regions would fall under entirely. After the correction of the light curves from intervals affected by high background flares we obtained over 50 kiloseconds of useful exposure from MOS detectors and nearly 40 kiloseconds of exposure from PN detector.

The source spectra were extracted from circular extraction regions with radius of 30 arcsec. For each detector we extracted two background spectra using circular aperture with radius of 30 arcsec from source-free regions on the same chip of the detector, as the source is located. However, the galaxy was not recognized visually at any of the three produced images corresponding to the three EPIC detectors. Detection of GP3 was also refuted by `edetect_chain` task, a SAS tool performing source detection. In neither of all three EPIC produced scientific images the galaxy was detected. Redistribution and ancillary files were created using `rmfgen` and `arfgen` tasks. For further spectral analysis we combined the source spectra from

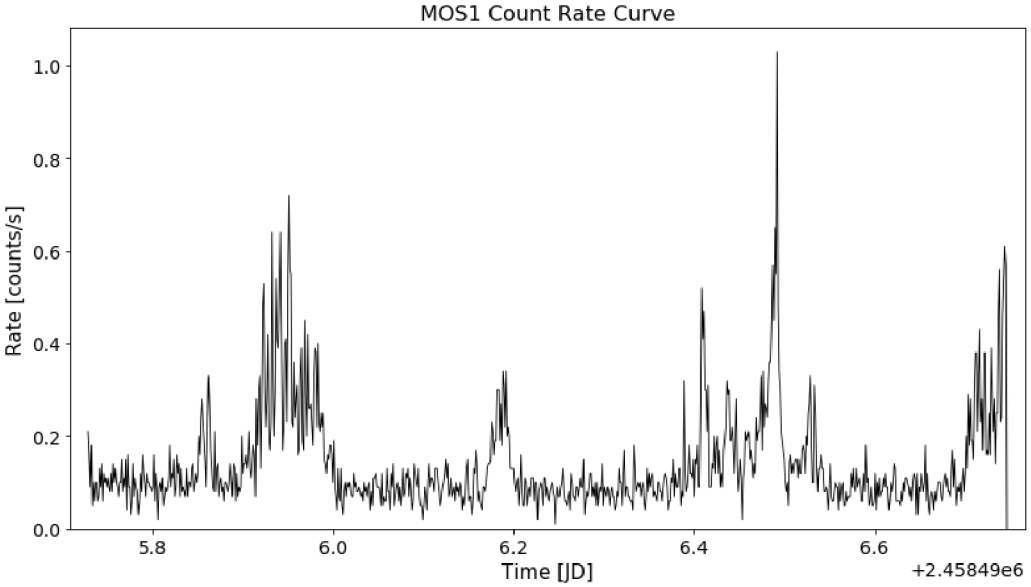


Figure 3.5: MOS1 count rate curve for GP3 new observation.

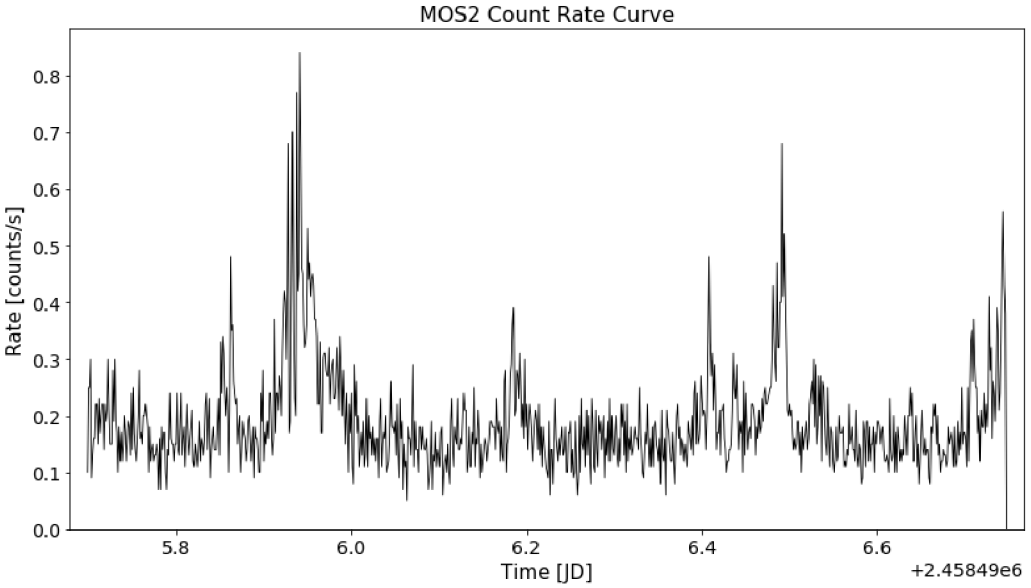


Figure 3.6: MOS2 count rate curve

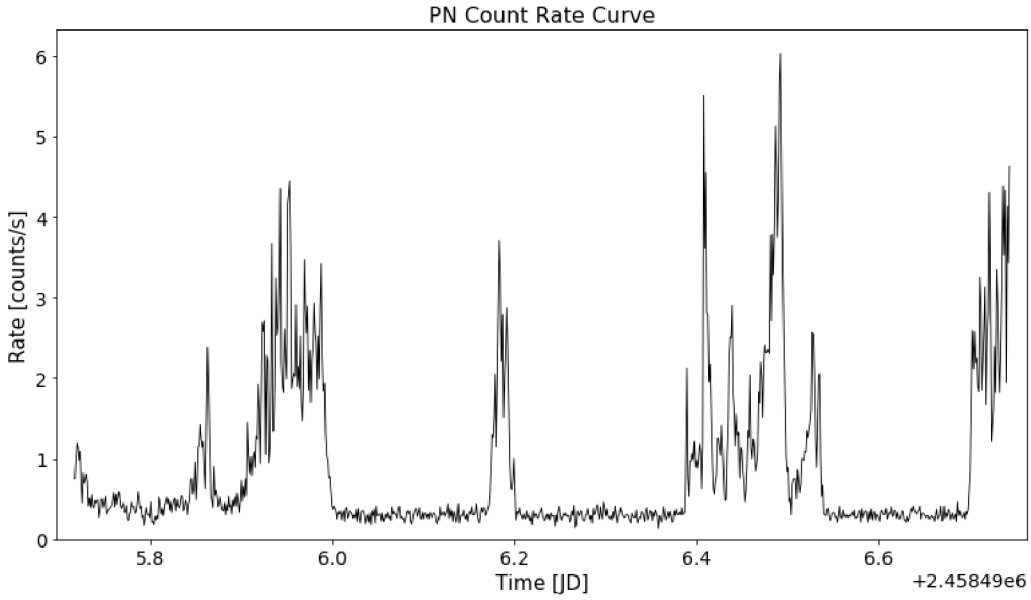
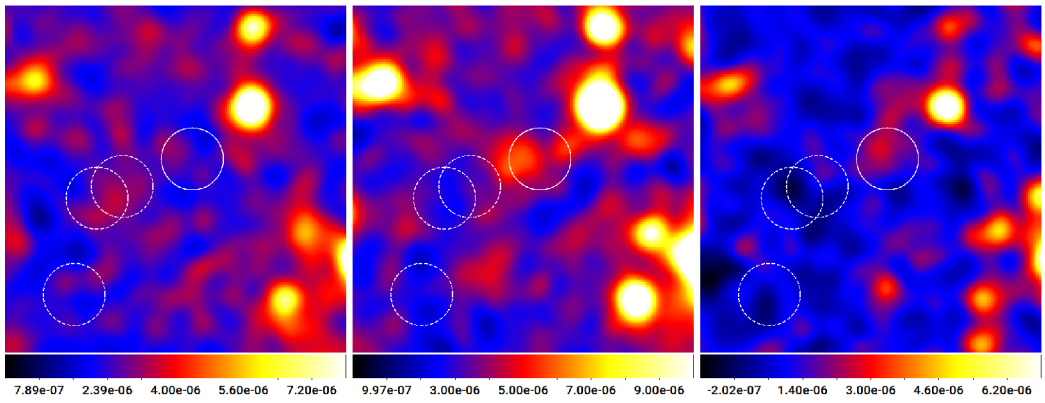


Figure 3.7: PN count rate curve

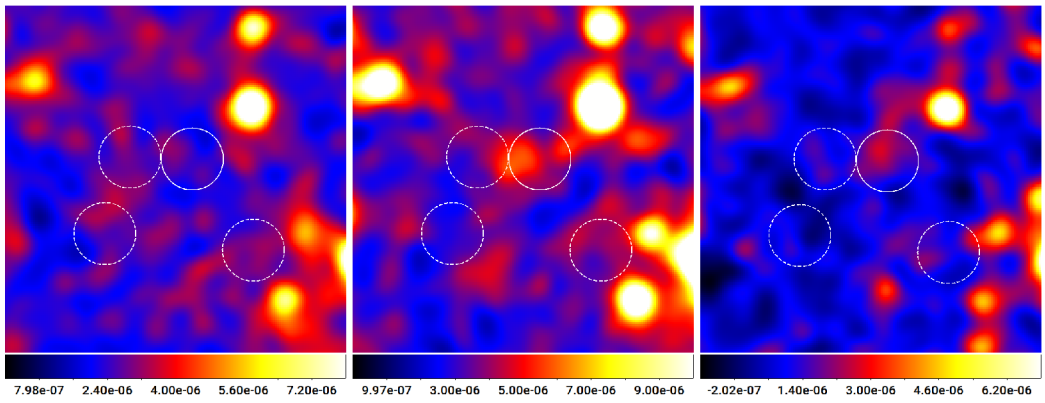
all detectors into single one spectrum with corresponding response files using `epicspeccombine` tool. We applied this method twice using both background spectra obtaining two data sets with different X-ray backgrounds. When applying the combining tool to response matrix and effective area file, the output is single response matrix, no ancillary file is generated for the merged spectrum. Similar approach was applied on the scientific images using SAS tool `eimagecombine` producing combined images for energy ranges 0.2 – 0.5 keV, 0.5 – 2.0 keV and 2.0 – 8.0 keV, respectively. These images are shown in Figures 3.8 for background number 1 and 3.9 for background number 2. Visually the galaxy could not be recognised on these images, yet there is an indication of a source on the edge of the solid aperture circle. This fluctuation was not detected on any of the single images from the detectors. However, the center of the fluctuation is probably located between 15 and 30 arcsec from the GP3 coordinates, which corresponds to about 60 to 120 kpc at the GP3's redshift. As Green Peas are very compact galaxies with diameter usually not larger than 5 kpc (Cardamone et al., 2009), the fluctuation is more probably a source on a cosmic background and Green Pea 3 remains undetected.

The combined data of all three detectors produced by `epicspeccombine` task can be directly loaded to XSPEC for spectral analysis, where we focused on specifying the upper limit of the galaxy X-ray luminosity, as it was not detected on



(a) 0.2 to 0.5 keV (b) 0.5 to 2.0 keV (c) 2.0 to 8.0 keV

Figure 3.8: X-ray images of GP3 new observation in three different energy bands. Solid circle denotes extraction region of the source spectrum and dashed circles denote extraction regions of background for all three detectors. The radius of all of the extraction regions is 30 arcsec and the width of each sub-figure is 5.45″.



(a) 0.2 to 0.5 keV (b) 0.5 to 2.0 keV (c) 2.0 to 8.0 keV

Figure 3.9: The same images as in Figure 3.8 but for different background extraction regions. These are denoted by dashed circles for each detector.

the images. We decided to fit the raw unbinned spectrum as well as binned one and compare the results. There are several tools available in SAS allowing grouping of channels in the spectrum along with preparing the data for analysis in XSPEC. Those tasks used in this work for binning our spectra are listed below.

grppha The `grppha`[3] is an interactive command driven task allowing the user to define or redefine the header of source spectrum file. Subsequent `grppha` commands allow grouping of channels in a spectrum. In our work we used this tool primarily to define the background spectrum and response matrix in header of the source spectrum in order to be able to load raw spectrum to XSPEC. With objective of comparing different binning tools we also used `grppha` for binning of the source spectrum to minimum 1 count per each bin using command `min counts 1` in the `grppha` environment.

specgroup The SAS tool `specgroup` allows a flexible grouping of channels in a spectrum extending the functionality of `grppha` tool by providing more options of binning the spectrum. In this work we use `specgroup` to bin the source spectrum to minimum 1 count per each bin not allowing oversampling by factor larger than 3. If the `oversample` parameter is not set, the output spectrum corresponds to output file created by `grppha` with the same minimum counts per bin.

ftgrouppha This tool provides several ways to group the channels in a spectrum. In this work we used the following two group-types.

`bmin` Bins the source spectrum so that there is minimum number of counts in each bin in the background spectrum. The `groupscale` parameter defines the minimum number of counts per bin; in this work we use the `groupscale` parameter equal to 1 count per bin.

`opt` The binning is based on optimal binning scheme by Kaastra, J. S. et al. (2016).

The spectral fitting was performed in energy range 0.3 – 10 keV. The only fitted parameter was the normalization factor, the rest of the parameters of the model were fixed. The redshift of GP3 is in Table 3.2, the column density of hydrogen was obtained by command `nh` in XSPEC, its value is $N_{\mathrm{H}} = 1.97 \times 10^{20} \mathrm{cm}^{-2}$, and the value of the photon index of the power law taken from Svoboda et al. (2019) is $\alpha = 2.0$. The result for fitted parameter with corresponding unabsorbed X-ray flux and X-ray luminosity measurements are in Table 3.6 for the background 1 and in Table 3.7 for the background 2. The values of flux and luminosity are

[3] https://heasarc.gsfc.nasa.gov/ftools/caldb/help/grppha.txt

calculated for the 0.5 – 8.0 keV band in the rest frame of the galaxy, derived from the observed flux in 0.42 – 6.71 keV band. The last columns in Tables 3.6 and 3.7 show C-statistics fit goodness and number of degrees of freedom.

From the Tables 3.6 and 3.7 it can be seen, that there is a significant difference between raw unbinned data and the binned ones. From modelling the unbinned spectrum grouped by `grppha` tool we could constrain only the upper limit of the normalization factor and therefore only the upper limit of observed X-ray luminosity. The results from modelling the binned spectra, where the oversampling acts also like a binning, predict X-ray luminosity to be as high as for detected source and are comparable with the X-ray luminosities obtained for GP1 and GP2 galaxies. As the GP3 is not clearly detected in the images and since it is the nearest object from the three studied GPs, we rely more on the analysis of the raw data.

Table 3.6: Results for modelling the spectrum of GP3 using background no. 1

Grouping tool	Net count rate [10^{-3} cts/s] 0.3 – 10 keV	Normalization factor [10^{-6} phot./keV/cm^2/s]	X-ray Flux [10^{-15} ergs cm^{-2} s^{-1}] 0.5 – 8.0 keV	X-ray Luminosity [10^{41} erg/s] 0.5 – 8.0 keV	C-statistic & Degrees of freedom
`GRPPHA` no binning	3.11 ± 0.85	< 0.14	< 0.43	< 0.46	1146.29/1938
`SPECGROUP` `oversample=3` no binning	3.21 ± 0.84	$1.19^{+0.54}_{-0.53}$	$3.71^{+1.68}_{-1.66}$	$3.93^{+1.78}_{-1.76}$	258/190
`SPECGROUP` `oversample=3` min cts/bin = 1	3.13 ± 0.84	1.22 ± 0.53	$3.83^{+1.66}_{-1.65}$	$4.06^{+1.76}_{-1.75}$	176/155
`GRPPHA` `group min=1`	3.15 ± 0.85	$1.44^{+0.51}_{-0.50}$	$4.51^{+1.60}_{-1.55}$	$4.78^{+1.70}_{-1.65}$	557/565
`FTGROUPPHA` `grouptype=bmin`	3.11 ± 0.85	$1.01^{+0.55}_{-0.54}$	3.14 ± 1.70	$3.33^{+1.81}_{-1.80}$	523/486
`FTGROUPPHA` `grouptype=opt`	3.19 ± 0.81	$1.36^{+0.54}_{-0.53}$	$4.25^{+1.68}_{-1.67}$	$4.50^{+1.78}_{-1.76}$	60/56

Table 3.7: Results for modelling the spectrum of GP3 using background no. 2

Grouping tool	Net count rate [10^{-3} cts/s] 0.3 – 10 keV	Normalization factor [10^{-6} phot./keV/cm^2/s]	X-ray Flux [10^{-15} ergs cm^{-2} s^{-1}] 0.5 – 8.0 keV	X-ray Luminosity [10^{41} erg/s] 0.5 – 8.0 keV	C-statistic & Degrees of freedom
GRPPHA no binning	2.16 ± 0.86	< 0.11	< 0.34	< 0.36	1186/1938
SPECGROUP oversample=3 no binning	2.12 ± 0.85	0.92 ± 0.53	2.87 ± 1.67	$3.04^{+1.77}_{-1.76}$	171/190
SPECGROUP oversample=3 min cts/bin=1	2.16 ± 0.85	0.96 ± 0.53	3.02 ± 1.65	3.20 ± 1.75	136/168
GRPPHA group min=1	2.24 ± 0.86	1.15 ± 0.55	3.62 ± 1.72	3.94 ± 1.82	751/799
FTGROUPPHA grouptype=bmin	2.16 ± 0.86	$0.61^{+0.56}_{-0.57}$	$1.90^{+1.75}_{-1.79}$	2.01 ± 1.90	690/718
FTGROUPPHA grouptype=opt	2.10 ± 0.83	$0.96^{+0.55}_{-0.54}$	$3.00^{+1.71}_{-1.70}$	$3.18^{+1.81}_{-1.80}$	62/69

4 | Discussion

Our results from spectral analysis of data obtained for GP1 and GP2 are consistent with the results provided by Svoboda et al. (2019). For GP1 the measured unabsorbed X-ray Luminosity in this work $\left(1.09^{+0.32}_{-0.31}\right) \times 10^{42}$ erg/s is consistent within the statistical uncertainty with the value of $(1.2 \pm 0.4) \times 10^{42}$ erg/s obtained by Svoboda et al. (2019). For GP2 our result $\left(1.15^{+0.24}_{-0.23}\right) \times 10^{42}$ erg/s is also consistent to the value $\left(1.2^{+0.2}_{-0.3}\right) \times 10^{42}$ erg/s provided by Svoboda et al. (2019). Moreover, Svoboda et al. (2019) modelled raw data grouped by `grppha` tool with no binning, in this work a `specgroup` tool with oversample parameter set to 3 was applied. The consistency of the results obtained by using different tools thus increases the credibility of obtained values for X-ray luminosities, although it must be mentioned that GP1 and GP2 both were detected easily. The possible explanations of the X-ray excess are discussed in section 4.3.1.

The analysis of archival observation of GP3 executed in this work provides slightly different results than those provided by Svoboda et al. (2019). In their work the authors obtained only the upper limit for X-ray luminosity of 1.3×10^{41} erg/s for GP3 modelling raw unbinned data grouped by `grppha` tool. Using the same approach we obtained upper limit for X-ray luminosity of 8.83×10^{40} ergs/s. This variance may have resulted from usage of different background spectra; the effect of different background we discuss in section 4.2. However, the results from analysis of binned data with oversample parameter equal to 3 differ significantly. For GP3 we obtained X-ray luminosity corresponding to best fit for modelled spectrum of $\left(1.33^{+0.31}_{-0.29}\right) \times 10^{42}$ ergs/s which is higher than X-ray luminosity obtained for easily detected galaxies GP1 and GP2. This result arouses suspicion about suitability of used method in process of obtaining the upper limit of luminosity for undetected source. This topic is discussed further in section 4.1.

The results of data analysis from the new observation of GP3 shows difference between the outputs obtained by fitting the raw unbinned data and the binned one. Using two different background extraction regions we obtained similar upper limits of 4.6×10^{40} erg/s and 3.6×10^{40} erg/s from raw data spectral fitting. All the binning tools we used provide about an order higher result for the X-ray luminosity of GP3. Discussion about different methods of data preparation is in the section 4.1.

4.1 Different binning approaches

Our spectral analysis described in Chapter 3 shows that results provided by different binning approaches may vary significantly. Disagreement between diverse methods appears if we try to set the upper limit of X-ray luminosity for the third, undetected Green Pea 3 galaxy. The main difference is between the results from spectral fitting of raw and binned data.

The results originating from different methodology are inconsistent. There is a significant difference between the upper limits obtained from raw data and the results of the binning methods that behave as if there was a source detection. The binning tools provide results more than an order higher as compared to the unbinned data spectral analysis. Also the comparison of the obtained X-ray luminosity for undetected GP3 with X-ray luminosities obtained for detected GP1 and GP2 galaxies, using `specgroup` tool with `oversample=3` arouses suspicion. Using this tool, for GP3 we obtained value of X-ray luminosity higher than for two detected galaxies, but since the GP3 was not detected in the scientific images, this result seems not to be reasonable.

This shows, that the upper limit for undetected source may be very sensitive to the methodology of binning. As all of the binning tools seem to be fitting a spectrum of a detected source and provide much higher outputs, this methodology may also lead to an artificial detection. There is only a little difference between the individual binning tools, as the result they provide are corresponding with themselves. However, there is opinion, held by our consultant Peter Boorman, that C-statistic works flawless if there are no zero bins in the background spectrum. This is achieved by applying the `ftgrouppha groupstyle=bmin` binning tool to the data. From Tables 3.6 and 3.7 can be seen, that this specific binning tool provides results lower than other binning tools for both extracted backgrounds. However in case of background number 1 the luminosity uncertainty does not include the value of upper limit obtained from raw data. In case of background number 2 the upper limit is included in the luminosity uncertainty of this particular binning tool, but the results still predict a source detection. The reason, why different methods of data preparation provide such variant results is an objective of our future work.

4.2 Dependence on the background extraction region

For new Green Pea 3 observation, we decided to provide results for several different backgrounds to see, how much is the output of spectral fitting affected by them. In both cases of the new GP3 observation we selected background extraction regions with radius of 30″ in a source-free circular area not far from the source extraction region. For MOS detectors, we managed to select two distinct regions without an

overlap. In principle, the background extraction region for PN detector should be located in similar distance from the readout node as the source extraction region (Smith et al., 2016), however as the chip has shape of a narrow strip with the shorter edge being adjacent with the readout node, our extraction regions partially overlap. From Tables 3.6 and 3.7 it can be seen, that the values are corresponding within the statistical uncertainties and for the upper limits the difference is small. We can conclude, that the results of spectral analysis are in the case of different background usage with the same radius congruent if their intensity is similar.

In addition, to be able to directly compare our results for X-ray luminosity of GP3 galaxy from the archival observation with the result provided in the work by Svoboda et al. (2019), we used the same background extraction region together with the same method of preparing the data for spectra fitting. Their background extraction regions have bigger radius, specifically 120 arcsec for MOS detectors and 60 arcsec for PN detector (for more details see Table 3 in Appendix A). Using the same background coordinates and radii and fitting raw unbinned data grouped by `grppha` tool we obtained the following values of the upper limits for the normalization factor and corresponding X-ray flux and luminosity

$$\begin{aligned} K &< 8.80 \times 10^{-7}\ \text{photons/keV/cm}^2\text{/s}, \\ F_{\mathrm{X}} &< 2.75 \times 10^{-15}\ \text{erg/cm}^2\text{/s}, \\ L_{\mathrm{X}} &< 2.92 \times 10^{41}\ \text{erg/s}. \end{aligned}$$

Even if the value of X-ray luminosity is higher than 1.3×10^{41} erg/s as in the work of Svoboda et al. (2019), there can be also seen the increase of the luminosity comparing to our previous result 8.83×10^{40} erg/s using low-intensity but smaller background region. Using `specgroup` tool with `oversmaple=3` we obtained very similar results

$$\begin{aligned} K &< 1.16 \times 10^{-6}\ \text{photons/keV/cm}^2\text{/s}, \\ F_{\mathrm{X}} &< 3.63 \times 10^{-15}\ \text{erg/cm}^2\text{/s}, \\ L_{\mathrm{X}} &< 3.85 \times 10^{41}\ \text{erg/s}. \end{aligned}$$

From these results it can be seen, that larger background regions are much more robust and the differences between the outputs from raw and binned data are significantly reduced. However, deeper analysis of different binning methods and different background extraction regions and their dependency on the upper limits of undetected source is an objective of our future work.

4.3 Difference between studied Green Pea galaxies

Figure 1.3 shows the studied Green Peas in metallicity-dependent L_X-SFR plane by Brorby; Kaaret; Prestwich; Mirabel (2016). The position of Green Peas results from analysis in the work by Svoboda et al. (2019), who discovered the X-ray excess in galaxies GP1 and GP2. Despite similar physical properties of the three studied Green Pea galaxies, taking into account that both GP2 and GP3 are in the Figure 1.2 located almost in the same position, in terms of absolute numbers there is big difference in their X-ray emission. The first two galaxies show an enhanced X-ray emission, however, the third galaxy remained undetected. Our analysis revealed, that the third Green Pea galaxy was undetected not due to the shrank net exposure time caused by subtraction of the intervals affected by the high background flares, but due to its physical nature. Here we provide a several possible explanations of the enhanced X-ray emission in two detected galaxies.

4.3.1 Origin of enhanced X-ray emission in GP1 and GP2

Low mass X-Ray binaries

The X-ray emission of star-forming galaxies is proportional to number of relatively short-lived HMXBs, hence it is proportional to the SFR as follows

$$L_{0.5-8.0\,\mathrm{keV}}[\mathrm{erg\,s^{-1}}] \approx 4 \times 10^{39}\,\mathrm{SFR}\,[M_\odot\,\mathrm{yr^{-1}}]. \tag{4.1}$$

On the other hand, contribution of older LMXB populations is proportional to the stellar mass as

$$L_{\mathrm{LMXB},\,0.5-8.0\,\mathrm{keV}}[\mathrm{erg\,s^{-1}}] \approx 10^{39} \times \frac{M_*}{10^{10} M_\odot}, \tag{4.2}$$

obtained by Gilfanov; Merloni (2014). Therefore if the SFR of a galaxy is higher than 1 $M_\odot\,\mathrm{yr^{-1}}$ (Gilfanov; Grimm, et al., 2004), the contribution of LMXBs to total X-ray emission is not significant. As our sample of Green Peas has SFR $> 10\ M_\odot\,\mathrm{yr^{-1}}$, the contribution of LMXBs is truly insignificant.

Ultra-luminous and Hyper-luminous X-ray sources

The origin of unexplained X-ray emission could be attributed to so-called ultra-luminous X-ray sources (ULXs), a point-like sources and non-nuclear objects found in external galaxies with X-ray luminosities of $10^{39} < L_X < 10^{41}$ erg/s (Bachetti et al., 2014). The majority of ULXs are probably an X-ray binaries hosting stellar-mass compact object with accretion near or above the Eddington

limit (Kaaret et al., 2017). However, a presence of at least 10 such objects would be required to explain an excess of almost 10^{42} erg/s, measured from GP1 and GP2.

Additionally, there are non-nuclear X-ray sources even brighter than ULXs, with luminosities $L_X > 10^{41}$ erg/s named hyper-luminous X-ray sources (Barrows et al., 2019, HLXs). The origin of such luminous X-ray emission from these objects could be explained by accretion onto more massive black holes. Therefore HLXs are candidates for so-called intermediate black holes (IMBH) with masses about $10^2 < M_\odot < 10^5$ and may prove the existence of these objects (Kaaret et al., 2017; Barrows et al., 2019). Such powerful X-ray sources might explain the measured excess in GP1 and GP2, but we can only speculate about individual X-ray sources characteristics in our sample of studied galaxies due to lack of spatial resolution.

Possible AGN contribution

Another explanation of enhanced X-ray emission would be the presence of an AGN. In the L_X-SFR-metallicity plane (see Figure 1.3) based on study by Brorby; Kaaret; Prestwich; Mirabel (2016) the GP1 and GP2 galaxies are located in a region, where an AGNs could be expected. Even thought our Green Peas are in the BTP diagram (see Figure 1.2) located in a pure star-forming region, especially GP2 lies near the line separating starbursts from AGNs. This may raise the questions about a possible AGN contribution, but to prove or contradict this theory a multi-spectral observations in infrared and radio bands would be required.

Summary

In this work we analysed the archival observations of the three Green Peas galaxies observed by the space X-ray observatory *XMM-Newton* and also a new observation of previously undetected third Green Pea galaxy. Our results showed increased X-ray emission from two detected GP galaxies and confirmed the results provided by Svoboda et al. (2019).

From our analysis of the new observation of GP3 galaxy we found out, that the galaxy was not detected not due to the insufficient observational time, but due to its physical nature. We also learned, that the upper limit of its X-ray luminosity is sensitive to the methodology of preparation of the data for the spectral fitting and the binning methods might lead to artificial detection. The choice of the background extraction region, especially its radius, is also a factor on which the final results are dependent. Nevertheless we specified the upper limit of GP3's X-ray luminosity and found out, that it is more than an order lower than the X-ray luminosities of the two detected GP galaxies.

The difference between these three objects could be explained by the presence of so-called hyper-luminous X-ray sources, which are candidates for unconfirmed intermediate black holes or multiple ultra-luminous X-ray sources, probably an X-ray binaries with accretion near or above the Eddington limit in the first two galaxies. Another explanation would be an accretion of matter onto a central black hole in Green Pea 1 and Green Pea 2 galaxies, which does not occur in the third studied Green Pea galaxy.

Bibliography

Arnaud, K. A., 1996. XSPEC: The First Ten Years. In: Jacoby, George H.; Barnes, Jeannette (eds.). *Astronomical Data Analysis Software and Systems V*. Vol. 101, p. 17. Astronomical Society of the Pacific Conference Series.

Bachetti, M. et al., 2014. An ultraluminous X-ray source powered by an accreting neutron star. *Nature*. Vol. 514, pp. 202–204. ISSN 1476-4687. Available from DOI: 10.1038/nature13791.

Baldwin, J. A.; Phillips, M. M.; Terlevich, R., 1981. Classification parameters for the emission-line spectra of extragalactic objects. *Publications of the Astronomical Society of the Pacific*. Vol. 93, pp. 5–19. Available from DOI: 10.1086/130766.

Barrows, R. Scott; Mezcua, Mar; Comerford, Julia M., 2019. A Catalog of Hyperluminous X-Ray Sources and Intermediate-mass Black Hole Candidates out to High Redshifts. *The Astrophysical Journal*. Vol. 882, no. 2, pp. 181. ISSN 1538-4357. Available from DOI: 10.3847/1538-4357/ab338a.

Basu-Zych, A. R.; Gonçalves, Thiago S.; Overzier, Roderik; Law, David R.; Schiminovich, David; Heckman, Tim; Martin, Chris; Wyder, Ted; O'Dowd, Matt, 2009. An Osiris Study of the Gas Kinematics in a Sample of UV-Selected Galaxies: Evidence of "Hot and Bothered" Starbursts in the Local Universe. *The Astrophysical Journal*. Vol. 699, no. 2, pp. L118–L124. Available from DOI: 10.1088/0004-637x/699/2/l118.

Basu-Zych, A. R.; Lehmer, B. D.; Bret, D., et al., 2013. Evidence fro elevated X-ray emission in local Lyman Break Galaxy Analogs. *The Astrophysical Journal*. Vol. 774, no. 2, pp. 152. ISSN 1538-4357. Available from DOI: 10.1088/0004-637x/774/2/152.

Basu-Zych, A. R.; Lehmer, B. D.; Fragos, T.; Hornschemeier, A.; Yukita, M.; Zezas, A.; Ptak, A., 2016. Exploring the overabundance of ULXs in metal- and dust-poor local Lyman Break Aanalogs. *The Astrophysical Journal*. Vol. 818, no. 2, pp. 140. Available from DOI: 10.3847/0004-637x/818/2/140.

Basu-Zych, A. R.; Lehmer, B. D.; Hornschemeier, A. E., et al., 2012. The X-ray Star formation story as told by Lyman Break Galaxies in the 4 Ms CDF-S. *The Astrophysical Journal*. Vol. 762, no. 1, pp. 45. ISSN 1538-4357. Available from DOI: 10.1088/0004-637x/762/1/45.

Brorby, M.; Kaaret, P., 2017. X-rays from Green Pea analogues. *Monthly Notices of the Royal Astronomical Society*. Vol. 470, no. 1, pp. 606–611. ISSN 1365-2966. Available from DOI: 10.1093/mnras/stx1286.

Brorby, M.; Kaaret, P.; Prestwich, A., 2014. X-ray binary formation in low-metallicity blue compact dwarf galaxies. *Monthly Notices of the Royal Astronomical Society*. Vol. 441, no. 3, pp. 2346–2353. ISSN 0035-8711. Available from DOI: 10.1093/mnras/stu736.

Brorby, M.; Kaaret, P.; Prestwich, A.; Mirabel, I. F., 2016. Enhanced X-ray emission from Lyman break analogues and a possible LX–SFR–metallicity plane. *Monthly Notices of the Royal Astronomical Society*. Vol. 457, no. 4, pp. 4081–4088. ISSN 0035-8711. Available from DOI: 10.1093/mnras/stw284.

Cardamone, C. et al., 2009. Galaxy Zoo Green Peas: Discovery of A Class of Compact Extremely Star-Forming Galaxies. *Monthly Notices of the Royal Astronomical Society*. Vol. 399, no. 3, pp. 1191–1205. Available from DOI: 10.1111/j.1365-2966.2009.15383.x.

Cash, W., 1979. Parameter estimation in astronomy through application of the likelihood ratio. *The Astrophysical Journal*. Vol. 228, pp. 939–947. Available from DOI: 10.1086/156922.

Douna, V. M.; Pellizza, L. J.; Mirabel, I. F.; Pedrosa, S. E., 2015. Metallicity dependence of high-mass X-ray binary populations. *Astronomy & Astrophysics*. Vol. 579, pp. A44. ISSN 1432-0746. Available from DOI: 10.1051/0004-6361/201525617.

Gabriel, C. et al., 2004. The XMM-Newton SAS - Distributed Development and Maintenance of a Large Science Analysis System: A Critical Analysis. In: Ochsenbein, F.; Allen, M. G.; Egret, D. (eds.). *Astronomical Data Analysis Software and Systems XIII*. Strasbourg, France: Astronomical Society of the Pacific Conference Series, vol. 314, p. 759. ISBN 1-58381-169-9.

Gilfanov, M.; Grimm, H.-J.; Sunyaev, R., 2004. Lx-SFR relation in star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*. Vol. 347, no. 3, pp. L57–L60. ISSN 0035-8711. Available from DOI: 10.1111/j.1365-2966.2004.07450.x.

Gilfanov, M.; Merloni, A., 2014. Observational Appearance of Black Holes in X-Ray Binaries and AGN. *Space Science Reviews*. Vol. 183, pp. 121–148. ISSN 1572-9672. Available from DOI: 10.1007/s11214-014-0071-5.

Heckman, Timothy M. et al., 2005. The Properties of Ultraviolet-luminous Galaxies at the Current Epoch. *The Astrophysical Journal*. Vol. 619, no. 1, pp. L35–L38. Available from DOI: 10.1086/425979.

Hoopes, Charles G. et al., 2007. The Diverse Properties of the Most Ultraviolet-Luminous Galaxies Discovered by GALEX. *The Astrophysical Journal Supplement Series*. Vol. 173, no. 2, pp. 441–456. Available from DOI: 10.1086/516644.

Izotov, Y. I.; Guseva, N. G.; Thuan, T. X., 2011. Green Pea Galaxies and Cohorts: Luminous Compact Emission-line Galaxies in the Sloan Digital Sky Survey. *The Astrophysical Journal*. Vol. 728, no. 2, pp. 161. ISSN 1538-4357. Available from DOI: 10.1088/0004-637x/728/2/161.

Jeon, M.; Pawlik, A. H.; Bromm, V.; Milosavljević, M., 2014. Radiative feedback from high-mass X-ray binaries on the formation of the first galaxies and early reionization. *Monthly Notices of the Royal Astronomical Society*. Vol. 440, no. 4, pp. 3778–3796. ISSN 0035-8711. Available from DOI: 10.1093/mnras/stu444.

Kaaret, P.; Feng, H.; Roberts, T. P., 2017. Ultraluminous X-Ray Sources. *Annual Review of Astronomy & Astrophysics*. Vol. 55, no. 1, pp. 303–341. Available from DOI: 10.1146/annurev-astro-091916-055259.

Kaastra, J. S.; Bleeker, J. A. M., 2016. Optimal binning of X-ray spectra and response matrix design. *Astronomy & Astrophysics*. Vol. 587, pp. A151. Available from DOI: 10.1051/0004-6361/201527395.

Kauffmann, G. et al., 2003. The host galaxies of active galactic nuclei. *Monthly Notices of the Royal Astronomical Society*. Vol. 346, pp. 1055–1077. Available from DOI: 10.1111/j.1365-2966.2003.07154.x.

Kewley, L. J.; Dopita, M. A.; Sutherland, R.S.; Heisler, C.A.; Trevena, J., 2001. Theoretical modeling of starburst galaxies. *The Astrophysical Journal*. Vol. 556, pp. 121–140. Available from DOI: 10.1086/321545.

Knevitt, G.; Wynn, G. A.; Power, C.; Bolton, J. S., 2014. Heating and ionization of the primordial intergalactic medium by high-mass X-ray binaries. *Monthly Notices of the Royal Astronomical Society*. Vol. 445, no. 2, pp. 2034–2048. ISSN 0035-8711. Available from DOI: 10.1093/mnras/stu1803.

Kunth, Daniel; Östlin, Göran, 2000. The most metal-poor galaxies. *Astronomy & Astrophysics Review*. Vol. 10, pp. 1–79. Available from DOI: 10.1007/s001590000005.

Lehmer, B. D.; Alexander, D. M.; Bauer, F. E.; Brandt, W. N.; Goulding, A. D.; Jenkins, L. P.; Ptak, A.; Roberts, T. P., 2010. A CHANDRA Perspective on Galaxy-wide X-ray Binary Emission and its Correlation with Star Formation Rate and Stellar Mass: New Results from Luminous Infrared Galaxies. *The Astrophysical Journal*. Vol. 724, no. 1, pp. 559–571. ISSN 1538-4357. Available from DOI: 10.1088/0004-637x/724/1/559.

Mikulášek, Zdeněk; Krtička, Jiří, 2005. *Základy fyziky hvězd*.

Mineo, S.; Gilfanov, M.; Sunyaev, R., 2012. X-ray emission from star-forming galaxies — II. Hot interstellar medium. *Monthly Notices of the Royal Astronomical Society*. Vol. 426, no. 3, pp. 1870–1883. ISSN 0035-8711. Available from DOI: 10.1111/j.1365-2966.2012.21831.x.

Mirabel, I. F.; Dijkstra, M.; Laurent, P.; Loeb, A.; Pritchard, J. R., 2011. Stellar black holes at the dawn of the Universe. *Astronomy & Astrophysics*. Vol. 528, pp. A149. Available from DOI: 10.1051/0004-6361/201016357.

Overzier, Roderik A. et al., 2008. Hubble Space TelescopeMorphologies of Local Lyman Break Galaxy Analogs. I. Evidence for Starbursts Triggered by Merging. *The Astrophysical Journal*. Vol. 677, no. 1, pp. 37–62. Available from DOI: 10.1086/529134.

Power, Ch.; James, G.; Combet, C.; Wynn, G., 2013. Feedback from High-Mass X-Ray Binaries on the High-Redshift Intergalactic Medium: Model Spectra. *The Astrophysical Journal*. Vol. 764, no. 1, pp. 76. Available from DOI: 10.1088/0004-637x/764/1/76.

Schneider, P., 2015. *Extragalactic Astronomy and Cosmology*. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN 978-3-642-54082-0. Available from DOI: 10.1007/978-3-642-54083-7.

Smith, M.; Guainazzi, M., 2016. *XMM-SOC-CAL-TN-0018*.

Svoboda, J.; Douna, V.; Orlitová, I.; Ehle, M., 2019. Green Peas in X-Rays. *The Astrophysical Journal*. Vol. 880, no. 2, pp. 144. Available from DOI: 10.3847/1538-4357/ab2b39.

Ward-Thompson, D.; Whitworth, A. P., 2011. *An Introduction to Star Formation*. Cambridge University Press. ISBN 978-0-521-63030-6. Available from DOI: 10.1017/CBO9780511974021.

Wilms, J.; Allen, A.; McCray, R., 2000. On the Absorption of X-Rays in the Interstellar Medium. *The Astrophysical Journal*. Vol. 542, no. 2, pp. 914–924. Available from DOI: 10.1086/317016.

York, Donald G. et al., 2000. The Sloan Digital Sky Survey: Technical Summary. *The Astronomical Journal*. Vol. 120, no. 3, pp. 1579–1587. ISSN 0004-6256. Available from DOI: 10.1086/301513.

Online resources

Arnaud, K.; Gordon, C.; Dorman, B., 2020. *Xspec: An X-Ray Spectral Fitting Package*. Available from: https://heasarc.gsfc.nasa.gov/docs/xanadu/xspec/manual/manual.html.

How to Generate Vignetting-Corrected Background-Subtracted Images, 2020. Available from: https://www.cosmos.esa.int/web/xmm-newton/sas-thread-images.

SAS Threads, 2020. Available from: https://www.cosmos.esa.int/web/xmm-newton/sas-threads.

SpecGroup, 2019. Available from: https://xmm-tools.cosmos.esa.int/external/sas/current/doc/specgroup.

XMM-Newton, 2020. Available from: https://sci.esa.int/web/xmm-newton.

XMM-Newton Current Calibration Files, 2020. Available from: https://www.cosmos.esa.int/web/xmm-newton/current-calibration-files.

Yusaf, Rehana, 1994. *GRPPHA*. Available from: https://heasarc.gsfc.nasa.gov/ftools/caldb/help/grppha.txt.

Appendix

A Details of the extraction regions

The source and extraction regions used in this work during the data reduction are shown in the Table 1 for 2019 observation of all three Green Peas and in Table 2 for 2019 observation of GP3. Additional background regions discussed in section 4.2 are in Table 3.

Table 1: Extraction regions for 2013 observation

Source	Detector	Source extraction region	Background extraction region
GP1	MOS1	117° 24′ 11.574″, +33° 37′ 16.388″	117° 25′ 02.000″, +33° 36′ 40.213″
	MOS2	"	117° 25′ 07.430″, +33° 36′ 43.830″
	PN	"	117° 25′ 17.261″, +33° 37′ 01.876″
GP2	MOS1	125° 41′ 54.924″, +22° 41′ 44.081″	125° 42′ 48.164″, +22° 41′ 13.941″
	MOS2	"	125° 40′ 59.466″, +22° 41′ 34.772″
	PN	"	125° 40′ 50.218″, +22° 41′ 59.578″
GP3	MOS1	204° 52′ 4.558″, +15° 16′ 42.128″	204° 48′ 23.828″, +15° 17′ 43.605″
	MOS2	"	204° 52′ 51.506″, +15° 19′ 42.888″
	PN	"	204° 54′ 04.938″, +15° 16′ 23.096″

Table 2: Extraction regions for GP3 2019 observation

Detector	Source extraction region	Background extraction region
MOS1	204° 52′ 04.558″, +15° 16′ 42.128″	204° 53′ 39.803″, +15° 16′ 04.359″
		204° 51′ 03.759″, +15° 15′ 14.428″
MOS2	204° 52′ 04.558″, +15° 16′ 42.128″	204° 54′ 02.328″, +15° 14′ 31.878″
		204° 53′ 31.939″, +15° 15′ 30.558″
PN	204° 52′ 04.558″, +15° 16′ 42.128″	204° 53′ 14.721″, +15° 16′ 15.483″
		204° 53′ 06.688″, +15° 16′ 44.070″

Table 3: Additional background extraction regions

Source and observation year	Detector	Background extraction region	Radius arcsec
GP3 2013	MOS	204° 53′ 41.910″, +15° 14′ 29.700″	120
	PN	204° 53′ 41.910″, +15° 16′ 48.430″	60
GP3 2019	MOS1	204° 53′ 59.198″, +15° 13′ 52.042″	80
	MOS2	204° 53′ 33.402″, +15° 14′ 32.655″	80
	PN	204° 53′ 14.787″, +15° 17′ 34.854″	50

B Bash skript for data analysis

There is an example of bash skript used for the data analysis. The following skript is written for data reduction of data originating from MOS1 detector for GP1 archival observation.

```
evselect table=2434_0690470101_EMOS1_S001_ImagingEvts.ds
        withrateset=Y rateset=MOS1rates.ds maketimecolumn=Y
        timebinsize=100 makeratecolumn=Y
        expression='#XMMEA_EM && (PI>10000) && (PATTERN==0)'
tabgtigen table=MOS1rates.ds expression='RATE<=0.35'
        gtiset=MOS1gti.ds
evselect table=2434_0690470101_EMOS1_S001_ImagingEvts.ds
        withfilteredset=Y filteredset=MOS1clean.ds destruct=Y
        keepfilteroutput=T expression='#XMMEA_EM &&
        gti(MOS1gti.ds,TIME) && (PI>150)'
evselect table=MOS1clean.ds imagebinning=binSize
        imageset=MOS1image.ds withimageset=yes xcolumn=X
        ycolumn=Y ximagebinsize=80 yimagebinsize=80
evselect table=MOS1clean.ds withspectrumset=yes
        spectrumset=MOS1source_spectrum.ds energycolumn=PI
        spectralbinsize=5 withspecranges=yes specchannelmin=0
        specchannelmax=11999 expression='#XMMEA_EM &&
        (PATTERN<=12) && ((X,Y) IN circle(26555.13,27912.801,480))'
evselect table=MOS1clean.ds withspectrumset=yes
        spectrumset=MOS1bg_spectrum.ds energycolumn=PI
        spectralbinsize=5 withspecranges=yes specchannelmin=0
        specchannelmax=11999 expression='#XMMEA_EM &&
        (PATTERN<=12) && ((X,Y) IN circle(25715.301,27189.256,480))'
backscale spectrumset=MOS1source_spectrum.ds
        badpixlocation=MOS1clean.ds
backscale spectrumset=MOS1bg_spectrum.ds
        badpixlocation=MOS1clean.ds
rmfgen spectrumset=MOS1source_spectrum.ds rmfset=MOS1.rmf
arfgen spectrumset=MOS1source_spectrum.ds arfset=MOS1.arf
        withrmfset=yes rmfset=MOS1.rmf
        badpixlocation=MOS1clean.ds detmaptype=psf
```