

Metallicity dependence of HMXB populations

V. M. Douna^{1,2,*}, L. J. Pellizza³, I. F. Mirabel^{1,4}, and S. E. Pedrosa¹

¹ Instituto de Astronomía y Física del Espacio (IAFE), UBA-CONICET, CC 67, Suc. 28, (C1428ZAA), Buenos Aires, Argentina.

² Facultad de Cs. Exactas y Naturales (FCEN), UBA, Buenos Aires, Argentina.

³ Instituto Argentino de Radioastronomía (IAR), CCT-La Plata, CONICET, C.C.5., 1894, Villa Elisa, Argentina.

⁴ CEA-Saclay, IRFU/DSM/Service d'Astrophysique, 91191 Gif/Yvette, France.

Received; accepted

ABSTRACT

Context. High-mass X-ray binaries (HMXBs) might have contributed a non-negligible fraction of the energy feedback to the interstellar and intergalactic media at high redshift, becoming important sources for the heating and ionization history of the Universe. However, the importance of this contribution depends on the hypothesized increase in the number of HMXBs formed in low-metallicity galaxies and in their luminosities.

Aims. In this work we test the aforementioned hypothesis, and quantify the metallicity dependence of HMXB population properties.

Methods. We compile from the literature a large set of data on the sizes and X-ray luminosities of HMXB populations in nearby galaxies with known metallicities and star formation rates. We use Bayesian inference to fit simple Monte Carlo models that describe the metallicity dependence of the size and luminosity of the HMXB populations.

Results. We find that HMXBs are typically ten times more numerous per unit star formation rate in low-metallicity galaxies ($12 + \log(\text{O}/\text{H}) < 8$, namely $< 20\%$ solar) than in solar-metallicity galaxies. The metallicity dependence of the luminosity of HMXBs is small compared to that of the population size.

Conclusions. Our results support the hypothesis that HMXBs are more numerous in low-metallicity galaxies, implying the need to investigate the feedback in the form of X-rays and energetic mass outflows of these high-energy sources during cosmic dawn.

Key words. X-rays: binaries – galaxies: star formation – galaxies: abundances

1. Introduction

It has been proposed that the energy feedback of high-mass X-ray binaries (HMXBs) plays a major role in the thermal and ionization histories of the Universe (Power et al. 2009; Mirabel et al. 2011; Jeon et al. 2013; Power et al. 2013; Kneivitt et al. 2014), and that it significantly affects the cosmic star formation rate and the evolution of low-mass galaxies (Justham & Schawinski 2012; Artale et al. 2015). The importance of these sources is based on the hypothesis that they are more numerous, and possibly more luminous, in low-metallicity stellar populations. This, coupled with the chemical evolution of the Universe, suggests that in the early Universe the contribution of HMXBs to the energy feedback into the interstellar and intergalactic media (ISM and IGM) would have been at least as large as that of other high-energy feedback sources, such as supernovae.

However, direct evidence for the increase in the number and luminosity of HMXBs in low-metallicity stellar populations is still scarce. Grimm et al. (2003) and Mineo et al. (2012) have carried out exhaustive observational studies, compiling statistically significant samples of HMXBs in local galaxies. They found that the total X-ray luminosity of a HMXB population (L_X) scales with the star formation rate (SFR) of its host galaxy, which at first sight seems to imply that there are no metallicity effects. However, their sample includes mainly high-SFR galaxies that, owing to the correlation between SFR and metallicity, would be

high-metallicity systems. Hence, no metallicity effects are expected. Trying to overcome this limitation, Kaaret et al. (2011) and Brorby et al. (2014) studied a sample of local low-SFR, low-metallicity blue compact dwarfs (BCDs). They found that these galaxies host HMXB populations ten times more numerous per unit SFR, than galaxies in the sample of Mineo et al. (2012).

On the other hand, Basu-Zych et al. (2013a) have shown that the L_X -SFR relation evolves with redshift, which could be due to metallicity effects. Fragos et al. (2013) used population synthesis models coupled to large scale cosmological simulations to investigate the emission of HMXB populations at different redshifts. They found that the integrated X-ray luminosity per unit SFR increases with redshift, exceeding that of other X-ray sources (low-mass X-ray binaries, active galactic nuclei) at high redshifts. In addition, Basu-Zych et al. (2013b) showed that the relation between the ratio L_X/SFR and the gas-phase metallicity of a sample of $z < 0.1$ Lyman-break analogs with slightly sub-solar metallicities ($12 + \log(\text{O}/\text{H}) \gtrsim 8.1$) is consistent with the theoretical expectations of Fragos et al. (2013).

Although previous works showed the existence of a metallicity dependence of HMXB populations, a discrimination of its effects on the size of these populations from those on the intrinsic luminosities is still lacking. To explore this issue, in this work we compile from the literature a large sample of local galaxies for which individual HMXBs can be resolved. Our sample extends those of previous works (Mineo et al. 2012; Brorby et al. 2014) by including homogeneous estimations of the metallicities of the

* e-mail: vdouna@iafe.uba.ar

galaxies, hence allowing the study of a larger metallicity range. Moreover, our analysis extends the work of previous authors by requiring our models to match both the data of the HMXB population sizes and luminosities. In this paper we present the construction of the observational sample (Sect. 2), and investigate the dependence on metallicity of the properties of HMXB populations (Sect. 3). We discuss the results and present our conclusions in Sect. 4.

2. The sample

In order to investigate the dependence of the properties of HMXB populations on the metallicity of their progenitors, for 49 galaxies in the Local Universe we compiled from the literature the number of binaries, the integrated X-ray luminosities, star formation rates (SFR) and metallicities. Following Mineo et al. (2012), we selected galaxies that satisfy two criteria: a) high specific star formation rate ($sSFR = SFR/M_*$), and b) small distance. As the production of HMXBs depends on the SFR of the galaxy (Grimm et al. 2003), and the number of low-mass X-ray binaries is proportional to its stellar mass M_* (Gilfanov 2004), the first criterion is useful to select galaxies in which HMXBs dominate the population of X-ray binaries. In particular, we chose $sSFR > 10^{-10} \text{ yr}^{-1}$. The distance criterion was imposed because we are interested in galaxies where individual HMXBs can be resolved. The galaxies in our sample were chosen by requiring that their distance is $D < 65 \text{ Mpc}$, or their redshift is $z < 0.015$.

The main source of objects for our sample is the work of Mineo et al. (2012). This work compiles the number of sources, luminosities (both of individual XRBs and the total integrated luminosity of the galaxy), and star formation rates for 29 nearby ($D < 40 \text{ Mpc}$), star-forming galaxies with $SFR \sim 0.1 - 20 M_\odot$. The sample of these authors is based on a homogeneous set of observations in X-rays, infrared and UV from *Chandra*, *Spitzer*, *GALEX* and *2MASS*, with very specific selection criteria. We extended their sample in two ways. First, we compiled from the literature the metallicities of the galaxies when available. This left us with only 19 galaxies, most of them being high-metallicity systems. As a second step we added to this sample a set of low-metallicity galaxies because the investigation of a large abundance range is needed to assess clearly the effects of metallicity on HMXB populations. The low-metallicity galaxies come mainly from the work of Brorby et al. (2014), who studied a set of 25 nearby ($D < 30 \text{ Mpc}$) low-metallicity BCDs. Another 5 galaxies from different works (Ghigo et al. 1983; Thuan et al. 2004; Winter et al. 2006; Kaaret et al. 2011) were also added. We note that 20 of the 49 galaxies in our sample have no HMXBs detected; however these data still provide useful constraints on the metallicity dependence, as we will show in Appendix A. The compilation of abundances, together with the extension of the metallicity range and the use of both the observed population sizes and integrated luminosities, would allow us to go beyond the results of previous works. As the low- and high-metallicity samples were taken from different sources, with different uncertainties and selection effects, a standardization of the data is needed to make statistical comparisons and draw meaningful conclusions. In the following sections we describe the steps followed to make the data homogeneous, and the caveats involved in the process.

2.1. Number of binaries and their luminosity

As all the works from which we construct our sample give the data of individual sources in the galaxy, we took directly the observed number of sources N and their individual X-ray luminosities from the aforementioned works (the complete data of our sample is given in Table 1). The luminosity threshold L_{th} above which sources are detected was also recorded. Although we cannot ensure that all the sources listed are indeed HMXBs, the selection criteria imposed minimize the contamination by other types of sources. In the case of the data taken from Mineo et al. (2012), we only used objects with location flag equal to 1 in their Table A1, which are those sources lying in the HMXB-dominated regions of the galaxies. For the data of low-metallicity galaxies, the association of the few X-ray sources found with star-forming regions makes highly improbable of them being other kind of sources.

X-ray luminosities of individual sources are given in different energy ranges in the literature, according to the instrument used to observe them. To make the data homogeneous, we corrected all the observations to the $0.5 - 8 \text{ keV}$ band, which is that of most observations of our sample. To perform the correction, we assumed that the X-ray spectrum of individual sources is a power-law, and used for each source the spectral index reported in the literature. In the case it is not available, we took the standard value -1.7 (Fabbiano 2006). The luminosity thresholds of the observations were corrected accordingly. Throughout this paper, we will express all the luminosities in the $0.5 - 8 \text{ keV}$ band.

The total X-ray luminosity L_X^G of each galaxy in our sample is taken as the sum of the luminosities L_i of individual sources above L_{th} . However, the latter differs for each galaxy. Aiming at obtaining a homogeneous luminosity sample and following Mineo et al. (2012), we also computed the total luminosities of the galaxies corrected to a common threshold $L_{\text{th},1} = 10^{36} \text{ erg s}^{-1}$. In the case of NGC 1569, as $L_{\text{th}} < L_{\text{th},1}$, this correction was made by subtracting the luminosity of the sources below $L_{\text{th},1}$. For the remaining galaxies, we used a power-law X-ray luminosity function (XLF) with the index ($\alpha = -1.58$) and normalization factor ($q = 1.49 (10^{38} \text{ erg s}^{-1})^{0.58} M_\odot^{-1} \text{ yr}$) given by Mineo et al. (2012), hence obtaining the corrected luminosity

$$L_X^{G,36} = L_X^G + q \text{ SFR} \int_{L_{\text{th},1}}^{L_{\text{th}}} L^{\alpha+1} dL. \quad (1)$$

In most cases, the integral in the previous equation amounts to less than 11% of the final value, except for UGC 5720 (19%), NGC 4194 (28%), and NGC 3079 (29%). We conservatively assumed here that the XLF does not depend on metallicity. This choice is not crucial because the corrections are small, owing to the weak dependence of the integrated luminosity on L_{th} . However, if the higher normalization factor given by Brorby et al. (2014) were used for low-metallicity galaxies, a slightly larger total luminosity would have been found, making our conclusions stronger.

It should be pointed out that the aforementioned correction on the luminosities was not applied to the galaxies without detected sources. In this case, any trend found in the corrected values would reflect the corresponding trends of the model used in Eqn. 1 for the XLF, rather than those of the original data, hence biasing the results.

Following the above reasoning, we also corrected the number of sources detected in each galaxy to a common threshold $L_{\text{th},2} = 10^{38} \text{ erg s}^{-1}$. The choice of this threshold value was dictated by the fact that the observed number of sources does de-

Galaxy	D [Mpc]	$12 + \log(\text{O}/\text{H})$	SFR [$M_{\odot}\text{yr}^{-1}$]	$\log L_X^G$ [erg s^{-1}]	$\log L_{\text{th}}$ [erg s^{-1}]	N	$\log L_X^{G,36}$ [erg s^{-1}]	$\log L_X^{G,38}$ [erg s^{-1}]	N^{38}
DDO 68 ⁱ	5.9	7.15	0.020	38.04	37.5	2	38.06	-	0
IZw 18 ⁱ	18.2	7.18	0.068	39.58	37.87	1	39.58	39.58	1
SBS 0335-052	54.3 ^l	7.29 ^a	0.44 [*]	39.52 ^l	38.95 ^l	1	39.57	39.55	2
SBS 1129+576 ⁱ	26.3	7.41	0.063	39.01	38.62	1	39.02	39.01	1
SBS 0940+544 ⁱ	22.1	7.48	0.022	38.99	38.45	1	39.00	38.99	1
RC2A 1116+51 ⁱ	20.8	7.51	0.040	39.51	38.52	1	39.51	39.51	1
HS1442+4250 ⁱ	8.67	7.64	0.015	38.34	38.11	1	38.35	38.34	1
RC2A 1228+12 ⁱ	21.2	7.64	0.031	38.53	38.53	1	38.55	38.54	1
VIIZw 403 ⁱ	3.87	7.69	0.014	38.62	37.14	2	38.62	38.60	1
Mrk 71	3.4 ²	7.85 ^b	0.041 ^{**}	38.34 ²	38.3 ²	1	38.37	38.35	1
NGC 4631 ⁱⁱ	7.6	8.13 ^c	4	39.71	37.09	23	39.74	39.65	6
NGC 1569 ⁱⁱ	1.9	8.19 ^d	0.078	37.95	35.71	10	37.93	-	0
NGC 1313 ⁱⁱ	4.1	8.23 ^e	0.44	39.69	36.82	8	39.69	39.68	4
NGC 5253 ⁱⁱ	4.1	8.24 ^f	0.38	38.60	36.07	17	38.60	38.31	1
NGC 4625 ⁱⁱ	8.2	8.27 ^c	0.09	38.01	37.01	4	38.04	-	0
NGC 2403 ⁱⁱ	3.1	8.34 ^g	0.52	39.41	36.16	42	39.41	39.30	2
NGC 4214 ⁱⁱ	2.5	8.36 ^h	0.17	38.41	36.19	14	38.42	38.17	1
NGC 4490 ⁱⁱ	7.8	8.36 ⁱ	1.8	40.06	37.1	32	40.07	40.03	11
NGC 3034 ⁱⁱ	3.9	8.36 ^c	10.5	39.86	36.86	54	39.90	39.75	12
NGC 4038-39 ⁱⁱ	13.8	8.4 ^j	5.4	40.25	36.92	83	40.26	40.20	20
NGC 7793 ⁱⁱ	4	8.4 ^k	0.29	38.36	36.55	9	38.38	38.20	1
NGC 3310 ⁱⁱ	19.8	8.44 ^l	7.1	40.77	37.84	23	40.79	40.77	22
UGC 5720 ⁱⁱ	24.9	8.4 ^m	1.8	39.57	38.36	4	39.66	39.60	6
NGC 5457 ⁱⁱ	6.7	8.55 ⁿ	1.5	39.60	36.36	96	39.61	39.49	11
NGC 3079 ⁱⁱ	18.2	8.57 ^h	6	39.65	37.98	14	39.80	39.65	14
NGC 5194 ⁱⁱ	7.6	8.58 ^o	3.7	40.01	37.05	69	40.02	39.91	16
NGC 5775 ⁱⁱ	26.7	8.64 ^p	5.3	40.52	38.01	25	40.54	40.52	25
NGC 4194 ⁱⁱ	39.1	8.67 ^h	16.8	40.42	38.64	4	40.56	40.49	29
NGC 1672 ⁱⁱ	16.8	8.97 ^q	12	40.45	37.74	25	40.49	40.45	22
J081239.52+483645.3 ⁱ	9.04	7.16	0.002	-	38.21	0	-	-	0
UGC 772 ⁱ	11.5	7.24	0.012	-	38.38	0	-	-	0
J 210455.31-003522.2 ⁱ	13.7	7.26	0.007	-	38.56	0	-	-	0
UGCA 292 ⁱ	3.5	7.27	0.002	-	37.34	0	-	-	0
J 141454.13-020822.9 ⁱ	24.6	7.32	0.011	-	38.53	0	-	-	0
6dFJ 0405204-364859 ⁱ	11	7.34	0.013	-	38.33	0	-	-	0
HS 0822+3542 ⁱ	12.7	7.35	0.004	-	38.48	0	-	-	0
J 085946.92+392305.6 ⁱ	10.9	7.45	0.002	-	38.36	0	-	-	0
KUG 0937+298 ⁱ	11.2	7.45	0.015	-	38.36	0	-	-	0
UGC 4483 ⁱ	3.44	7.54	0.004	-	37.55	0	-	-	0
J 120122.3+021108.5 ⁱ	18.4	7.55	0.008	-	38.58	0	-	-	0
KUG 0201-103 ⁱ	22.7	7.56	0.015	-	38.54	0	-	-	0
KUG 1013+381 ⁱ	19.6	7.58	0.073	-	38.56	0	-	-	0
SBS 1415+437 ⁱ	13.7	7.59	0.040	-	38.52	0	-	-	0
SBS 1102+606 ⁱ	19.9	7.64	0.035	-	38.54	0	-	-	0
KUG 0942+551 ⁱ	24.4	7.66	0.019	-	38.52	0	-	-	0
KUG 0743+513 ⁱ	8.6	7.68	0.032	-	38.14	0	-	-	0
Mrk 209 ⁱⁱⁱ	5.7	7.77 ^a	0.025	-	37.96	0	-	-	0
UM 461 ⁱⁱⁱ	13.4	7.85 ^r	0.010	-	37.66	0	-	-	0
II Zw 40	14 ³	8 ^b	1.32 ^{***}	-	39.86 ³	0	-	-	3

Notes. References: ⁽ⁱ⁾ Brorby et al. (2014), ⁽ⁱⁱ⁾ Mineo et al. (2012), ⁽ⁱⁱⁱ⁾ Kaaret et al. (2011), ^(a) Izotov & Thuan (1999), ^(b) Izotov & Thuan (2011), ^(c) Calzetti et al. (2007), ^(d) Kobulnicky & Skillman (1997), ^(f) López-Sánchez & Esteban (2010), ^(g) Esteban et al. (2009), ^(h) Engelbracht et al. (2008), ⁽ⁱ⁾ Pilyugin & Thuan (2007), ^(j) Mirabel et al. (1992), ^(k) Bibby & Crowther (2010), ^(l) Denicoló et al. (2002), ^(m) Hirashita et al. (2002), ⁽ⁿ⁾ Cedres et al. (2004), ^(o) Bresolin et al. (2004), ^(p) Werk et al. (2011), ^(q) Storchi-Bergman et al. (1996), ^(r) Kniazev et al. (2004), ^(s) Thuan et al. (2004), ⁽²⁾ Winter et al. (2006), ⁽³⁾ Ghigo et al. (1983), ^(*) Prestwich et al. (2013), ^(**) Hopkins et al. (2002), ^(***) Sage et al. (1992).

Table 1. Distances, oxygen abundances, SFRs, X ray luminosities, and observed number of HMXBs for the galaxies in our sample.

pend strongly on L_{th} . Hence, we took a large value so that for most galaxies $L_{\text{th}} \leq L_{\text{th},2}$, and the number of sources N^{38} above $L_{\text{th},2}$ is determined by simple counting. Only for 26 of our 49 galaxies is $L_{\text{th}} > L_{\text{th},2}$ and then a correction based on the XLF is needed,

$$N^{38} = N + q \text{SFR} \int_{L_{\text{th},2}}^{L_{\text{th}}} L^\alpha dL. \quad (2)$$

For 22 of them, the integral in the right-hand side of Eqn. 2 is $\ll 1$, and $N^{38} = N$. In the few remaining cases, the corrections amount to 33% (UGC 5720), 50% (SBS 0335-052), 86% (NGC 4194), and 100% (II Zw 40) of the final value. The luminosity $L_X^{G,38}$ was determined in the same way as $L_X^{G,36}$, but using $L_{\text{th},2}$ instead of $L_{\text{th},1}$ in Eqn. 1. The corrections in this case are lower than 7%, except for NGC 4194 (15%).

The corrections made by Eqns. 1 and 2 allow us to obtain a uniform sample of the numbers of sources and luminosities of the galaxies. It is important to note that most of these corrections are small, and that they do not affect our final results, as N^{38} , $L_X^{G,38}$, and $L_X^{G,36}$ are used only to visualize and derive preliminary trends. The full Bayesian analysis of the data, from which our final results are obtained and our conclusions are drawn, relies on the original data (N , L_X^G , and L_{th}).

2.2. Star formation rates

We extracted the star formation rates (SFRs) of 19 of the galaxies of our sample from Mineo et al. (2012). These authors use a combined estimator for the SFR based on the near-UV (NUV) radiation from young stars, corrected by the IR emission due to heating of dust,

$$\text{SFR} = \text{SFR}_{\text{NUV}} + (1 - \eta)\text{SFR}_{\text{IR}}, \quad (3)$$

where η is a correction factor related to the IR emission coming from the old stellar population, whose value is null for starburst galaxies and 0.4 for normal disk galaxies. The NUV and IR contributions to SFR (in $M_\odot \text{yr}^{-1}$) can be calculated from the 2312Å and 8 – 1000µm luminosities (in erg s^{-1}) as $\text{SFR}_{\text{IR}} = 4.6 \times 10^{-44} L_{\text{IR}}$ and $\text{SFR}_{\text{NUV}} = 1.2 \times 10^{-43} L_{\text{NUV}}$, respectively. In addition to the galaxies taken from the work of Mineo et al. (2012), we adopted this estimator for those galaxies for which NUV and IR observations are available. For the rest of the galaxies we used far-UV (FUV) SFR_{FUV} estimations corrected by a factor of 1.23. According to Prestwich et al. (2013) and Brorby et al. (2014), this factor gives the ratio of FUV estimations to those based on NUV and IR observations. For a single galaxy (II Zw 40) the only available SFR estimation is based on H_α measurements. In this case, we transformed this value into a FUV SFR using the conversion of Hunter et al. (2010), and then once again using the factor 1.23 to transform it to the NUV/IR SFR scale. In this way, we obtained a sample as homogeneous as possible to compare the effects of SFR in the production of HMXBs.

2.3. Metallicities

The metallicity estimations for our sample were taken directly from the literature. As we are dealing with star-forming galaxies, in which emission lines are used to estimate abundances, we

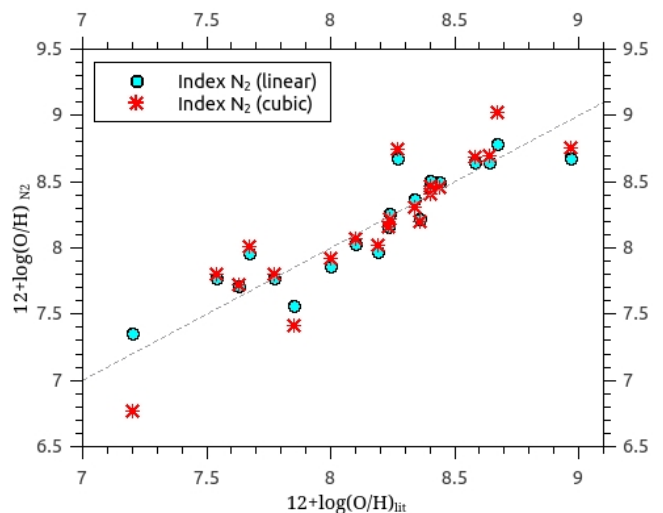


Fig. 1. Oxygen abundances calculated from raw spectral data by using both fits to the N_2 ratio given by Pettini & Pagel (2004), as a function of the abundances taken directly from the literature. Cyan dots (red stars) correspond to the linear (cubic) calibration of the N_2 index; the dashed line represents the identity.

use the oxygen abundances $12 + \log(\text{O}/\text{H})$ as a metallicity indicator. The abundance values compiled are *a priori* far from homogeneous, as they were obtained with different methods whose systematic differences are still unclear (Bresolin et al. 2004; Garnett et al. 2004; Liang et al. 2006; Yin et al. 2007a,b; Kewley & Ellison 2008; López-Sánchez et al. 2012). Although it is still not certain which method is better to estimate the absolute oxygen abundance of a galaxy, it is believed that the relative metallicities between galaxies estimated with the same calibration are reliable (Kewley & Ellison 2008). However, due to the large metallicity range of the galaxies in our sample, it was not possible to calculate the oxygen abundances with a single method because of the lack of homogeneity of the spectral lines observed.

It is important to note that in this work we are not interested in the computation of absolute oxygen abundances, but in the correlation of metallicity with the X ray emission of galaxies. With this aim and in order to evaluate the impact of comparing abundances estimated by different calibrations, we searched the literature for the raw spectral information of each galaxy, and built a homogeneous subsample of abundances computed by a single method. To this aim we used the intensities of the spectral lines $[\text{NII}] \lambda 6583$ and H_α , and computed the abundances of 21 galaxies using the N_2 index (Pettini & Pagel 2004). This method is single-valued, and is based on a ratio between lines that are very near in the spectrum, hence it does not depend on the reddening correction or flux calibration.

For those 21 galaxies of our homogeneous subsample, we calculated the abundances both using the linear and cubic fits for the N_2 ratio given by Pettini & Pagel (2004). The correlation between both N_2 estimations and the abundances taken directly from the literature are shown in Fig. 1. The variance between them is 0.17 dex in the case of the linear method and 0.22 dex for the cubic fit. Hence, despite the possible systematic deviations between the different calibrations, the general trend observed supports the use of the complete sample to estimate relative metallicities.

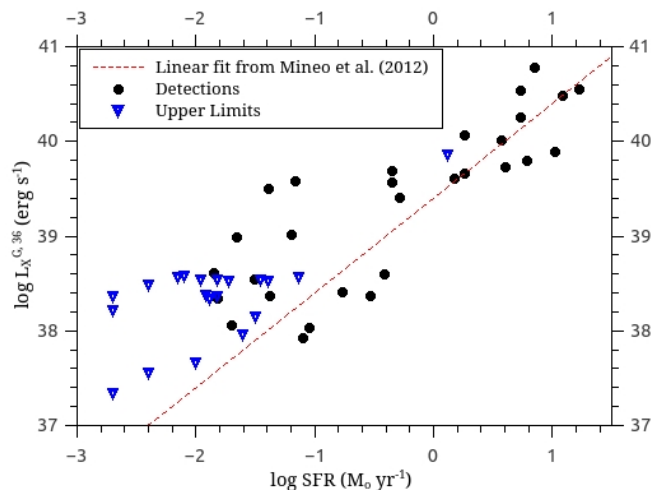


Fig. 2. X-ray luminosity as a function of SFR for the sample of galaxies compiled in this work. Black circles correspond to those galaxies with measured X-ray luminosities, while blue triangles represent the detection limits for single sources. These can be regarded as upper limits for the galaxy luminosity in the case it had only one source. The red dashed line ($\log L_X = \log \text{SFR} + 39.4$) is the relation found by Mineo et al. (2012). It describes well the data at high SFRs, whereas it fails to do so in the opposite SFR range. For $\log \text{SFR} \lesssim -1.1$, all the data fall above the fit, which suggests the existence of metallicity effects in the sample.

3. Analysis of the sample

3.1. Observed trends

In Fig. 2 we show the L_X –SFR relation for the galaxies in our sample, together with the best fit of Mineo et al. (2012, $\log L_X = \log \text{SFR} + 39.4$). In order to make a meaningful comparison, in this plot we use $L_X^{G,36}$ so that all galaxies have the same luminosity threshold, which is also the one used by these authors. It is evident that their best fit describes properly the mean behavior of high-SFR galaxies, mostly taken from their work, but fails to describe the correlation at low SFRs. In this limit, the X-ray luminosities are higher than those predicted by the fit. Indeed, all galaxies with $\log \text{SFR} \lesssim -1.1$ fall above the fit of Mineo et al. (2012). The fit also fails to describe the large dispersion of the data in the whole metallicity range. As the SFR of galaxies is correlated to their mean metallicity, the increase in the X-ray luminosity with respect to the fit of Mineo et al. (2012) observed at low SFRs could be due to metallicity effects.

However, it could also be due to statistical fluctuations, as low-SFR galaxies have small HMXB populations. The increase in the dispersion at low SFRs suggests that this effect is present in our sample. Disentangling both effects is crucial to reach a meaningful conclusion about the metallicity dependence of HMXB populations.

To further investigate the influence of metallicity on the L_X –SFR relation, in Fig. 3 we show the dependence of the $L_X^{G,36}/\text{SFR}$ ratio on the oxygen abundance of the galaxies. Despite the high dispersion, the observed behavior is consistent with an anticorrelation between the X-ray luminosity per unit SFR and the oxygen abundance. It is also apparent that a linear relation is a good description of the correlation. The best linear fit (not including upper limits) is $\log L_X^{G,36} = 39.26 + \log \text{SFR} - 1.01[\log(\text{O}/\text{H}) - \log(\text{O}/\text{H})_\odot]$, taking $12 + \log(\text{O}/\text{H})_\odot = 8.69$ from Asplund et al. (2009). The $L_X^{G,36}/\text{SFR}$ ratio is similar to that of Mineo et

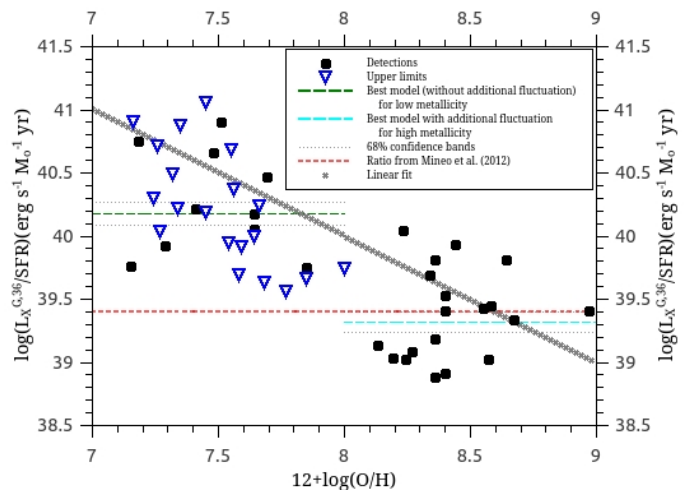


Fig. 3. X-ray luminosity per unit SFR as a function of the oxygen abundance of the galaxies in our sample. Black filled squares correspond to those galaxies with measured X-ray luminosities, while blue triangles represent the detection limits for single sources. These can be regarded as upper limits for the galaxy luminosity in the case it had only one source. The red short-dashed line represents the value found by Mineo et al. (2012) for solar-metallicity galaxies ($L_X^{G,36}/\text{SFR} = 39.4$), while the gray dotted line is the best linear fit to our sample, $\log L_X^{G,36} = 39.26 + \log \text{SFR} - 1.01[\log(\text{O}/\text{H}) - \log(\text{O}/\text{H})_\odot]$. A clear trend of decreasing $L_X^{G,36}/\text{SFR}$ with oxygen abundance is observable, which is a direct evidence of the existence of metallicity effects in HMXB populations. Stochasticity effects are also evident, and require a more refined treatment to make a robust prediction of the variation of the $L_X^{G,36}/\text{SFR}$ ratio with metallicity. The cyan and green long-dashed lines above and below $12 + \log(\text{O}/\text{H}) = 8$ respectively, represent our best models for the high- and low-metallicity subsamples, with the 68% confidence bands for the parameter Q (small-dotted gray lines).

al. (2012) at solar metallicity, while it increases about 1 dex per dex of decrease in abundance. However, the dispersion around the regression line is still important (~ 0.5 dex), suggesting that stochasticity effects should be investigated to make a more robust prediction of the variation with metallicity of the X-ray luminosity per unit SFR of galaxies.

The increase in the L_X/SFR value toward low oxygen abundances might be due to an increase either in the HMXB population size, or in the mean HMXB luminosity. To disentangle these two physical effects, in Fig. 4 we show the observed number of sources N^{38} with luminosities above the fixed threshold $L_{\text{th},2} = 10^{38} \text{ erg s}^{-1}$, per unit SFR, as a function of the metallicity of the galaxies. A clear trend is seen, in the sense that this number is lower for high-metallicity galaxies ($\log(\text{O}/\text{H}) > 8$) than for low-metallicity ones ($\log(\text{O}/\text{H}) < 8$). The difference is about 1 dex. The presence of many data points with error bars much smaller than this difference suggests that it cannot be explained by statistical effects. As the observed number of sources depends mainly on the population size, and only weakly in the HMXB luminosities, this result suggests that the main metallicity effect is a change in the total number of HMXBs per unit SFR. Indeed, the observed difference in N^{38}/SFR from low- to high-metallicity galaxies is roughly the same amount as the change in $L_X^{G,36}/\text{SFR}$ derived from Fig. 3. This implies that the increase in the integrated luminosity of galaxies at low metallicities can be explained by the increase in the size of their HMXB populations. An increase in the mean luminosity of HMXBs, if

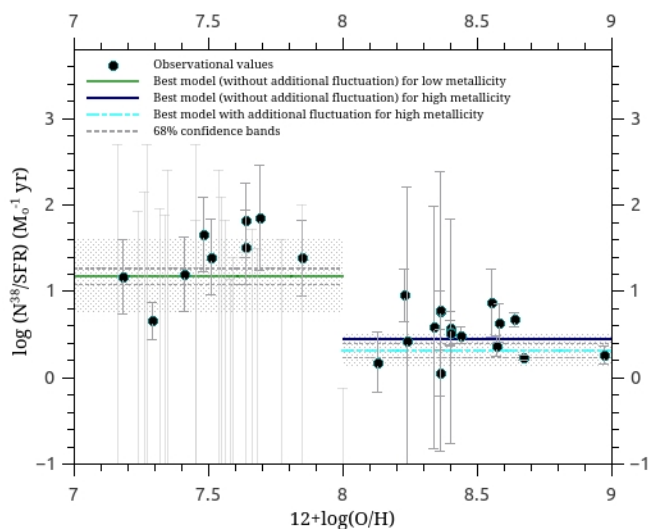


Fig. 4. Observed number of HMXBs with luminosities higher than 10^{38} erg s $^{-1}$, per unit SFR, as a function of the oxygen abundance. Black circles are the data points, with error bars representing Poissonian uncertainties. Large error bars reaching the lower end of the plot represent galaxies with no detected sources; their upper ends correspond to $N^{38} = 1$. Our best model for low-metallicity galaxies ($12 + \log(\text{O}/\text{H}) < 8$, green solid line, $Q = 3.13$, $\sigma_{\text{extra}} = 0$) predicts ten times more HMXBs than the corresponding best model for high-metallicity galaxies (cyan dot-dashed line, $Q = 2.14$, $\sigma_{\text{extra}} = 22$). For both models, the 68% confidence bands for the parameter Q and the dispersion predicted for the data are shown (gray dashed lines and shaded regions, respectively). The mean $Q = 2.28$ for high-metallicity galaxies with $\sigma_{\text{extra}} = 0$ is also plotted (blue solid line). The difference between low- and high-metallicity galaxies cannot be explained by statistical effects.

any, is certainly much lower. Once again, the dispersion of the N^{38}/SFR data is high, indicating the presence of stochasticity effects which must be taken into account to obtain a more robust prediction of the amount of the observed variations with metallicity.

3.2. Stochasticity effects

As pointed out before, all the data shown in this section present large dispersions. Part of them may be due to the stochasticity of the observed HMXB populations, which affects both their sizes and observed luminosities. Indeed, it has been shown by Gilfanov et al. (2004) that the collective luminosity of a small population of randomly chosen sources can grow non-linearly with the number of sources, contrary to the prediction of models based on infinite populations. To obtain a robust description of the trends present in the data, we developed a simple stochastic model for the number of sources and total luminosity of the population of HMXBs of a star-forming galaxy with a given SFR and oxygen abundance, which attempts to describe the observed behavior (see Appendix A for its details). Following Brorby et al. (2014), we assume that the mean number of sources in a galaxy follows a Poisson distribution whose mean scales with its SFR,

$$\langle N_{\text{HMXB}} \rangle = Q \text{SFR}, \quad (4)$$

where Q is a free parameter that may depend on the oxygen abundance. We also assume for the HMXBs a standard power-law XLF with index $\alpha = -1.6$ in the range $[L_{\text{min}}, L_{\text{max}}] =$

$[10^{35}, 10^{40}]$ erg s $^{-1}$ (Fabbiano 2006; Mineo et al. 2012). The observed number of sources and integrated luminosity of a particular galaxy above any specified threshold luminosity can be computed using the XLF. We note that our model includes only the dependence of the population size on metallicity, and not that of HMXB luminosities, as the XLF is fixed. In addition, our model includes two fluctuation sources, an intrinsic one due to the stochasticity of the star formation process, and another one given by the detection process, which selects only those binaries with luminosities above L_{th} . Model fitting was done by a Markov-chain Monte Carlo method, which gives the best-fit value of Q , together with its variance.

As Brorby et al. (2014) proposed that Q is different for low- and high-metallicity galaxies, we divide our sample into two subsamples comprising galaxies below and above $12 + \log(\text{O}/\text{H}) = 8$, respectively. We fit our model to each sample, obtaining $\log Q_{\text{low}} = 3.13 \pm 0.13$ and $\log Q_{\text{high}} = 2.28 \pm 0.02$ (which correspond to $q_{\text{low}} = 12.8$ and $q_{\text{high}} = 1.81$ using the notation of Brorby et al. (2014)¹. This results agree reasonably well with those of Mineo et al. (2012) ($q = 1.49$) and Brorby et al. (2014) ($q = 14.5$), and imply that low-metallicity galaxies produce ~ 10 times more HMXBs per unit SFR than their high-metallicity counterparts.

The dispersion given by our model describes the one of the low-metallicity subsample very well, as shown in Fig. 4. However, this is not true for the high-metallicity data. To investigate this issue, we added an extra dispersion σ_{extra} to the number of sources predicted by our model, and fit it to the high-metallicity subsample for different (fixed) values of σ_{extra} . A Bayesian model comparison shows that the data are best described by $\sigma_{\text{extra}} = 22$, typically 3 times that given by Poissonian effects, and a slightly lower $\log Q_{\text{high}} = 2.14$. The decrease in the latter value compared to the $\sigma_{\text{extra}} = 0$ model is due to the fact that the extra dispersion (constant and larger than the Poissonian uncertainties of the data) increases the weight of lower- N values relative to higher- N ones, hence lowering the best-fit value. The fact that an extra dispersion is not observed in the low-metallicity sample could be explained by the different sample sizes (556 high- vs. 12 low-metallicity sources). Indeed, the test of picking random samples of 12 high-metallicity sources and performing the same analysis shows that the extra dispersion remains undetected if the size of the high-metallicity sample is small. Moreover, the mean best-fit value of Q_{high} obtained for the random samples of 12 high-metallicity sources agrees to within uncertainties with the best-fit value of the whole high-metallicity sample. This fact implies that the increase in the number of HMXBs per unit SFR at low metallicities is not an effect of the difference in the sample sizes.

The aforementioned results confirm that the metallicity dependence observed in our sample is not an effect of statistical fluctuations, but a real physical trend, as the value of Q is different for the low- and high-metallicity subsamples even when the dispersion of the data is well described. Indeed, the probability that the low-metallicity data comes from random fluctuations around the best high-metallicity model (including the extra dispersion) is $\sim 10^{-47}$ times that given by the best low-metallicity model. Hence, in agreement with Brorby et al. (2014), our analysis suggests that the number of HMXBs per unit SFR increases by a factor of ~ 10 at low metallicities, even if we take into account the statistical fluctuations. As an interesting by-product,

¹ The factor Q in our work differs from that used by Brorby et al. (2014) in the normalization of the luminosity function.

these fluctuations are well described by Poissonian effects at low metallicities but are larger at high metallicities.

The models that best fit the number of sources of the galaxies also describe nicely their integrated luminosities (Fig. 3). In the high metallicity limit, the predictions are also in well agreement with the relation of Mineo et al. (2012). This result suggests that the most numerous HMXB populations produced at low metallicities can explain not only the different observed number of sources, but also the enhanced integrated X-ray luminosity of galaxies. Hence, the dependence of the intrinsic luminosity of the sources with metallicity would be small. However, given that the *observed* number of HMXBs depends on the XLF through the detectability of the sources, and that the XLF is fixed in our model, a dependence of the luminosity on metallicity cannot be completely discarded.

To test the consistency of the observed luminosities with those predicted by our model in a statistical way, we extended our model to predict the luminosities of the observed HMXBs by sampling the XLF with the restriction $L > L_{\text{th}}$ (see Appendix A for details). We fit the model to the high- and low-metallicity subsamples, now including the observed luminosities of the galaxies together with their numbers of HMXBs in the comparison. We included the previously obtained extra dispersion in the high-metallicity model. The resulting values of the free parameter Q are in full agreement with the ones obtained before, confirming that the luminosity dependence on metallicity is small, if any.

To further explore this possibility, we analyzed the correlation between the observed mean luminosity of HMXBs in each galaxy ($L_X^{G,38}/N^{38}$) and its metallicity (Fig. 5). This figure shows a very weak trend (red, dashed line) in the sense that HMXBs are brighter in low-metallicity populations, masked by large-amplitude fluctuations. If this trend could be confirmed by more numerous and more precise data, then it might indicate that the observed variation of the total X-ray luminosity of the HMXB populations results from a large contribution due to the change in the population size, plus a minor one originated in the variation of the mean HMXB luminosity.

Finally, it is important to point out that although the simplest model that describes the dependence of the number of HMXBs on the oxygen abundance is a split function, there is no physical evidence to justify the break at $12 + \log(\text{O}/\text{H}) \simeq 8$. Hence, following the apparent linear trend shown by the number of observed HMXBs in Fig. 4, as a second approach we also parameterized the factor Q by a linear function of the oxygen abundance. A fit to our whole sample confirms the existence of a dependence of Q on metallicity with a confidence greater than $1 - 5 \times 10^{-5}$. However, the best-fit parameters suggest that the anticorrelation between the number of binaries and the metallicity of the galaxies is milder than that predicted by a sharp break. In this case the change in the population size per unit SFR is only -0.14 dex per dex. The discrepancy is originated in the better statistics of high-metallicity data, which have more weight in the fit than low-metallicity ones, hence biasing the slope to shallower values. A deeper exploration of this issue would require larger samples, and a more detailed treatment of population size fluctuations.

4. Conclusions

We compiled from the literature a large sample of 49 galaxies with measured star-formation rates, oxygen abundances and HMXB population properties (observed number of sources and

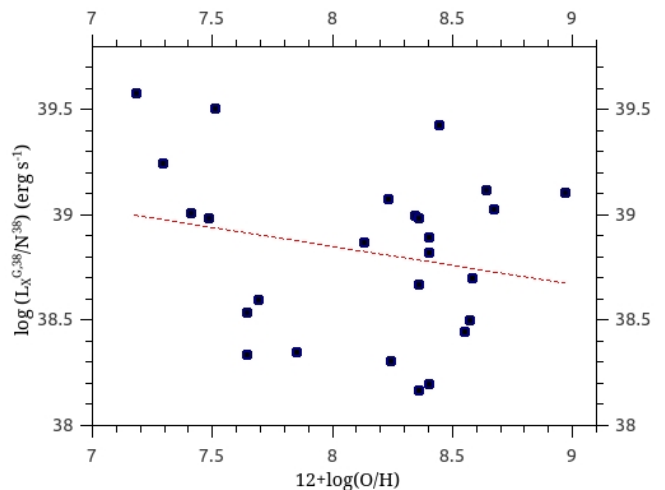


Fig. 5. Mean luminosity of HMXBs above a threshold of $10^{38} \text{ erg s}^{-1}$, as a function of the oxygen abundance. Black, filled squares are the data points, while the red dashed line is the best linear fit to them, $\log L_X^{G,38}/N^{38} = (38.7 \pm 1.3) - (0.18 \pm 0.16)[\log(\text{O}/\text{H}) - \log(\text{O}/\text{H})_{\odot}]$. A low-significance anticorrelation between these variables is observed, which might suggest a brightening of HMXBs at low metallicities. As in previous figures, the dispersion around the fit is large (0.4 dex).

luminosities above some threshold luminosity). This sample enhances those previously compiled by Mineo et al. (2012) and Brorby et al. (2014) by giving the oxygen abundances of the galaxies, and by including both the number of sources in each galaxy together with its integrated luminosity. As the low- and high-metallicity galaxies were taken from different works, with different uncertainties and selection effects, the data were homogenized to make statistical comparisons and draw meaningful conclusions. The SFRs in our sample were corrected to a common estimator. For the oxygen abundances, we used a single estimator in the largest possible subsample and the rest of the sample was shown to be unbiased with respect to this homogeneous set. Any residual non-uniformity that may have been left over by this process is small, and does not affect our conclusions. Hence, our complete sample is then suitable to investigate the effects of metallicity in the properties of HMXB populations.

Our main results are the following:

- The size of the HMXB populations of low-metallicity ($12 + \log(\text{O}/\text{H}) < 8$) galaxies, per unit SFR, is ~ 10 times larger than that of solar metallicity galaxies. This result is robust; it describes well the observed number of HMXBs and integrated X-ray luminosities of nearby star-forming galaxies, taking into account the observed fluctuations. It does not depend on the different sizes of the low- and high-metallicity subsamples. It is also in excellent agreement with those of Kaaret et al. (2011) and Brorby et al. (2014), who found similar results using slightly different methods and data.
- The dispersion in the observed number of HMXBs of low-metallicity, low-SFR galaxies can be explained by Poissonian fluctuations associated with the initial mass function and XLF sampling. This is not the case for high-metallicity galaxies, which show fluctuations typically half an order of magnitude higher than those expected from these effects. Mineo et al. (2012) have reported an analog dispersion in the L_X –SFR relation of the high-metallicity subsample, which cannot be attributed to selection effects. Our results indicate

that the large dispersion is not an effect of the IMF or XLF sampling. Therefore it could be due to a different physical origin, most probably the internal metallicity dispersion of the galaxies and the age distribution of HMXBs. Indeed, numerical simulations of stellar populations which compute the effect of these fluctuations give dispersion values that agree with our findings (Artale et al., in prep.). This effect is hidden in the low-metallicity subsample behind the large Poissonian fluctuations that arise from the small number of sources.

- Present evidence suggests that the aforementioned metallicity effects occur below $12 + \log(O/H) \approx 8$. Although a smooth linear trend in the correlations is apparent, it is not supported by statistical tests on the present sample. A sharp break at this metallicity provides the best description of our data. However, the smoothness of the metallicity dependence could be affected by differences in the sample size or residual inhomogeneities. A larger sample including galaxies near the break and more low-metallicity galaxies is crucial to solve this issue. A more thorough description of the data fluctuations, especially those from non-Poissonian sources, would also be desirable.
- The dependence of the luminosity of HMXBs with metallicity, if any, is much smaller than that of the population size. However, refined models are needed to determine its significance.

Our results support the general picture describing the enhancement of the HMXB population size at low metallicities (Dray 2005; Linden et al. 2010; Kaaret et al. 2011; Basu-Zych et al. 2013a,b; Fragos et al. 2013; Brorby et al. 2014). This implies that these energetic sources were more numerous in the Early Universe, hence a promising source of heating and partial reionization of the IGM, as proposed by several authors (Power et al. 2009; Mirabel et al. 2011; Jeon et al. 2013; Power et al. 2013; Kneivitt et al. 2014). It also implies that the energy feedback from HMXBs to the ISM could play an important role in the formation and evolution of low-mass galaxies, supporting recent results (Artale et al. 2015). To assess the exact magnitude of these effects, a more detailed description of the variation of the number of HMXBs and their luminosities must be given. This is only possible with further deep observations that unveil the low-luminosity end of the XLF in nearby galaxies, together with homogeneous metallicity and SFR estimations. Age measurements of the star-forming regions would also be valuable to investigate the effects of this important parameter.

Acknowledgements. We would like to thank the referee for useful comments that helped to clarify the manuscript. This work was partially supported by Argentine grants PICT 2011-0959 (ANPCyT), PIP 2009-0305 and PIP 2012-0396 (CONICET). Our MC models were run in the Fenix cluster of the Numerical Astrophysics Group at the Institute for Astronomy and Space Physics (CONICET/UBA).

References

Artale M., Tissera, P.B., Pellizza, L.J. 2015, MNRAS, 448, 3071.
 Asplund M., Grevesse N., Sauval A.J., Scott P. 2009, ARA&A, 47, 481.
 Basu-Zych, A., Lehmer, B., Hornschemeier, A., Bouwens, R., Fragos, T., Oesch, P., Belczynski, K., Brandt, W., Kalogera, V., Luo, B., Miller, N., Mullaney, J., Tzanavaris, P., Xue, Y., Zezas, A. 2013, ApJ, 762, 45.
 Basu-Zych, A., Lehmer, B., Hornschemeier, A., Gonçalves, T., Fragos, T., Heckman, T., Overzier, R., Ptak, A., Schiminovich, D. 2013, ApJ, 774, 152.
 Belczynski K., Sadowski A., Rasio F., 2004, ApJ, 11, 1068.
 Belczynski, K., Kalogera, V., Rasio, F.A., Taam, R.E., Zezas, A., Bulik, T., Maccarone, T.J., Ivanova, N. 2008, ApJS, 174, 223-260.
 Belczynski, K., Bulik, T., Fryer, C., Ruiters, A., Valsecchi, F., Vink, J., Hurley, J., 2010, ApJ, 714(2), pp. 1217-1226.

Bibby J., Crowther P., 2010, MNRAS, 405, 2737.
 Bresolin F., Garnett D., Kennicutt R., 2004, ApJ, 615, 228.
 Brorby, M., Kaaret, P., Prestwich, A., 2014, MNRAS, 441, 2346.
 Calzetti D., Kennicutt R., Engelbracht C. et al., 2007, ApJ, 666, 870.
 Cedres B., Urbanaja M., Cepa J., 2004, A&A, 422, 511.
 Denicoló G., Terlevich R., Terlevich E., 2002, MNRAS, 330, 69.
 Dray L. M., 2005, MNRAS, 370, 2079.
 Engelbracht C., Rieke G., Gordon K., Smith T. et al., 2008, ApJ, 678, 804.
 Esteban C., Bresolin F., Peimbert M., Garcia-Rojas J., Peimbert A., Mesa-Delgado A., 2009, ApJ, 700, 654.
 Fabbiano, G. 2006, Annu. Rev. Astron. Astrophys., 44, 323.
 Fragos, T., Lehmer, B., Naoz, S., Zezas, A., Basu-Zych, A. 2013, ApJ, 776, L31.
 Garnett D., Kennicutt R. Jr., Bresolin F., 2004, ApJ, 607, L21.
 Ghigo, F., Wardle, J., Cohen, N., 1983, AJ, 88, 1587.
 Gilfanov, M., 2004, MNRAS, 349, 146.
 Gilfanov, M., Grimm, H., Sunyaev, R., 2004, MNRAS, 351, 1365.
 Grimm, H., Gilfanov, M., Sunyaev, R. 2003, MNRAS, 339, 793.
 Hirashita H., Tajiri Y., Kamaya H., 2002, A&A, 388, 439.
 Hopkins A., Schulte-Ladbeck R., Drozdovsky I., 2002, AJ, 124, 862.
 Hunter, D., Elmegreen, B., Ludka, B., 2010, AJ, 139, 447.
 Izotov Y., Thuan T., 2011, ApJ, 734, 82.
 Izotov Y., Thuan T., 1999, ApJ, 511, 639.
 Jeon M., Pawlik A.H., Bromm V., Milosavljević M. 2013, MNRAS, 440, 3778
 Justham S., Schawinski K. 2012, MNRAS, 423, 1641
 Kaaret P., Schmitt J., Gorski M., 2011, ApJ, 741, 10.
 Kewley L., Ellison S., 2008, ApJ, 681, 1183.
 Kneivitt G., Wynn G., Power C., Bolton J. 2014, arXiv:1409.1267
 Justham S., Pustilnik A., Grebel E., Lee H., Pramskij A., 2004, ApJS, 153, 429.
 Kobulnicky H., Skillman E., 1997, ApJ, 489, 636.
 Liang Y., Yin S., Hammer F., Deng L., Flores H., Zhang B., 2006, ApJ, 652, 257.
 Linden, T., Kalogera, V., Sepinsky, J. F., Prestwich, A., Zezas, A., Gallagher, J. 2010, ApJ, 725, 1984.
 López-Sánchez A., Esteban C., 2010, A&A, 521, 31.
 López-Sánchez A., Dopita M., Kewley L., Zahid H., Nicholls D., Scharwachter J., 2012, MNRAS, ArXiv Astrophysics e-prints, arXiv:1203.5021.
 Mineo, S., Gilfanov, M., Sunyaev, R. 2012, MNRAS, 419, 2095.
 Mirabel I., Dottori H., Lutz D., 1992, A&A, 256, L19.
 Mirabel, I., Dijkstra, M., Laurent, P., Loeb, A., Pritchard, P. 2011, A&A, 528, A149.
 Pettini M., Pagel B., 2004, MNRAS, 348, 59.
 Pilyugin L., Thuan T., 2007, ApJ 669 299
 Power C., Wynn G.A., Combet C., Wilkinson M.I. 2009, MNRAS, 295, 1146
 Power C., James G., Combet C., Wynn G.A. 2013, ApJ, 764, 76
 Prestwich, A., Tsantaki, M., Zezas, A., Jackson, F., Roberts, T., Foltz, R., Linden, T., Kalogera, V., 2013, ApJ, 769, 92.
 Sage L. J., Salzer J. J., Loose H.-H., Henkel C., 1992, A&A, 265, 19.
 Storchi-Bergman T., Wilson A., Baldwin J., 1996, ApJ, 460, 252
 Shtykovskiy, P., Gilfanov, M., 2005, A&A, 431, p.597-614.
 Shtykovskiy P., Gilfanov M., 2007, AstL, 33, 437.
 Thuan T., Bauer F., Papaderos P., Izotov Y., 2004, ApJ, 606, 213.
 Werk J., Putman M., Meurer G., Santiago-Figueroa N., 2011, ApJ, 735, 71.
 Winter L., Mushotzky R., Reynolds C., 2006, ApJ, 649, 730.
 Yin S., Liang Y., Hammer F., Brinchmann J., Zhang B., Deng L., Flores H., 2007a, A&A, 462, 535.
 Yin S., Liang Y., Zhang B., 2007b, ASPC, 373, 686.

Appendix A: Monte Carlo model for HMXB populations

Appendix A.1: The model

Our model aims at describing the observed number of sources and total luminosity of the population of HMXBs of a star-forming galaxy with a given SFR and oxygen abundance, to perform a robust comparison with observations. We start from the fact that the progenitors of HMXBs are massive stars, and that these sources have short evolutionary times (less than ~ 100 Myr; Belczynski et al. 2004, 2008, 2010; Shtykovskiy & Gilfanov 2005, 2007). Hence, their number should correlate with the SFR, as Brorby et al. (2014) propose in their Eqn. 4. However, star formation can be regarded as a stochastic process in which the final result (the formation of a single or binary star, and in the last case the initial masses of the components, the

semimajor axes and eccentricities, etc.) follows a certain probability distribution. The sampling of this distribution in a particular stellar population is a source of fluctuations in the number of HMXBs (N_{HMXB}) of a galaxy. If we assume that the probability p_{HMXB} of producing a HMXB is fixed, then N_{HMXB} has a Bernoulli distribution. As the total number of stars N_{\star} in any stellar population is large, and $p_{\text{HMXB}} \ll 1$ because the initial mass function is a strongly decreasing function, we can approximate N_{HMXB} as a Poisson variable. The mean value of N_{HMXB} is then $\langle N_{\text{HMXB}} \rangle = N_{\star} p_{\text{HMXB}}$. However, as it is not our aim to compute the distribution of N_{HMXB} from first principles, following Brorby et al. (2014) we parameterize it as

$$\langle N_{\text{HMXB}} \rangle = Q(\log(O/H)) \text{SFR}, \quad (\text{A.1})$$

as N_{\star} scales with SFR. The factor Q is the same as q in the Eqn. 4 of Brorby et al. (2014), apart from the normalization factor of the XLF. We note that the difference between our model and that of these authors is that the above equation describes the mean number of sources in a Poisson process, while the corresponding equation of Brorby et al. (2014) describes the exact number. When $N \sim 1$ (as in low-SFR galaxies) differences arise, and the stochasticity must be taken into account. Our factor Q may depend continuously on the oxygen abundance, or just assumed to be different for low and high metallicities.

A second source of stochasticity is given by the detection process, which selects only those binaries with luminosities above some threshold L_{th} . The XLF sampling in the HMXB population introduces statistical fluctuations in the number of sources detected. Given that detection is a Bernoulli process, the composition with the Poisson process described above produces a new Poisson process with mean

$$\langle N_{\text{det}} \rangle = p_{\text{det}} Q(\log(O/H)) \text{SFR}, \quad (\text{A.2})$$

where

$$p_{\text{det}} = k_{\text{norm}} \int_{L_{\text{th}}}^{L_{\text{max}}} \psi(L) dL. \quad (\text{A.3})$$

Here the XLF is $\psi(L) = k_{\text{norm}}^{-1} L^{\alpha}$ for $L \in [L_{\text{min}}, L_{\text{max}}]$, and $k_{\text{norm}} = \int_{L_{\text{min}}}^{L_{\text{max}}} L^{\alpha} dL$ is its normalization factor.

Our model allows, for a fixed set of free parameters, the computation of the number of sources detected in a galaxy N_{det} , by sampling a Poissonian with mean given by Eqn. A.2. The luminosities of these sources are computed numerically by sampling the XLF N_{det} times, with the restriction that $L > L_{\text{th}}$. The sum of the sampled luminosities is L_{det} , the total luminosity of the sources detected in the galaxy. In this way, we can compute a realization of the observables $(N_{\text{det}}, L_{\text{det}})$ for any galaxy with known SFR and $12 + \log(O/H)$. We stress that our model is stochastic, i.e., each realization of our model gives a different set $(N_{\text{det}}, L_{\text{det}})$. Therefore, the comparison with observations must be made with care. We also stress that it is possible for our model to predict that no sources will be detectable in a given galaxy ($N_{\text{det}} = L_{\text{det}} = 0$), hence upper limits in the observed luminosity of galaxies can be included in the comparison.

Appendix A.2: Model fitting

The stochasticity of our model prevents us to use classical statistical methods such as least-squares fitting to compare the model

to observations and fix the values of the free parameters; instead, we use a Bayesian approach. We note that the free parameter of the model may be Q itself, or any set introduced to parameterize Q as a function of metallicity (e.g., the slope and intercept of a linear function). We will refer to them collectively as θ . From Bayes' theorem, the posterior probability density f_{post} of the model parameters θ , given the observed data $(\{N, L_{\text{X}}^{\text{G}}\}_i$ with $i \in [1, 49]$) is

$$f_{\text{post}}(\theta | \{N, L_{\text{X}}^{\text{G}}\}) = \frac{\mathcal{L}(\{N, L_{\text{X}}^{\text{G}}\} | \theta) f_{\text{pri}}(\theta)}{\int \mathcal{L}(\{N, L_{\text{X}}^{\text{G}}\} | \theta) f_{\text{pri}}(\theta) d\theta}, \quad (\text{A.4})$$

where f_{pri} is the prior probability density of the parameters, and \mathcal{L} the likelihood of the data given the values of the parameters. The best-fit parameters are those that maximize the posterior probability density f_{post} .

Assuming that we have no previous information on the parameters, we adopt a uniform prior. On the other hand, the likelihood is the product of the individual joint probability distributions of the number and luminosities of HMXBs of each galaxy,

$$\mathcal{L}(\{N, L_{\text{X}}^{\text{G}}\} | \theta) = \prod_{i=1}^{49} f_L(L_{\text{X},i}^{\text{G}} | N_i, \theta) P_N(N_i | \theta), \quad (\text{A.5})$$

where f_L is the conditional probability density for the observed luminosity given the observed number of sources, P_N the probability of the observed number of sources, and the product arises because the data of different galaxies are independent. As described in the previous section, the best fit gives in this case the values

$$P_N(N_i | \theta) = \frac{e^{-\langle N_{\text{det}} \rangle} \langle N_{\text{det}} \rangle^{N_i}}{N_i!}. \quad (\text{A.6})$$

Assuming that the measurement uncertainty for the luminosity $L_{\text{X},i}^{\text{G}}$ of galaxy i is Gaussian with standard deviation $\sigma_{L,i}$,

$$f_L(L_{\text{X},i}^{\text{G}} | N_i) = \frac{1}{\sqrt{2\pi}\sigma_{L,i}} \int_0^{\infty} \exp\left(-\frac{(L_{\text{X},i}^{\text{G}} - L_{\text{det}})^2}{2\sigma_{L,i}^2}\right) g_L(L_{\text{det}} | N_i) dL_{\text{det}}. \quad (\text{A.7})$$

Here $g_L(L_{\text{det}} | N_i)$ is the probability density of the luminosity of the detected HMXBs. This factor has no analytical expression, but the integral in Eqn. A.7 can be computed numerically by a Monte Carlo method as the mean of the Gaussian in the integrand over a large number $M \gg 1$ of realizations of our model, i.e.,

$$f_L(L_{\text{X},i}^{\text{G}} | N_i) = \frac{1}{\sqrt{2\pi}\sigma_{L,i} M} \sum_{j=1}^M \exp\left(-\frac{(L_{\text{X},i}^{\text{G}} - L_{\text{obs},j})^2}{2\sigma_{L,i}^2}\right). \quad (\text{A.8})$$

In this way, the likelihood \mathcal{L} and hence the posterior probability density f_{post} can be computed within our model. The maximization of the latter was performed by a Markov-chain Monte Carlo scheme, giving the best-fit parameters and their uncertainties.