Mon. Not. R. Astron. Soc. **351,** 1365–1378 (2004) doi:10.1111/j.1365-2966.2004.07874.x

Statistical properties of the combined emission of a population of discrete sources: astrophysical implications

M. Gilfanov,[1,2]★ H.-J. Grimm[1] and R. Sunyaev[1,2]

[1]*Max-Planck-Institut für Astrophysik, 85741 Garching bei München, Germany*
[2]*Space Research Institute, Moscow, Russia*

Accepted 2004 March 23. Received 2004 March 8; in original form 2003 December 23

ABSTRACT

We study the statistical properties of the combined emission of a population of discrete sources (for example, the X-ray emission of a galaxy due to its population of X-ray binaries). Namely, we consider the dependence of their total luminosity $L_{\rm tot} = \sum L_k$ and of the fractional $\rm rms_{tot}$ of their variability on the number of sources n or, equivalently, on the normalization of the luminosity function. We show that, as a result of small number statistics, a regime exists in which $L_{\rm tot}$ grows non-linearly with n, in apparent contradiction with the seemingly obvious prediction $\langle L_{\rm tot}\rangle = \int L\,({\rm d}N/{\rm d}L)\,{\rm d}L \propto n$. In this non-linear regime, $\rm rms_{tot}$ decreases with n significantly more slowly than expected from the rms $\propto 1/\sqrt{n}$ averaging law. For example, for a power-law luminosity function with a slope of $\alpha = 3/2$, in the non-linear regime, $L_{\rm tot} \propto n^2$ and $\rm rms_{tot}$ does not depend at all on the number of sources n. Only in the limit of $n \to \infty$ do these quantities behave as intuitively expected, $L_{\rm tot} \propto n$ and $\rm rms_{tot} \propto 1/\sqrt{n}$. We give exact solutions and derive convenient analytical approximations for $L_{\rm tot}$ and $\rm rms_{tot}$.

Using the total X-ray luminosity of a galaxy due to its X-ray binary population as an example, we show that the $L_{\rm X}$–star formation rate and $L_{\rm X}$–M_* relations predicted from the respective 'universal' luminosity functions of high- and low-mass X-ray binaries are in good agreement with observations. Although caused by small number statistics, the non-linear regime in these examples extends as far as SFR $\lesssim$ 4–5 $\rm M_\odot\ yr^{-1}$ and $\log(M_*/\rm M_\odot) \lesssim$ 10.0–10.5, respectively.

Key words: methods: data analysis – methods: statistical – X-rays: binaries – X-rays: galaxies.

1 INTRODUCTION AND QUALITATIVE CONSIDERATIONS

In many astrophysical situations a problem arises to predict or interpret results of measurements of the total (combined) luminosity of a population of discrete sources. Among many examples are the total luminosity of X-ray binaries in a galaxy, or the total flux of background sources detected above the sensitivity limit inside the field of view of a telescope.

In the following discussion we will use high-mass X-ray binaries (HMXB) in star-forming galaxies as an example. As was shown by Grimm, Gilfanov & Sunyaev (2003), the luminosity distribution of HMXB sources in a galaxy is described to first approximation by a 'universal' luminosity function whose shape is the same in all galaxies and whose normalization is proportional to the star formation rate (SFR) of the parent galaxy:

$$\frac{{\rm d}N}{{\rm d}L} \propto {\rm SFR}\, f(L). \qquad (1)$$

In a broad luminosity range, $\log(L_{\rm X}) \sim$ 35.5–40.5, the shape of the HMXB 'universal' luminosity function is close to a power law, $f(L) = L^{-\alpha}$, with the slope $\alpha \approx 1.6$. Importantly, the luminosity of the compact sources in star-forming galaxies appears to be restricted by a maximum value of $L_{\rm cut} \sim$ few $\times\ 10^{40}$ erg $\rm s^{-1}$. This cut-off luminosity can be defined, for example, by the Eddington luminosity limit for the most massive objects associated with the star-forming regions. Obviously, on the faint side, the luminosity distribution equation (1) must become flatter or have a cut-off as well, in order to keep the total number of sources finite. This low-luminosity cut-off may be caused for example by the propeller effect (Illarionov & Sunyaev 1975) as discussed by Shtykovskii & Gilfanov (2004).

The expectation value for the total number of sources in a galaxy is given by

$$\langle N_{\rm tot}\rangle = \int_0^{+\infty} \frac{{\rm d}N}{{\rm d}L}\,{\rm d}L \propto {\rm SFR}, \qquad (2)$$

and, naturally, is directly proportional to its star formation rate. The number of sources actually observed in a given galaxy obeys the Poisson distribution $p_\mu(N)$ with μ defined by equation (2). Apart from effects of counting statistics, the number of HMXB sources

★E-mail: gilfanov@mpa-garching.mpg.de

found in an arbitrarily chosen galaxy will be close to the above expectation value.

The problem considered in this paper is the behaviour of the total luminosity of high-mass X-ray binaries in a galaxy,

$$L_{\rm tot} = \sum_{k=1}^{N} L_k, \tag{3}$$

as a function of its star formation rate.

An apparently obvious expression for $L_{\rm tot}$ can be obtained by integrating the luminosity distribution (1):

$$\langle L_{\rm tot}\rangle = \int_0^{+\infty} L\,\frac{{\rm d}N}{{\rm d}L}\,{\rm d}L \propto {\rm SFR}. \tag{4}$$

Hence, one might expect that the total X-ray luminosity of HMXB sources is also directly proportional to the star formation rate of the host galaxy, as is the total number of sources. This problem, however, involves some subtleties related to the statistical properties of the power-law distribution of the sources over luminosity, which appear not to have been recognized previously, at least in an astrophysical context [a somewhat related problem has been considered by Kalogera et al. (2001) in connection with estimating the coalescence rate for neutron star–neutron star binaries in the Galaxy]. Although equation (4) correctly predicts *the average X-ray luminosity computed for a large number of galaxies* with similar values of star formation rate, it fails to describe the relation between *the most probable value of X-ray luminosity of an arbitrarily chosen galaxy* and its SFR. The main surprise of the study presented here is that, in the low SFR regime, the relation between SFR of the host galaxy and the total luminosity of its HMXBs is non-linear – with increase of the star formation rate, the luminosity appears to grow faster than linear. The relation becomes linear only for sufficiently high star formation rates, when the total number of objects with a luminosity close to the maximum possible value, defined by $L_{\rm cut}$, becomes sufficiently large.

This can be illustrated by the following simple consideration. For an arbitrarily chosen galaxy, the brightest source is most likely going to have a luminosity $\tilde{L}_{\rm max}$ defined by the condition[1]

$$N(L > \tilde{L}_{\rm max}) = \int_{\tilde{L}_{\rm max}}^{+\infty} \frac{{\rm d}N}{{\rm d}L}\,{\rm d}L \sim 1. \tag{5}$$

For a power-law luminosity distribution with slope α and with a cut-off at $L_{\rm cut}$, equation (1), the above expression yields

$$\begin{aligned} &\tilde{L}_{\rm max} \propto {\rm SFR}^{1/(\alpha-1)} && \text{for low SFR,}\\ &\tilde{L}_{\rm max} = L_{\rm cut} && \text{for high SFR.} \end{aligned} \tag{6}$$

As might be intuitively expected, at low SFR, the most probable luminosity of the brightest source increases with SFR, until it reaches the maximum value of $L_{\rm cut}$. The threshold value of the star formation rate, separating low- and high-SFR regimes, is defined by the condition $N(L \sim L_{\rm cut}) \sim 1$, i.e. that there are about a few sources expected with luminosity close to the cut-off value $L_{\rm cut}$.

The most probable value of the total luminosity, $\tilde{L}_{\rm tot}$, can then be computed by integrating the luminosity function from $L_{\rm min}$ to $\tilde{L}_{\rm max}$:

$$\tilde{L}_{\rm tot} \approx \int_{L_{\rm min}}^{\tilde{L}_{\rm max}} \frac{{\rm d}N}{{\rm d}L}\,L\,{\rm d}L, \tag{7}$$

which, for $1 < \alpha < 2$ and $L_{\rm min} \ll \tilde{L}_{\rm max}$, leads to

$$\tilde{L}_{\rm tot} \propto \begin{cases} {\rm SFR}^{1/(\alpha-1)} & \text{for low SFR,}\\ {\rm SFR} & \text{for high SFR,} \end{cases} \tag{8}$$

i.e. is non-linear in the low-SFR regime and becomes linear only at high star formation rates.

This can be qualitatively understood as follows. For a slope of the luminosity distribution $1 < \alpha < 2$, the total luminosity of a galaxy is defined by the brightest sources. The non-linear behaviour in the low-SFR limit is caused by the fact that an increase of the SFR leads to a non-linear increase of the luminosity of the brightest sources. Therefore their total luminosity grows faster than the star formation rate. This non-linear growth continues until the maximum possible value of the luminosity of the compact sources is achieved. Further increase of the star formation rate leads to a linear increase of the number of the brightest sources in the galaxy, but not of their individual luminosities. Consequently, the $L_{\rm X}$–SFR relation becomes linear.

In the more formal language of statistics, such behaviour is related to the properties of the probability distribution of the collective luminosity $p(L_{\rm tot})$. In particular, it can be understood in terms of the difference between the expectation mean and the mode of the probability distribution. The expectation mean is defined as

$$\langle L_{\rm tot}\rangle = \int_0^{\infty} p(L_{\rm tot}) L_{\rm tot}\,{\rm d}L_{\rm tot}, \tag{9}$$

and is given by equation (4). The mode of the statistical distribution, $\tilde{L}_{\rm tot}$, is defined as the value of the random variable ($L_{\rm tot}$ in our case) at which the probability distribution $p(L_{\rm tot})$ reaches its maximum value. Whereas the expectation mean $\langle L_{\rm tot}\rangle$ describes the result of averaging of the X-ray luminosities of many galaxies having similar values of SFR, it is the mode of the $p(L_{\rm tot})$ distribution that predicts the most probable value of the total luminosity of a randomly chosen galaxy. The non-linear behaviour in the low-SFR regime is caused by the skewness of the probability distribution $p(L_{\rm tot})$ resulting in a difference between expectation mean and mode. In the high-SFR limit, the $p(L_{\rm tot})$ distribution asymptotically approaches the Gaussian distribution, in accord with the central limit theorem. The boundary value of the SFR, separating the non-linear and linear regimes of the $L_{\rm X}$–SFR relation is defined by the parameters of the luminosity function.

Interestingly, the fact of the existence of a linear regime in the $L_{\rm X}$–SFR relation is a direct consequence of the cut-off in the luminosity function. Only in the presence of a maximum possible luminosity for the sources, $L_{\rm cut}$ (for instance, Eddington limit for a neutron star), can a linear regime exist, when the total luminosity of a galaxy is defined by a sufficiently large number of bright sources near $L_{\rm cut}$, and any subsequent increase of the star formation rate results in linear growth of the total luminosity.

In the above discussion, we used high-mass X-ray binary populations in star-forming galaxies as an example. Obviously, the effect considered in this paper is of a broader general interest and is at work in many situations related to computing/measuring the integrated luminosity of a finite number of discrete sources. One example related to the application of stellar synthesis models to observations of stellar clusters has been independently discovered by Cervino & Luridana (2004).

This paper is structured as follows. In Section 2 we consider the statistical properties of the total luminosity, in particular in Section 2.1 we derive formulae for the probability distribution $p(L_{\rm tot})$, using two different approaches, and present results of numerical calculations in Section 2.2. The variability of the total emission is

[1] Indeed, for example, sources 10 times more luminous will appear on average in one out of $\sim 10^{(\alpha-1)}$ galaxies.

studied in Section 3. In Section 4 we discuss astrophysical applications, including the properties of the total X-ray emission of a galaxy due to its population of high-mass (Sections 4.2 and 4.3) and low-mass (Section 4.4) X-ray binaries. Our results are summarized in Section 5. In the Appendices we derive convenient approximations for $\tilde{L}_{\rm tot}$ and fractional rms of the total emission and consider their asymptotic behaviour. A casual reader, not interested in the mathematical aspects of the problem, can skip Section 2.1 and proceed with Section 2.2.

2 TOTAL LUMINOSITY

We consider a population of sources with a luminosity function (LF):

$$\frac{{\rm d}N}{{\rm d}L} = Af(L). \tag{10}$$

The expectation values for the total number of sources and the total luminosity are

$$\langle n\rangle = \int_0^\infty \frac{{\rm d}N}{{\rm d}L}\,{\rm d}L, \tag{11}$$

$$\langle L_{\rm tot}\rangle = \int_0^\infty L\,\frac{{\rm d}N}{{\rm d}L}\,{\rm d}L. \tag{12}$$

It is assumed that both quantities are well defined and finite.

2.1 The probability distribution of the total luminosity

Below we present two methods of computing the probability distribution of the total luminosity. The second of these two methods is somewhat more convenient computationally.

2.1.1 Method I

To compute the probability distribution for the total luminosity, $p(L_{\rm tot})$, we divide the L_1–L_2 luminosity range into intervals of infinitesimal width δL_k and express the combined luminosity of all sources as a sum

$$L_{\rm tot} = \sum_k L_{{\rm tot},k}, \tag{13}$$

where $L_{{\rm tot},k}$ is the combined luminosity of the sources in the kth interval, running from L_k to $L_k + \delta L_k$. The number of sources in the interval $(L_k, L_k + \delta L_k)$ obeys a Poisson distribution with mean

$$\mu_k = \frac{{\rm d}N}{{\rm d}L}(L = L_k)\delta L_k. \tag{14}$$

For $\delta L_k \to 0$, it is sufficient to consider the occurrence of either zero or one source per interval whose probabilities are, respectively,

$$\begin{aligned} p_k(0) &= 1 - \frac{{\rm d}N}{{\rm d}L}\delta L_k + O\left(\delta L_k^2\right),\\ p_k(1) &= \frac{{\rm d}N}{{\rm d}L}\delta L_k + O\left(\delta L_k^2\right). \end{aligned} \tag{15}$$

The probability distribution for the combined luminosity of the sources in the kth interval is

$$p(L_{{\rm tot},k}) = \delta(L_{{\rm tot},k})\left(1 - \frac{{\rm d}N}{{\rm d}L}\delta L_k\right) + \delta(L_{{\rm tot},k} - L_k)\frac{{\rm d}N}{{\rm d}L}\,\delta L_k, \tag{16}$$

where $\delta(x - x_0)$ is the delta function. The characteristic function of this distribution is

$$\begin{aligned} \hat{p}(L_{{\rm tot},k}) &= \int p(L_{{\rm tot},k})\,{\rm e}^{{\rm i}\omega L}\,{\rm d}L\\ &= 1 - \frac{{\rm d}N}{{\rm d}L}\delta L_k + \frac{{\rm d}N}{{\rm d}L}\delta L_k\,{\rm e}^{{\rm i}\omega L_k}. \end{aligned} \tag{17}$$

Using the convolution theorem, the characteristic function of the probability distribution of the total luminosity can be computed as a product of characteristic functions of $p(L_{{\rm tot},k})$:

$$\hat{p}(L_{\rm tot}) = \prod_k \hat{p}(L_{{\rm tot},k}), \tag{18}$$

$$\begin{aligned} \ln\hat{p}(L_{\rm tot}) &= \sum_k \ln\left(1 - \frac{{\rm d}N}{{\rm d}L}\delta L_k + \frac{{\rm d}N}{{\rm d}L}\delta L_k\,{\rm e}^{{\rm i}\omega L_k}\right)\\ &= -\sum_k \frac{{\rm d}N}{{\rm d}L}\delta L_k + \sum_k \frac{{\rm d}N}{{\rm d}L}\delta L_k\,{\rm e}^{{\rm i}\omega L_k} + O\left(\delta L_k^2\right)\\ &= -\int \frac{{\rm d}N}{{\rm d}L}\,{\rm d}L + \int \frac{{\rm d}N}{{\rm d}L}\,{\rm e}^{{\rm i}\omega L}\,{\rm d}L. \end{aligned} \tag{19}$$

Finally,

$$\begin{aligned} p(L_{\rm tot}) &= \int \hat{p}(L_{\rm tot})\,{\rm e}^{-{\rm i}\omega L_{\rm tot}}\,{\rm d}\omega\\ &= \int \exp\left(\int \frac{{\rm d}N}{{\rm d}L}\,{\rm e}^{{\rm i}\omega L}\,{\rm d}L - \langle n\rangle - {\rm i}\omega L_{\rm tot}\right)\,{\rm d}\omega, \end{aligned} \tag{20}$$

where $\langle n\rangle$ is given by equation (11).

2.1.2 Method II

The probability distribution of the total number of sources follows a Poisson distribution $P_\mu(n) = \mu^n\,{\rm e}^{-\mu}/n!$ with mean μ given by equation (11). The probability distribution of the total luminosity is

$$p(L_{\rm tot}) = \sum_{n=0}^{n=\infty} P_\mu(n)p_n(L_{\rm tot}), \tag{21}$$

where $p_n(L_{\rm tot})$ is the probability distribution of the total luminosity of exactly n sources. In the majority of practically interesting cases, the total number of sources is sufficiently large, $\langle n\rangle \gg 1$, and the Poisson distribution in equation (21) can be replaced by the delta function $\delta(n - \langle n\rangle)$:

$$p(L_{\rm tot}) \approx p_n(L_{\rm tot}), \qquad n = \langle n\rangle \gg 1, \tag{22}$$

where $n = \langle n\rangle$ is given by equation (11).

The probability distribution of the sum of n random variables, $L_{\rm tot} = \sum_{k=1}^n L_k$, can be calculated using the convolution theorem:

$$\begin{aligned} \widehat{p_n} &= \widehat{p_1}^{\,n},\\ \widehat{p_1}(\omega) &= \int_0^\infty p_1(L)\,{\rm e}^{{\rm i}\omega L}\,{\rm d}L,\\ p_n(L_{\rm tot}) &= \int_{-\infty}^{+\infty} \widehat{p_n}(\omega)\,{\rm e}^{-{\rm i}\omega L_{\rm tot}}\,{\rm d}\omega, \end{aligned} \tag{23}$$

where $p_1(L)$ is the probability distribution for the luminosity of one source, which equals the luminosity function with appropriate normalization:

$$p_1(L) = \frac{1}{\langle n\rangle}\,\frac{{\rm d}N}{{\rm d}L}. \tag{24}$$

The probability distribution $p_n(L_{\rm tot})$, defined by equation (23), describes the case when n sources are *observed*. In contrast, the

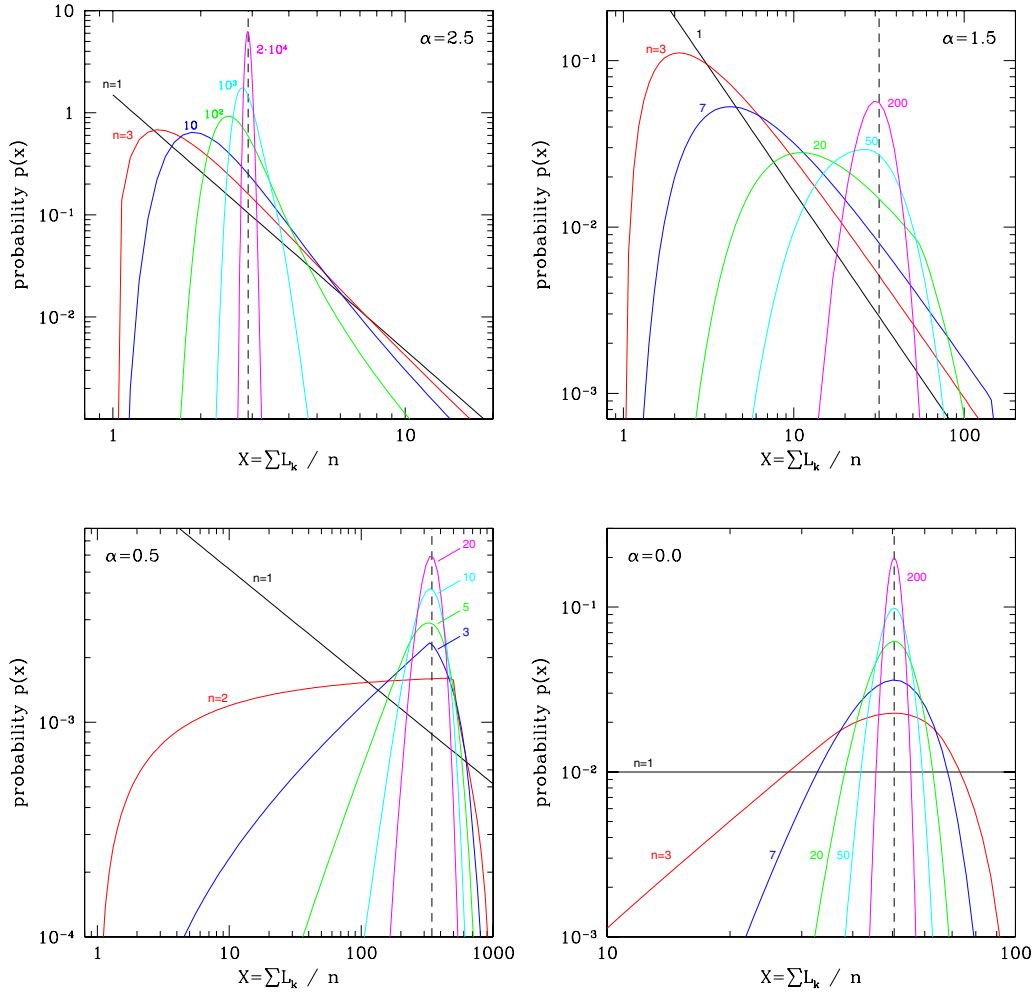

Figure 1. Probability distributions of the average luminosity, $\sum_{k=1}^{n} L_k/n$, of n discrete sources with a power-law LF (equation 25). The lower and upper luminosity cut-offs were fixed at $L_1 = 1$ and $L_2 = 10^3$ for all plots. The value of the slope α is indicated in each panel. Each curve is marked according to the number of sources n. The vertical dashed lines show the expectation mean $\langle L_{\rm tot}\rangle/n$.

distribution $p(L_{\rm tot})$, defined by equation (20), is parametrized via the normalization of the luminosity function or, equivalently, via the expectation value $\langle n\rangle$, and describes the case when $\langle n\rangle$ sources are *expected*. These two distributions are related via equation (21). For $n \gg 1$, which is often the case, they are nearly identical, and it can be assumed that $n = \langle n\rangle$ and both quantities are related to the normalization of the luminosity function via equation (11).

2.2 Illustrative examples

To illustrate the behaviour of the total luminosity, we assume a power-law luminosity function:

$$\frac{{\rm d}N}{{\rm d}L} = \begin{cases} AL^{-\alpha} & \text{if } L_1 \leqslant L \leqslant L_2, \\ 0 & \text{otherwise.} \end{cases} \qquad (25)$$

2.2.1 *Probability distribution $p_n(L_{\rm tot})$*

The probability distributions $p_n(L_{\rm tot})$ for various values of the slope α and the number of sources n are shown in Fig. 1. To facilitate the comparison of distributions for different values of n, the abscissa in these plots is the average luminosity $L_{\rm tot}/n = \sum_{k=1}^{n} L_k/n$. The values of the luminosity cut-offs were fixed at $L_1 = 1$ and $L_2 = 10^3$.

The skewness of the probability distribution $p_n(L_{\rm tot})$ leads to a deviation of its mode $\tilde{L}_{\rm tot}$ (the most probable value of the total luminosity) from the expectation mean $\langle L_{\rm tot}\rangle$ indicated in each panel by the vertical dashed line. The effect is most pronounced for $\alpha \sim 3/2$, is unimportant for shallow luminosity functions with $\alpha < 1$ and gradually diminishes with increasing α at $\alpha > 2$.

For illustration, we also show in Fig. 1 the probability distributions for a flat LF, $\alpha = 0$, in which case $\tilde{L}_{\rm tot}$ exactly equals $\langle L_{\rm tot}\rangle$. Naturally, for any value of α the $p_n(L_{\rm tot}) \rightarrow$ Gaussian distribution in the limit of $n \rightarrow \infty$, in accord with the central limit theorem. Correspondingly $\tilde{L}_{\rm tot} \rightarrow \langle L_{\rm tot}\rangle$ in this limit.

2.2.2 *The most probable value of total luminosity*

The dependence of the most probable value of the total luminosity $\tilde{L}_{\rm tot}$ on the total number of sources n for different values of α and the ratio L_2/L_1 is shown in Figs 2 and 3. Convenient analytical approximations for the $\tilde{L}_{\rm tot}$–n relation are derived in Appendix A and its asymptotic behaviour is considered in Appendix A2.

The $\tilde{L}_{\rm tot}$–n relation is significantly non-linear for a 'small' number of sources or, equivalently, for small values of the LF normalization A. For $L_1 \ll L_2$, $\alpha > 1$, the boundary between the non-linear and linear regimes, expressed in terms of the normalization A,

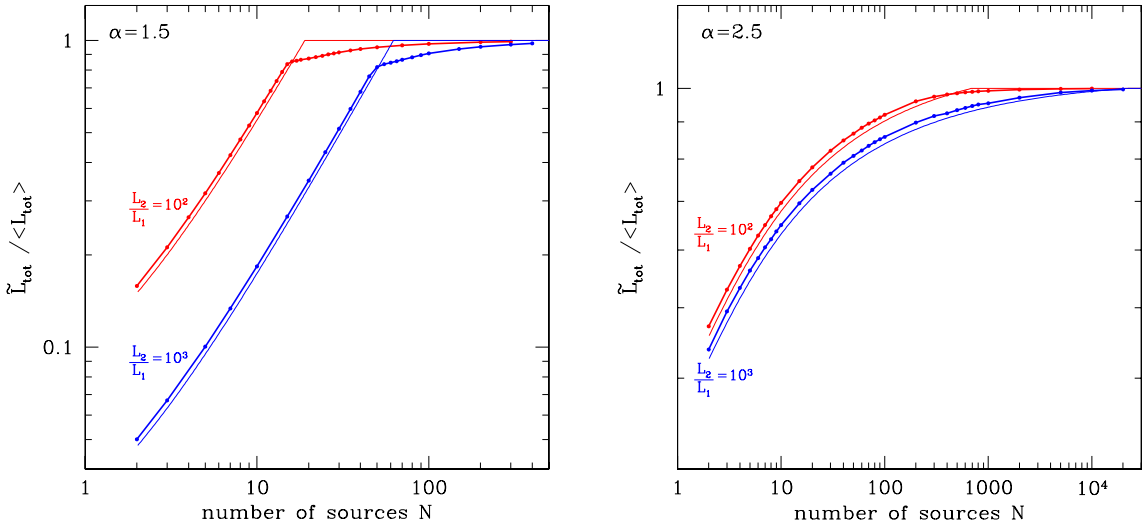

Figure 2. The ratio of the most probable value of the total luminosity $\tilde{L}_{\rm tot}$ to the expectation mean $\langle L_{\rm tot}\rangle$ versus number of sources observed in the L_1–L_2 luminosity range for different values of the LF slope α and the ratio L_2/L_1. The results of exact calculation using equation (23) are shown by the solid symbols connected with the solid line. The solid line without symbols shows the approximate relation calculated from equations (A7) and (A3), as described in Appendix A1.

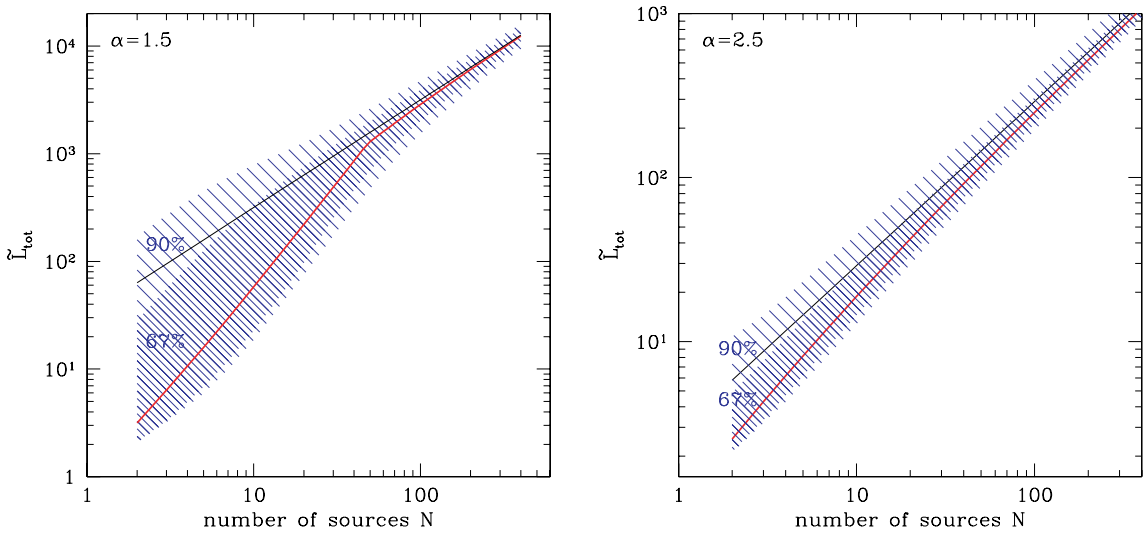

Figure 3. The most probable value of the total luminosity $\tilde{L}_{\rm tot}$ and its intrinsic dispersion versus number of observed sources for different values of the LF slope α. The lower and upper cut-off luminosities were fixed at $L_1 = 1$ and $L_2 = 10^3$. The results of exact calculation using equation (23) are shown. The behaviour of $\tilde{L}_{\rm tot}$ is shown by the thicker solid line. The narrower and broader dashed areas correspond to its 67 and 90 per cent intrinsic dispersion. The thinner solid line shows the linear relation for the expectation mean $\langle L_{\rm tot}\rangle \propto n$.

depends only on the LF slope α and its high-luminosity cut-off L_2 (see Appendix A2). This is due to the fact that the behaviour of $\tilde{L}_{\rm tot}$ is defined by the number of sources near the high-luminosity cut-off of the luminosity function, rather than by the total number of sources in the entire L_1–L_2 luminosity range, i.e. the condition for the linear regime is

$$N(L\sim L_2)\sim L_2\frac{{\rm d}N}{{\rm d}L}(L=L_2)\gtrsim 1. \qquad (26)$$

The total number of sources in the L_1–L_2 luminosity range, on the contrary, is defined by the low-luminosity cut-off L_1 (for $\alpha > 1$). For sufficiently large L_2/L_1, the non-linear regime can occur for an arbitrarily large total number of sources (cf. the curves corresponding to different values of L_2/L_1 in Fig. 2). Interestingly, for the LF slope in the range of $1 < \alpha < 2$, where the effect is strongest, there is a relatively sharp break separating the non-linear part of the dependence from the linear part.

Although for small n the most probable value of luminosity $\tilde{L}_{\rm tot}$ can deviate significantly from the expectation mean $\langle L_{\rm tot}\rangle$ (Figs 1 and 2), the condition $\int p(L_{\rm tot})L_{\rm tot}\,{\rm d}L_{\rm tot} = \langle L_{\rm tot}\rangle$ is satisfied for any n. Consequently, the average of $L_{\rm tot}$ over a large number of realizations with the same n always equals $\langle L_{\rm tot}\rangle$. This equality is achieved due to the existence of outliers, having values of $L_{\rm tot}$ significantly exceeding both $\tilde{L}_{\rm tot}$ and $\langle L_{\rm tot}\rangle$, in accordance with the skewness of the probability distribution $p_n(L_{\rm tot})$ for small n. This naturally leads to enhanced and asymmetric dispersion of the observed values of $L_{\rm tot}$ in the non-linear regime as illustrated by the shaded areas in Fig. 3.

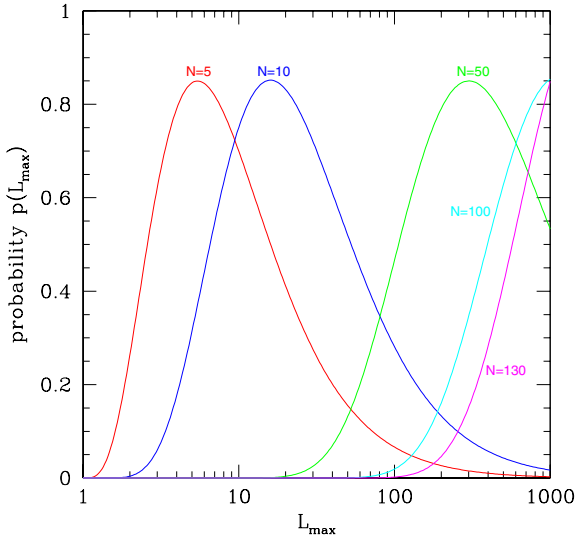

Figure 4. Probability distributions $p(L_{\max})$ of the luminosity of the brightest source in a sample of n sources with a power-law LF with slope $\alpha = 1.5$ and $L_1 = 1$, $L_2 = 10^3$. Each curve is marked according to the number of sources n.

2.3 The luminosity of the brightest source in a sample

The probability distribution for the luminosity of the brightest source observed in a sample of n sources equals

$$p(L_{\max}) = [p_1(L < L_{\max})]^{n-1} p_1(L_{\max}) n, \tag{27}$$

where $p_1(L)$ is the probability distribution for the luminosity of one source (e.g. equation 24) and $p_1(L < L_{\max})$ denotes the cumulative probability $p_1(L < L_{\max}) = \int_0^{L_{\max}} p_1(L)\,\mathrm{d}L$.

The $p(L_{\max})$ distribution for $\alpha = 1.5$ is shown in Fig. 4 and illustrates the intuitively obvious fact that if the number of sources is sufficiently small the brightest sources most likely will not reach the highest possible value of L_2. An analytical formula for the most probable value of the luminosity of the brightest source is given by equation (A1).

3 VARIABILITY OF THE TOTAL EMISSION

For a population of n sources with luminosities L_k and fractional rms of aperiodic variability rms_k, the fractional rms of the total emission equals

$$\mathrm{rms}^2_{\mathrm{tot}} = \frac{\sum_{k=1}^{n} L_k^2 \mathrm{rms}_k^2}{\left(\sum_{k=1}^{n} L_k\right)^2}, \tag{28}$$

assuming that variations of the source fluxes are uncorrelated with each other. In the following we also assume for simplicity that all sources have the same value of fractional rms, i.e. $\mathrm{rms}_k = \mathrm{rms}_0$.

In the limit of $n \to \infty$, corresponding to the linear regime in the $\tilde{L}_{\mathrm{tot}}$–n relation, the sums in equation (28) can be replaced by the respective integrals of the LF:

$$\frac{\mathrm{rms}^2_{\mathrm{tot}}}{\mathrm{rms}^2_0} = \frac{\int L^2 (\mathrm{d}N/\mathrm{d}L)\,\mathrm{d}L}{\left[\int L (\mathrm{d}N/\mathrm{d}L)\,\mathrm{d}L\right]^2} \propto \frac{1}{n} \propto \frac{1}{A} \propto \frac{1}{L_{\mathrm{tot}}}. \tag{29}$$

In the linear regime the fractional rms of the collective emission obeys a $\propto 1/\sqrt{n}$ averaging law, as expected for uncorrelated variations of individual sources.

In the non-linear regime, however, for a sufficiently flat luminosity function, the total luminosity is defined by a few brightest sources. To first approximation the number of such sources effectively contributing to the sums in equation (28) does not depend on the LF normalization. Consequently, the fractional rms of the total emission is constant, independently of the total number of sources or of their total luminosity. This contradicts the intuitive expectation that the fractional rms decreases with the number of sources as $\mathrm{rms} \propto 1/\sqrt{n}$.

To illustrate this behaviour we performed a series of Monte Carlo simulations for a power-law LF. For each set of parameters and a given value of the number of sources, in each run n sources were placed between L_1 and L_2 with a power-law luminosity distribution (equation 25). For each run, the ratio $\mathrm{rms}^2_{\mathrm{tot}}/\mathrm{rms}^2_0$ was computed following equation (28). From the results of many runs, the probability distribution for $\mathrm{rms}^2_{\mathrm{tot}}/\mathrm{rms}^2_0$ was obtained. The maximum of this distribution defines the most probable value of $\mathrm{rms}^2_{\mathrm{tot}}$ and its width characterizes the intrinsic dispersion of this quantity. Examples are shown in Fig. 5 for two values of the LF slope, $\alpha = 1.5$ and 2.5. For $\alpha < 1$ and $\alpha > 3$ there is no non-linear regime and the fractional rms of the total emission obeys the $\propto 1/\sqrt{n}$ law for any n. In plotting Fig. 5, we converted $\mathrm{rms}^2_{\mathrm{tot}}$ to $\mathrm{rms}_{\mathrm{tot}}$, so that the thick solid curves shows the behaviour of the square root of the most probable value of the $\mathrm{rms}^2_{\mathrm{tot}}/\mathrm{rms}^2_0$ ratio.

Approximate formulae for $\mathrm{rms}^2_{\mathrm{tot}}$ are derived in Appendix B.

4 ASTROPHYSICAL APPLICATIONS

4.1 Determining LF parameters from total emission of unresolved sources

The shape of the $\tilde{L}_{\mathrm{tot}}$–A relation (A is the normalization of the luminosity function) is defined by the parameters of the luminosity function. For $1 \lesssim \alpha \lesssim 2$, it has two distinct power-law regimes, separated by a break (Figs 2 and 3):

$$\tilde{L}_{\mathrm{tot}} \propto \begin{cases} A^{1/(\alpha-1)} & \text{for } A < A_{\mathrm{break}}, \\ A & \text{for } A > A_{\mathrm{break}}. \end{cases} \tag{30}$$

The position of the break between the non-linear and linear regimes is defined by the high-luminosity cut-off of the LF and its slope as described by equation (A13). This opens a possibility to determine the LF parameters without actually resolving the individual sources, but studying large samples of objects (e.g. galaxies) and the relation between their total emission and the normalization A. The value of A in many cases can be determined independently from observations at other wavelengths. For example, the normalizations of the luminosity functions of high- and low-mass X-ray binaries are proportional to the star formation rate (Grimm et al. 2003) and stellar mass (Gilfanov 2004) of the host galaxy, respectively. Both quantities can be determined from the conventional indicators, such as radio or near-infrared luminosities.

Of course with the subarcsecond angular resolution of *Chandra*, the luminosity distribution of point sources in nearby galaxies can be studied directly. However, for more distant galaxies, $D \gtrsim 30$–50 Mpc, even *Chandra* resolution becomes insufficient and only the total luminosity of the galaxy can be measured. Provided that the contaminating contribution of the emission of a central active galactic nucleus (AGN) and/or of hot X-ray-emitting gas can be constrained or separated, one can study the relation between the total luminosity L_{X} of a (unresolved) galaxy and its star formation rate or stellar mass. The L_{X}–SFR or L_{X}–M_* 'growth curves' constructed for large samples of galaxies and spanning a broad range of SFR and M_* can be used to constrain the XLF parameters of X-ray binaries in distant galaxies including those located at intermediate and high

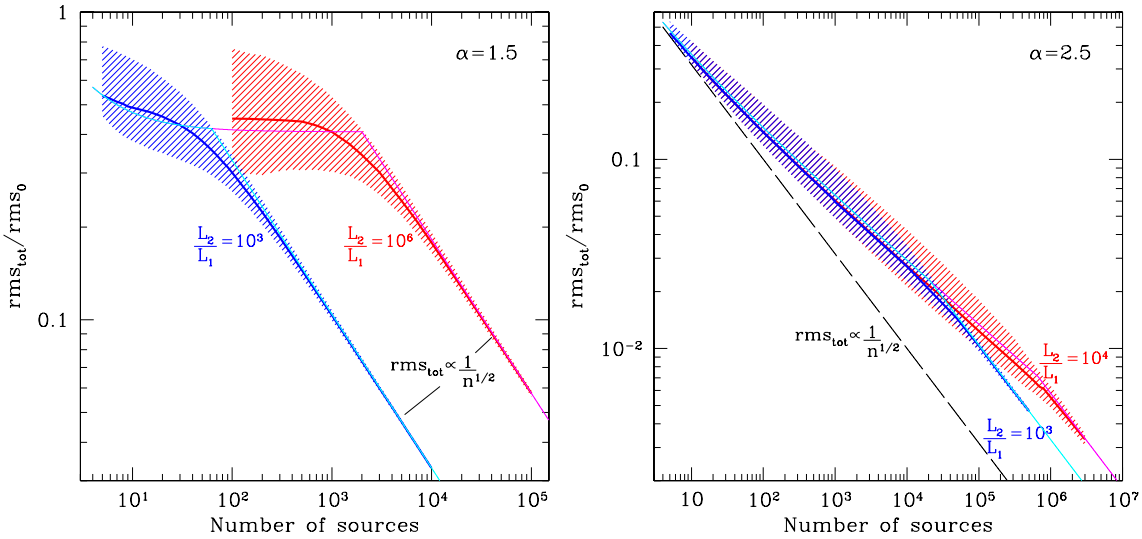

Figure 5. Variability of the total emission. The ratio of the fractional rms of the combined emission to that of one source, rms_0, as a function of the number of sources for different values of the LF slope α (indicated in the upper right corner of each plot) and different values of the ratio of L_2/L_1. The thick solid lines show the square root of the most probable value of $\mathrm{rms}^2_{\mathrm{tot}}/\mathrm{rms}^2_0$, and the shaded areas show the 67 per cent intrinsic dispersion, both obtained from the Monte Carlo simulations. The thin solid lines were computed using the approximation given by equation (B3). The dashed line shows $\mathrm{rms}_{\mathrm{tot}}/\mathrm{rms}_0 \propto 1/\sqrt{n}$ dependence.

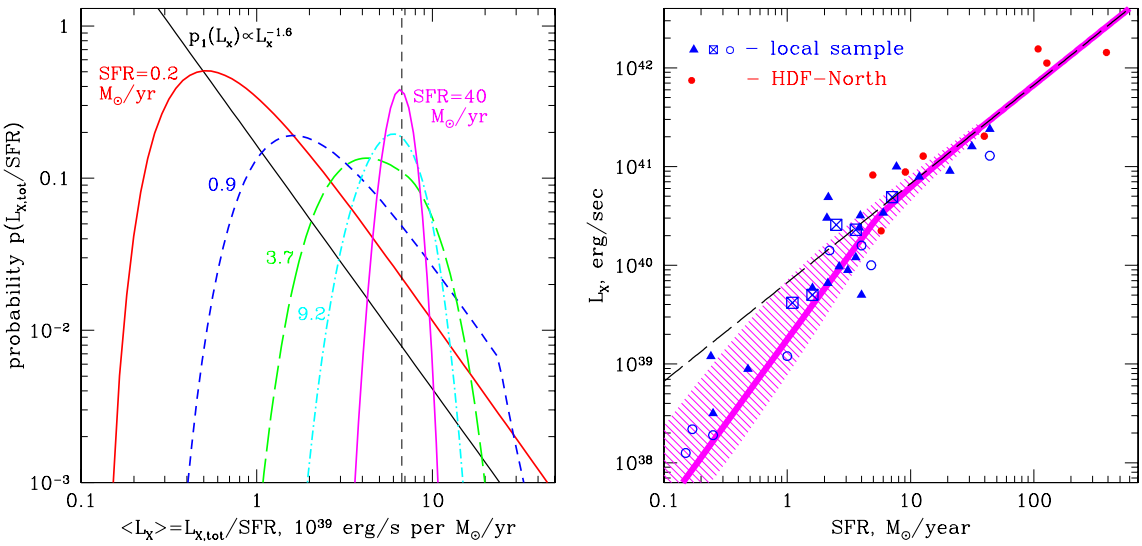

Figure 6. Left: The probability distributions $p(L_{\mathrm{tot}}/\mathrm{SFR})$ for different values of SFR. The vertical dashed line shows the expectation mean, defined by equation (4). Right: The L_{X}–SFR relation. The open circles are nearby galaxies observed by *Chandra*, the filled triangles are spatially unresolved nearby galaxies observed by *ASCA* and *BeppoSAX*, for which only total luminosity is available, and the filled circles are distant star-forming galaxies from the *Hubble Deep Field* North. The thick grey line is the relation between the SFR and the most probable value of the total luminosity, predicted from the 'universal' HMXB XLF, the shaded area shows its 67 per cent intrinsic spread, and the dashed line is the expectation mean, defined by equation (4). The five nearby galaxies, used by Grimm et al. (2003) to derive the 'universal' HMXB XLF, are marked as crossed boxes.

redshifts. With this, one can study the influence of a number of factors such as effects of binary evolution, metallicity, regimes of star formation, etc., on the luminosity distribution of X-ray binaries.

4.2 High-mass X-ray binaries in star-forming galaxies

4.2.1 L_{X}–SFR relation for star-forming galaxies

The slope of the 'universal' luminosity function of HMXBs, $\alpha = 1.6$, is in the range where the non-linear behaviour of the $\tilde{L}_{\mathrm{tot}}$–A relation is most pronounced. In the left panel of Fig. 6 we plot the probability distribution of the total luminosity $p(L_{\mathrm{tot}})$, computed for different values of the star formation rate. The distribution is strongly asymmetric in the non-linear low-SFR regime, which, for the parameters of the 'universal' HMXB X-ray luminosity function (XLF), corresponds to the formation rate of massive stars below $\mathrm{SFR} \lesssim 4$–$5\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$. Note that a small value of SFR does not necessarily imply a small total number of sources, which is defined by the (unknown) low-luminosity cut-off of the HMXB XLF. For example, for $\mathrm{SFR} = 0.2\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$, when the non-linear effect is

very pronounced, the total number of sources might be as large as ∼300 (∼1200) for the low-luminosity cut-off of 10^{34} (10^{33}) erg s^{-1}. These low values of the star formation rate correspond to the familiar examples of the Milky Way Galaxy and the Magellanic Clouds. On the opposite end among the relatively nearby and well-known galaxies are the Antennae interacting galaxies, which, with a star formation rate of SFR $\sim 7\,M_\odot\,yr^{-1}$, have a nearly symmetric $p(L_{tot})$ distribution, sufficiently close to the normal distribution.

The predicted L_X–SFR relation is shown in the right panel in Fig. 6, along with the measured values of X-ray luminosities and the star formation rates for a number of nearby galaxies and galaxies observed with *Chandra* at intermediate redshifts, $z \sim 0.2$–1.3, in the *Hubble Deep Field* North. The data shown in Fig. 6 are from Grimm et al. (2003), complemented with the local galaxies data from Ranalli, Comastri & Seti (2003). In plotting the latter we removed the duplications and the galaxies likely to be contaminated by the contribution of low-mass X-ray binaries, unrelated to the current star formation activity, as discussed by Gilfanov et al. (2004). The luminosities and star formation rates for the *Hubble Deep Field* North galaxies (Brandt et al. 2001) were computed for the following cosmological parameters: $H_0 = 70\,km\,s^{-1}\,Mpc^{-1}$, $\Omega_m = 0.3$ and $\Lambda = 0.7$, as described in Gilfanov et al. (2004).

Fig. 6 further illustrates the difference between the mode and the expectation mean of the $p(L_{tot})$ probability distribution. The thick solid line in the right panel shows the SFR dependence of the mode of $p(L_{tot})$ and predicts the *most probable* value of the X-ray luminosity of a randomly chosen galaxy. If observations of many (different) galaxies with similar values of SFR are performed, the obtained values of L_{tot} will obey the probability distribution depicted in the left panel. The average of the measured values of L_{tot} will be equal to the expectation mean given by equation (4) and shown by the dashed straight lines in the left and right panels. Owing to the properties of $p(L_{tot})$, these two quantities are not identical in the low-SFR limit when the total luminosity is defined by the small number of the most luminous sources. Only in the large SFR limit when there are sufficiently many sources with luminosities near the cut-off of the luminosity function, $\log(L_X) \sim 40$, does the L_X–SFR relation behave in the intuitively expected way.

Owing to the skewness of the probability distribution $p(L_{tot})$, large and asymmetric dispersion among the measured values of L_{tot} is expected in the non-linear low-SFR regime. This asymmetry is already seen in Fig. 6 – at low SFR values there are more points above the thick solid curve than below. Moreover, the galaxies lying significantly above the solid and dashed curves in Fig. 6 should be expected at low SFR and will inevitably appear as the plot is populated with more objects. Such behaviour differs from a typical astrophysical situation and should not be ignored when analysing and fitting the L_X–SFR relation in the low-SFR regime. In particular, owing to non-Gaussianity of the $p(L_{tot})$ distribution, the standard data analysis techniques – least-squares and χ^2 fitting – become inadequate.

4.2.2 *High-luminosity cut-off in the HMXB XLF*

The position of the break between the non-linear and linear parts of the L_X–SFR relation depends on the LF slope and its cut-off luminosity (Fig. 7, equation A13): $SFR_{break} \propto L_{cut}^{\alpha-1}$. This allows one to constrain the parameters of the luminosity distribution of compact sources using the data of spatially unresolved galaxies as discussed in Section 4.1.

Agreement of the predicted L_X–SFR relation with the data both in high- and low-SFR regimes confirms the universality of the HMXB

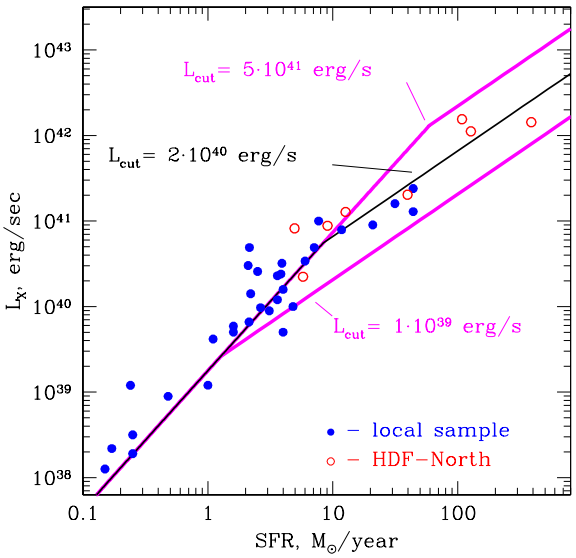

Figure 7. Dependence of the L_X–SFR relation on the XLF cut-off luminosity L_{cut}. The three curves corresponding to different values of L_{cut} coincide in the non-linear low-SFR regime but differ in the position of the break between linear and non-linear regimes. The data are the same as in Fig. 6.

luminosity function, derived by Grimm et al. (2003) from significantly fewer galaxies (shown as crossed boxes in Fig. 6) than plotted in Figs 6 and 7. It provides an independent confirmation of the existence of a cut-off in the HMXB XLF at $\log(L_{cut}) \sim 40.5$, including HMXBs in spatially unresolved high-redshift galaxies from the *Hubble Deep Field* North. This implies, in particular, that ultraluminous X-ray sources (ULXs) at redshifts of $z \sim 0.2$–1.3 were not significantly more luminous than those observed in the nearby galaxies.

4.2.3 *Aperiodic variability*

The X-ray flux from X-ray binaries is known to be variable in a broad range of time-scales, on the orders of milliseconds to years. In addition to a number of coherent phenomena and quasi-periodic oscillations, significant continuum aperiodic variability is often observed. The fractional rms of aperiodic variations depends on the nature of the binary system and the spectral state of the X-ray source and is usually in the range from a fraction of a per cent to ∼20–30 per cent. Correspondingly, the combined emission of X-ray binaries in a galaxy should also be variable in a similarly broad range of time-scales. It has been suggested by Grimm, Gilfanov & Sunyaev (in preparation) that, as a result of a large difference in the characteristic time-scales of the accretion flow on to a stellar-mass object and on to a supermassive black hole, variability of the X-ray emission from a galaxy can be used to distinguish between the combined emission of a population of X-ray binaries and that of an accreting supermassive black hole in the centre of a galaxy (AGN).

From results of Section 3 and Appendix B one should expect that in the non-linear low-SFR regime the fractional rms of the total X-ray emission of a star-forming galaxy is independent of SFR. This prediction is confirmed by the results of the Monte Carlo simulations performed as described in Section 3 and shown in Fig. 8. For moderate star formation rates, SFR $\lesssim$ 5–10 $M_\odot\,yr^{-1}$, we predict a rather large aperiodic variability of the total emission of HMXBs at the level of about one-third to one-half of the fractional rms of individual X-ray binaries. At larger values of SFR, corresponding to the linear

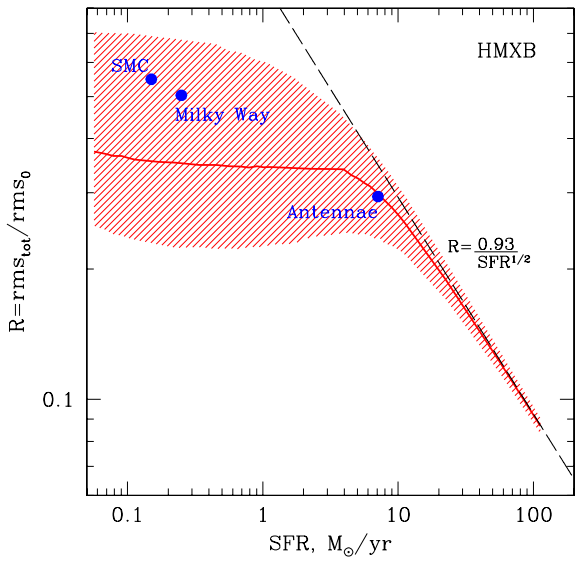

Figure 8. Variability of the total emission of HMXBs in star-forming galaxies. Dependence of the ratio $\mathrm{rms_{tot}}/\mathrm{rms_0}$ on the star formation rate. The thick solid line shows the most probable value of $\mathrm{rms_{tot}}/\mathrm{rms_0}$, and the shaded area shows its 67 per cent intrinsic dispersion, both obtained from the Monte Carlo simulations for a power-law LF with the slope of $\alpha = 1.6$ and a cut-off at $L_2 = 2 \times 10^{40}$ erg s^{-1}. The dashed line shows the asymptotic behaviour at large SFR. The filled circles correspond to HMXB sources in the Milky Way, the SMC and the Antennae galaxies, computed from equation (28), using the observed luminosities of X-ray sources in these galaxies.

regime in the $L_{\rm X}$–SFR relation, it decreases as rms $\propto 1/\sqrt{\rm SFR}$, in accord with equation (B3). Also shown in Fig. 8 are values of the fractional rms reduction, computed directly from equation (28) using the luminosities of the observed HMXBs in the Milky Way (Grimm et al. 2002), the Small Magellanic Cloud (Yokogawa et al. 2000) and the Antennae galaxies (Zezas et al. 2002). Note that the predicted rms–SFR relation can be modified by the luminosity dependence of the rms of individual sources. This factor might become especially important at large values of SFR when the total luminosity of a star-forming galaxy is dominated by ULXs whose variability properties we know little about.

4.2.4 Luminosity of the brightest source

As the first *Chandra* observations of compact sources in nearby galaxies have become available, it has been noted (e.g. Sarazin et al. 2001; Irwin et al. 2002; Fabbiano & White 2003) that the luminosity of the brightest X-ray binary in a galaxy might depend on its properties. In particular, in the case of high-mass X-ray binaries it appeared to correlate with the star formation rate of the host galaxy. For example, in the Antennae galaxies, a number of compact sources have been discovered with luminosities of $\sim 10^{40}$ erg s^{-1} (Zezas et al. 2002). On the other hand, the luminosities of the brightest HMXB sources in the Milky Way do not exceed $\lesssim 10^{38}$ erg s^{-1} (Grimm et al. 2002). It has been argued that this might reflect the difference in the intrinsic source properties, related to the difference in the galactic environment and in initial conditions for X-ray binary formation in starburst galaxies and in galaxies with weak and steady star formation.

However, as discussed in Section 2.3, the probability distribution for the luminosity of the brightest source in a galaxy, $p(L_{\rm max})$, depends on the LF normalization, i.e. on the SFR of the host galaxy in the case of HMXBs. The luminosity of the brightest source increases with SFR, until it reaches the maximum possible value, defined by the high-luminosity cut-off of the LF. The $p(L_{\rm max})$ distributions, computed from equation (27) for the parameters of the 'universal' HMXB XLF, are shown for different values of SFR in the left panel of Fig. 9. The right panel in Fig. 9 shows the dependence of the most probable value of the luminosity of the brightest HMXB on

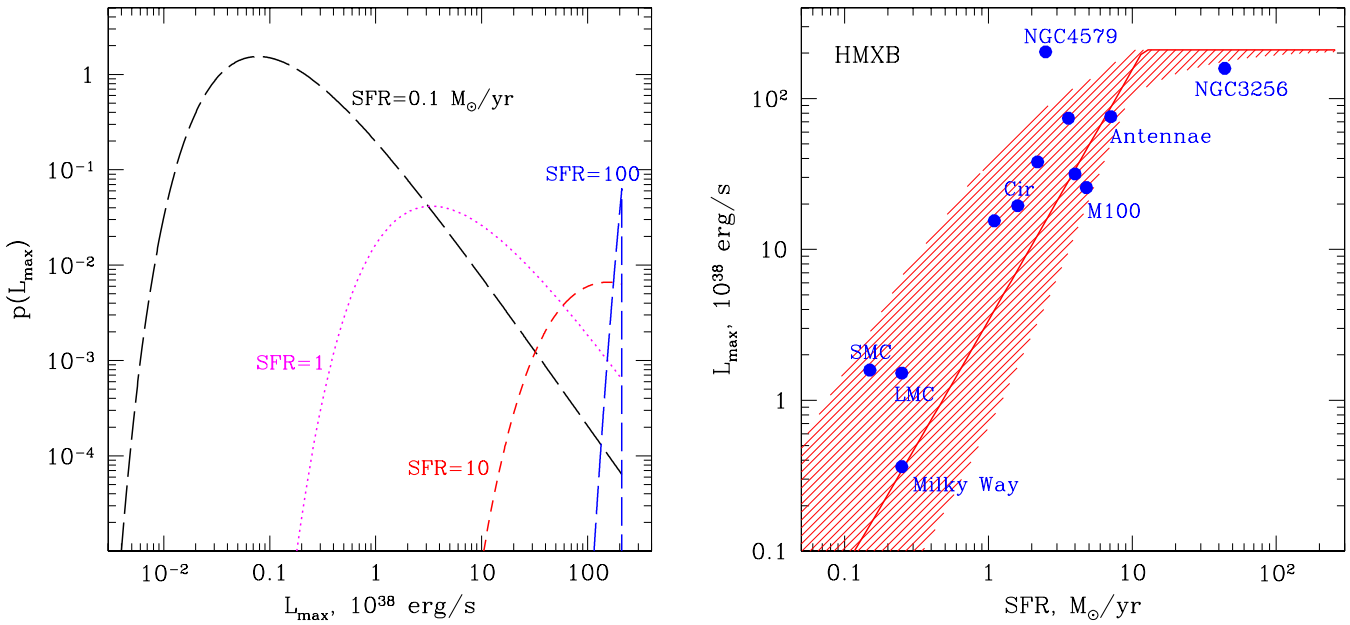

Figure 9. Left: The probability distribution of the luminosity of the brightest HMXB source for different values of the star formation rate, computed from equation (27) using the parameters of the 'universal' luminosity function of HMXBs. Right: Expected luminosity of the brightest HMXB versus star formation rate. The thick solid line shows the most probable value of $L_{\rm max}$ (equation A1), and the shaded area shows its 67 per cent intrinsic dispersion, obtained from the probability distribution given by equation (27). The filled circles show maximum observed luminosity of HMXBs in the Milky Way and several nearby star-forming galaxies.

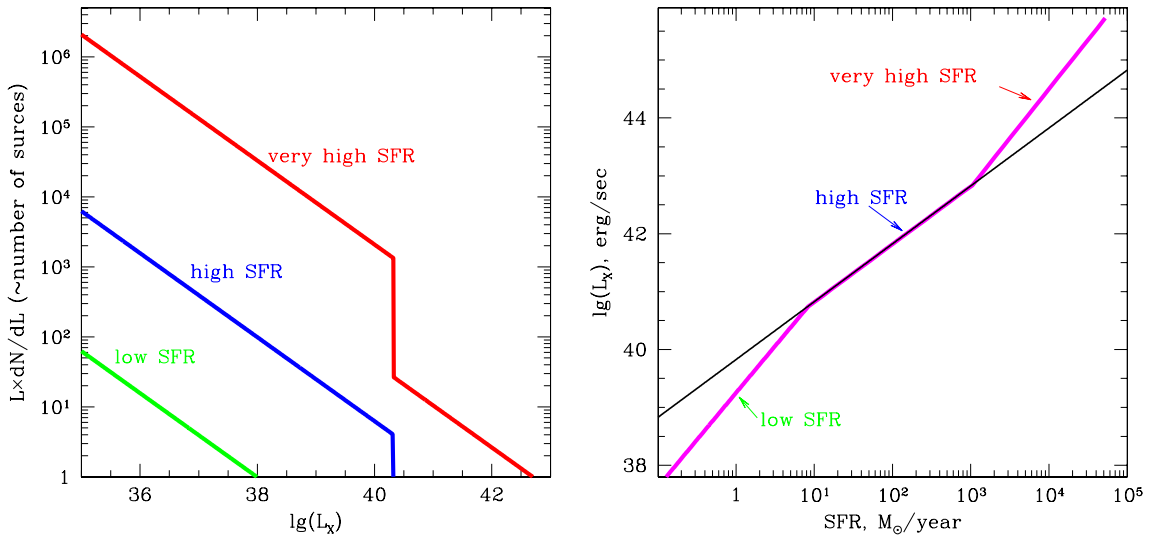

Figure 10. Illustration of the effect of hypothetical intermediate-mass black holes on the L_X–SFR relation. Left: The luminosity function of compact sources at different levels of star formation rate. Right: Corresponding L_X–SFR relation. The thin straight line shows the linear dependence.

the SFR of the host galaxy, described by equation (A1), along with its intrinsic 67 per cent uncertainty. Filled symbols in Fig. 9 are the luminosities of the brightest source observed in star-forming galaxies from the sample of Grimm et al. (2003, and references therein). As is clear from the plot, the large difference in the maximum luminosity between low- and high-SFR galaxies, e.g. between the Milky Way and the Antennae galaxies, can be naturally understood in terms of the properties of the probability distribution $p(L_{\max})$. No additional physical processes affecting HMXB formation in starburst galaxies need to be invoked.

4.3 Intermediate-mass black holes

The hypothetical intermediate-mass black holes (BHs), probably reaching masses of $\sim 10^{2-5} M_\odot$, might be produced, for example, via black hole mergers in dense stellar clusters, and can be associated with extremely high star formation rates. To accrete efficiently, they should form close binary systems with normal stars or be located in dense molecular clouds. It is natural to expect that such objects are significantly less frequent than stellar-mass black holes. The transition from stellar-mass BH HMXB to intermediate-mass BHs should manifest itself as a step in the luminosity distribution of compact sources (Fig. 10, left panel). If the cut-off in the HMXB XLF observed at $\log(L_{\rm cut}) \sim 40.5$ corresponds to the maximum possible luminosity of 'ordinary' stellar-mass black holes and if at $L > L_{\rm cut}$ a population of hypothetical intermediate-mass BHs emerges, it should lead to a drastic change in the slope of the L_X–SFR relation at extreme values of SFR (Fig. 10, right panel). Therefore, observations of the L_X–SFR relation for distant star-forming galaxies with very high SFR might be an easy way to probe the population of intermediate-mass black holes.

4.4 Low-mass X-ray binaries

4.4.1 L_X–stellar mass relation and maximum luminosity

As was shown by Gilfanov (2004) the luminosity distribution of low-mass X-ray binaries in nearby early-type galaxies and bulges of spiral galaxies can be described by a 'universal' XLF whose shape is approximately the same in different galaxies and whose normalization is proportional to the stellar mass of a galaxy. The shape of the 'universal' LMXB XLF is significantly more complex than that of HMXBs. It appears to follow the L^{-1} power law at low luminosities, gradually steepens at $\log(L_X) \gtrsim 37.0$–37.5, and has a rather abrupt cut-off at $\log(L_X) \sim 39.0$–39.5. In the $\log(L_X) \sim 37.5$–38.7 luminosity range, it approximately follows a power law with the differential slope of $\approx$1.8–1.9.

Given the shape of the XLF, the total luminosity of LMXB sources in a galaxy is defined by the sources with $\log(L_X) \sim 37$–38, the contribution of the brighter and, especially, weaker sources being less significant. Therefore, the non-linear regime in the L_X–M_* relation, although it does exist for $\log(M_*) \lesssim 10.0$–10.5, is less important than in the L_X–SFR relation for high-mass X-ray binaries (see, for example, fig. 14 in Gilfanov 2004).

Significantly more pronounced is the dependence of the luminosity of the brightest source on the LF normalization, i.e. on the stellar mass of the host galaxy. In order to study this dependence, we used the broken power-law approximation for the LMXB XLF from Gilfanov (2004) and performed a series of Monte Carlo simulations, similar to those described in Section 3. The probability distribution of the maximum luminosity $p(L_{\max})$, obtained from these simulations, is shown in the left panel in Fig. 11 for different values of the stellar mass of the host galaxy. The right panel shows the dependence of the most probable value of the maximum luminosity and of its 67 per cent intrinsic spread on the stellar mass. Its broken line shape is a result of the broken power-law approximation of the LMXB XLF used in the simulations. Solid symbols show the observed values of the maximum luminosity for the number of nearby early-type galaxies, bulges of spiral galaxies and for LMXBs in the Milky Way from the sample of Gilfanov (2004, and references therein).

Similarly to HMXBs in star-forming galaxies, it is obvious from Fig. 11 that the significant difference in the value of the luminosity of the brightest source can be naturally explained by the properties of the luminosity function of LMXBs. The same effect leads to an artificial (unphysical) dependence of the average luminosity of low-mass X-ray binaries in a galaxy on its stellar mass (e.g. fig. 17 in Gilfanov 2004). So far there is no evidence for a significant change

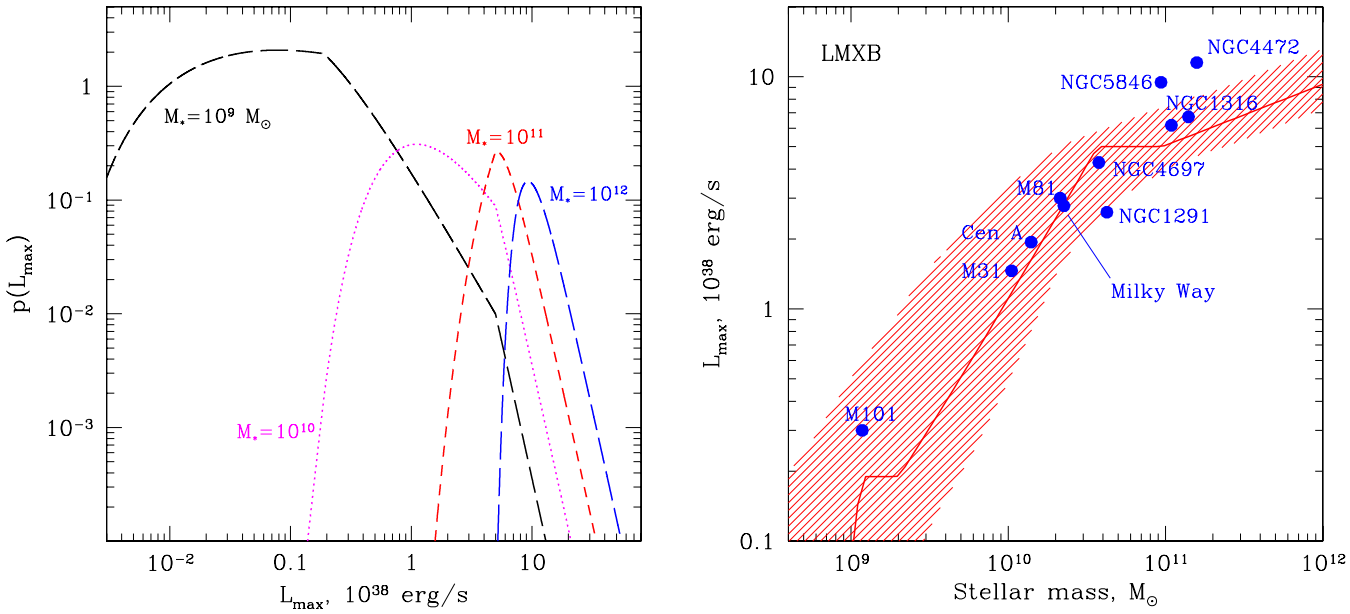

Figure 11. Left: Probability distribution for the luminosity of the brightest LMXB for different values of the stellar mass, computed from equation (27), using the parameters of the broken power-law approximation to the 'universal' luminosity function of LMXBs. Right: Expected luminosity of the brightest LMXB source versus stellar mass of the host galaxy. The thick solid line shows the most probable value of $L_{\rm max}$, and the shaded area shows its 67 per cent intrinsic dispersion, both obtained from the probability distribution given by equation (27). The filled circles show maximum observed luminosities of LMXB sources in the Milky Way and several nearby galaxies, studied with *Chandra*. The 'broken' shape of the predicted dependence is a consequence of the broken power-law approximation of the 'universal' LMXB XLF, used in the calculations.

of intrinsic properties of low-mass X-ray binaries with galactic environment. The difference between the luminosity of the brightest LMXB in massive elliptical galaxies and the bulges of spiral galaxies can be understood based on the probability arguments.

4.4.2 Variability

As in the case of HMXBs, in the linear large-mass limit, $\log(M_*) \gtrsim 10.5$, the fractional rms of the aperiodic variability of the combined emission of LMXBs follows the $\mathrm{rms}_{\rm tot} \propto 1/\sqrt{M_*}$ averaging law. Owing to the shape of the LMXB XLF, it decreases rather quickly with the stellar mass of the galaxy in the non-linear low-mass regime as well (Fig. 12). Consequently, in massive elliptical galaxies, with stellar mass $\log(M_*) \sim 11.0$–11.5, the fractional rms variability of the total emission will be suppressed by a factor of ∼10–15 with respect to the rms of individual sources. In a galaxy similar to the Milky Way, with $\log(M_*) \sim 10.5$–10.7, the suppression factor is ∼5. Considerable variability on the level of about one-quarter to one-half of that of individual X-ray binaries can be expected only for light bulges of spiral galaxies with masses in the $\log(M_*) \sim 9.5$–10.5 range.

Fig. 13 compares the dependence of the fractional rms on the most probable value of the total luminosity for high- and low-mass X-ray binaries. In the bright luminosity end, $\log(L_{\rm X}) \gtrsim 39.5$, the X-ray emission from early-type galaxies is expected to be significantly, up to a factor of ∼7, less variable than from star-forming galaxies.

5 SUMMARY

We studied the statistical properties of the combined emission of a population of discrete sources. Namely, we considered the properties of their total luminosity

$$L_{\rm tot} = \sum_{k=1}^{n} L_k, \qquad (31)$$

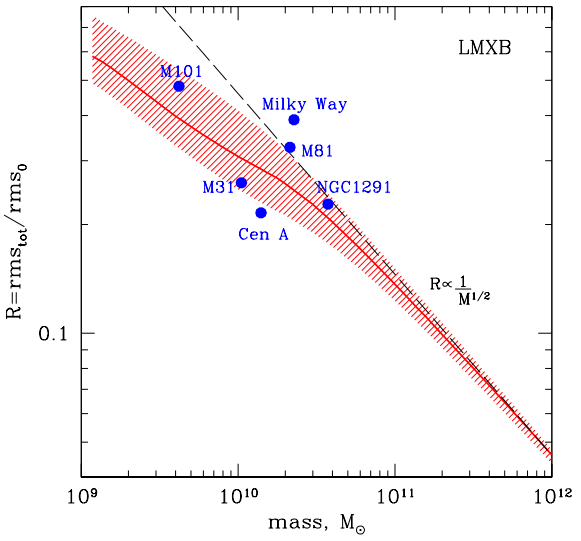

Figure 12. Variability of the total emission of low-mass X-ray binaries. The dependence of the ratio $\mathrm{rms}_{\rm tot}/\mathrm{rms}_0$ on the stellar mass of the galaxy. The thick solid line shows the most probable value of $\mathrm{rms}_{\rm tot}/\mathrm{rms}_0$, and the shaded area shows its 67 per cent intrinsic dispersion, both obtained from the Monte Carlo simulations for the 'universal' luminosity function from Gilfanov (2004). The dashed line shows the asymptotic behaviour at large M_*. The filled circles correspond to LMXBs in the Milky Way and several nearby galaxies, computed from equation (28), using the observed luminosities of X-ray sources in these galaxies.

and its dependence on the number of sources n or, equivalently, on the normalization of the luminosity function (LF). Using high-mass X-ray binaries in star-forming galaxies as an example, L_k are the luminosities of individual X-ray binaries in a galaxy and $L_{\rm tot}$ is its total X-ray luminosity due to HMXB population. In this example the

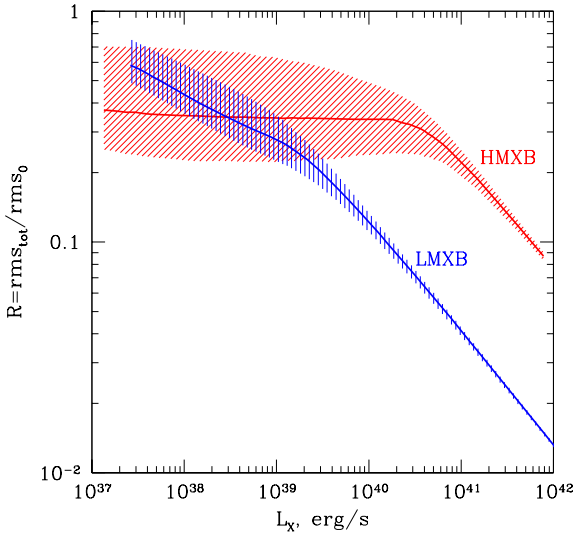

Figure 13. Comparison of the variability of low- and high-mass X-ray binaries. The dependence of the ratio $\mathrm{rms_{tot}}/\mathrm{rms_0}$ on the total X-ray luminosity of the galaxy due to X-ray binaries. The thick solid lines show the most probable value of $\mathrm{rms_{tot}}/\mathrm{rms_0}$, and the shaded areas show its 67 per cent intrinsic dispersion, both obtained from the Monte Carlo simulations for the respective 'universal' luminosity functions from Grimm et al. (2003) and Gilfanov (2004). For both curves, the linear parts at high L_X follow the rms $\propto 1/\sqrt{L_X}$ averaging law.

normalization of the luminosity function, i.e. the number of HMXBs in the galaxy n, is proportional to its star formation rate. We showed that, as a result of the statistical properties of the probability distribution $p(L_{\rm tot})$, the result of a measurement of the total luminosity of a randomly chosen galaxy might deviate significantly from the intuitively obvious expression

$$\langle L_{\rm tot}\rangle = \int_0^{+\infty} L\,\frac{{\rm d}N}{{\rm d}L}\,{\rm d}L \propto n \propto {\rm SFR}. \qquad (32)$$

These properties of $p(L_{\rm tot})$ can result in a surprising non-linear dependence of the total luminosity on n. They can also cause anomalous variability of the combined emission in an apparent violation of the rms $\propto 1/\sqrt{n}$ averaging law for uncorrelated variations of individual sources.

Our results can be summarized as follows.

(i) The probability distribution $p(L_{\rm tot})$ can be computed numerically from equation (20) or alternatively equation (23). Examples are shown in Fig. 1.

(ii) The relevant characteristics of the $p(L_{\rm tot})$ distribution are: its mode $\tilde{L}_{\rm tot}$ – the value of the random variable $L_{\rm tot}$ for which $p(L_{\rm tot})$ has the maximum – and the expectation mean $\langle L_{\rm tot}\rangle$ defined by equation (32). It is the mode of the $p(L_{\rm tot})$ distribution that predicts the most probable value of the total luminosity of a randomly chosen galaxy. If many galaxies with approximately the same total number of sources n are observed, the measured values of their total X-ray luminosities are distributed according to the $p(L_{\rm tot})$ distribution whose shape depends on the chosen value of n (Fig. 6, left panel). The average of the measured values of $L_{\rm tot}$ is equal to the expectation mean $\langle L_{\rm tot}\rangle$ and is always proportional to n in agreement with equation (32).

(iii) For small values of the LF normalization, $\tilde{L}_{\rm tot}$ and $\langle L_{\rm tot}\rangle$ do not equal each other (Fig. 1) and the $\tilde{L}_{\rm tot}$–n relation is non-linear (Figs 2, 3 and 6). Only in the limit of $n \gg 1$ does $\tilde{L}_{\rm tot} = \langle L_{\rm tot}\rangle$ and $\tilde{L}_{\rm tot}$ depend on n linearly. The threshold value of n depends on the LF shape and can be arbitrarily large.

(iv) The skewness of the $p(L_{\rm tot})$ probability distribution in the non-linear regime (Fig. 1) results in an enhanced and significantly asymmetric intrinsic dispersion of the measured values of $L_{\rm tot}$ (Figs 3 and 6). Its non-Gaussianity precludes the use of the standard fitting techniques in analysing the $L_{\rm tot}$ –n relation (e.g. L_X–SFR relation for star-forming galaxies), such as χ^2 minimization technique.

(v) The variability of the total emission (e.g. aperiodic variability of X-ray emission of a galaxy due to superposition of variabilities of individual sources) in the non-linear regime decreases with the number of sources more slowly than a rms $\propto 1/\sqrt{n}$ law for uncorrelated variations of individual sources, resulting in anomalously high variability of the total emission. In the linear regime the rms $\propto 1/\sqrt{n}$ dependence is restored (Section 3, Fig. 5).

(vi) The amplitude of the discussed effect depends on the shape of the luminosity function. For a power-law LF it is strongest for $1 < \alpha < 2$, is unimportant for shallow luminosity functions with $\alpha < 1$, and gradually diminishes with increase of α above $\alpha = 2$ (Figs 1–3).

We illustrate these results using the example of the combined emission of X-ray binaries in galaxies and its dependence on the star formation rate and on the stellar mass of the host galaxy.

(i) For the slope of the HMXB 'universal' XLF, $\alpha \approx 1.6$, the discussed effects are strongest with a significant non-linear regime in the L_X–SFR relation at SFR $\lesssim$ 4–5 $\mathrm{M_\odot\ yr^{-1}}$. The predicted L_X–SFR dependence is in good agreement with observations (Fig. 6). Given the shape of the 'universal' LMXB XLF, no significant non-linearity of the L_X–M_* relations is expected, also in a good agreement with observations.

(ii) The L_X–SFR relation can be used to constrain the XLF parameters of HMXBs in distant unresolved galaxies including the galaxies observed with *Chandra* in *Hubble Deep Field* North at redshifts $z \sim 0.2$–1.3 (Sections 4.2.2 and 4.3, Figs 7 and 10).

(iii) For both high- and low-mass X-ray binaries a strong dependence of the luminosity of the brightest source on the SFR and stellar mass of the host galaxy is expected. The $L_{\rm max}$–SFR and $L_{\rm max}$–M_* dependences predicted from the respective 'universal' XLFs explain well the results of *Chandra* observations of nearby galaxies (Figs 9 and 11). The significant difference in the luminosity of the brightest LMXB between bulges of spiral galaxies and giant ellipticals or between the brightest HMXB in the Milky Way and in starburst galaxies can be understood based solely on probability arguments.

(iv) We predict enhanced variability of X-ray emission from star-forming galaxies due to HMXBs, significantly above the $\propto 1/\sqrt{n}$ averaging law. For SFR $\lesssim 5\ \mathrm{M_\odot\ yr^{-1}}$ the expected fractional rms of variability of the combined emission of HMXBs does not depend on the star formation rate and approximately equals about one-third to one-half of the rms of individual sources (Fig. 8). On the contrary, variability of X-ray emission from early-type galaxies due to LMXBs will be significantly suppressed because of the averaging effect (Fig. 12), up to ~3–10 times in the $\sim 10^{10}$–$10^{11}\ \mathrm{M_\odot}$ stellar mass range. For the same total luminosity, star-forming galaxies are expected to have significantly larger fractional rms than massive elliptical and S0 galaxies assuming that the fractional rms of individual sources are comparable (Fig. 13).

ACKNOWLEDGMENTS

We thank the anonymous referee for comments that helped to improve the presentation of the paper.

REFERENCES

Brandt W. N. et al., 2001, AJ, 122, 2810
Cervino M., Luridana V., 2004, A& A, 413, 145
Fabbiano G., White N., 2003, in Lewin W., van der Klis M., eds, Compact Stellar X-ray Sources, Cambridge Univ. Press, Cambridge (astro-ph/0307077)
Gilfanov M., 2004, MNRAS, 349, 146
Gilfanov M., Grimm H.-J., Sunyaev R., 2004, MNRAS, 347, L57
Grimm H.-J., Gilfanov M., Sunyaev R., 2002, A& A, 391, 923
Grimm H.-J., Gilfanov M., Sunyaev R., 2003, MNRAS, 339, 793
Illarionov A. F., Sunyaev R. A., 1975, A& A, 39, 185
Irwin J. A., Sarazin C. L., Bregman J. N., 2002, ApJ, 570, 152
Kalogera V., Narayan R., Spergel D. N., Taylor J. H., 2001, ApJ, 556, 340
Ranalli P., Comastri A., Seti G., 2003, A& A, 399, 39
Sarazin C. L., Irwin J. A., Bregman J. N., 2001, ApJ, 556, 533
Shtykovskiy P., Gilfanov M., 2004, A&A, submitted (astro-ph/0404300)
Yokogawa J., Imanishi K., Tsujimoto M., Nishiuchi M., Koyama K., Nagase F., Corbet R. H. D., 2000, ApJS, 128, 491
Zezas A., Fabbiano G., Rots A. H., Murray S. S., 2002, ApJS, 142, 239

APPENDIX A: APPROXIMATE SOLUTION FOR THE TOTAL LUMINOSITY

We consider the case of a power-law LF (equation 25), with a slope $\alpha > 0$. The probability distribution for the luminosity of the brightest source in the sample of n sources is defined by equation (27). The maximum of this distribution gives the most probable value of the luminosity of the brightest sources:

$$\tilde{L}_{\max} = \min(L'_{\max}, L_2),$$
$$\left(\frac{L'_{\max}}{L_1}\right)^{\alpha-1} = 1 + \frac{\alpha-1}{\alpha}(n-1). \tag{A1}$$

Similarly, the probability distribution of the minimum luminosity in the sample is

$$p(L_{\min}) = [p_1(L > L_{\min})]^{n-1} p_1(L_{\min}) n. \tag{A2}$$

Contrary to $p(L_{\max})$, $p(L_{\min})$ declines steeply at $L > L_1$ for any n (for $\alpha > 0$). Within the accuracy of this approximation we can assume that $p(L_{\min}) = \delta(L_{\min} - L_1)$, i.e. $L_{\min} = L_1$.

The total luminosity of n sources distributed between L_1 and $L_{\max}$ according to the power law with slope α ($\alpha \neq 1, \alpha \neq 2$) can be approximated as

$$L_{\rm tot} \approx n \frac{1-\alpha}{2-\alpha} \frac{L_{\max}^{2-\alpha} - L_1^{2-\alpha}}{L_{\max}^{1-\alpha} - L_1^{1-\alpha}}. \tag{A3}$$

Knowing the probability distribution $p(L_{\max})$, the probability distribution $p_n(L_{\rm tot})$ can be calculated as

$$p_n(L_{\rm tot}) \approx p(L_{\max}) \left(\frac{{\rm d}L_{\rm tot}}{{\rm d}L_{\max}}\right)^{-1}, \tag{A4}$$

where $L_{\max} = L_{\max}(L_{\rm tot})$ is the inverse function to equation (A3) and $p(L_{\max})$ is given by equation (27).

The most probable value of $L_{\rm tot}$ is defined by the condition

$$\frac{{\rm d}p_n(L_{\rm tot})}{{\rm d}L_{\rm tot}} = 0. \tag{A5}$$

With equations (27), (A3) and (A4) the above equation can be transformed to:[2]

$$(\alpha-2)\xi^{2\alpha} - (\alpha-2)(1+\alpha-n+\alpha n)\xi^{1+\alpha} + (\alpha-1)^2(1+n)\xi^{\alpha} + [1+(\alpha-1)n]\xi^2 = 0, \tag{A6}$$

with $\xi = L_{\max}/L_1$, or, equivalently,

$$n = \frac{(\alpha-2)\xi^{2\alpha} + (2+\alpha-\alpha^2)\xi^{1+\alpha} + (\alpha-1)^2\xi^{\alpha} - \xi^2}{(\alpha-1)[(\alpha-2)\xi^{1+\alpha} - (\alpha-1)\xi^{\alpha} + \xi^2]} \tag{A7}$$

with $\alpha \neq 1, \alpha \neq 2$. Because of the simplifying assumption $p_{\min}(L_{\min}) = \delta(L_{\min} - L_1)$ and the approximate nature of equation (A3), the probability distribution defined by equation (A4) is valid only for $L_{\rm tot} < \langle L_{\rm tot}\rangle$ and is undefined otherwise. This, however, is sufficient for our purpose as $L_{\max} \sim L_2$ corresponds to the break in the $\tilde{L}_{\rm tot}$–n relation (Figs 2 and 3), above which $\tilde{L}_{\rm tot} = \langle L_{\rm tot}\rangle$.

A.1 The practical recipe

The $\tilde{L}_{\rm tot}$–n relation can be computed parametrically using equations (A7) and (A3) The practical recipe is for a set of values of $L_{\max}$, $L_1 < L_{\max} \leqslant L_2$, to compute n from equation (A7) and $L_{\rm tot}$ from equation (A3). The pairs of values ($L_{\rm tot}$, n) define the $\tilde{L}_{\rm tot}$–n relation before and up to the break. Above the break, $\tilde{L}_{\rm tot} = \langle L_{\rm tot}\rangle$ and can be computed from equation (A3) with $L_{\max} = L_2$ and $n > n_{\rm break}$ a free parameter. The n in the obtained $\tilde{L}_{\rm tot}$–n relation can be transformed to the normalization A via equation (11).

The approximation defined by the equations (A7) and (A3) is compared with the results of the exact calculation in Fig. 2. It is accurate within several per cent everywhere, except for the break region, where its accuracy is ~10–20 per cent.

A.2 Asymptotics

Using the approximate solution for $\tilde{L}_{\rm tot}$ from Appendix A we consider the asymptotic behaviour of the $\tilde{L}_{\rm tot}$–n relation in the limit of $L_2/L_1 \to \infty, \alpha > 1$.

From equation (A7) variable ξ is related to the number of sources by (see footnote 2):

$$\xi^{\alpha-1} = (\alpha-1)n + O(1), \qquad L_{\max} = \xi L_1, \tag{A8}$$

in the limit of $n \gg 1$ or equivalently $L_{\max} \gg L_1$. Although in the limit $n \to \infty$ this approximation is valid for any $\alpha > 1$, its accuracy deteriorates considerably for $\alpha \lesssim 1.6$ where it can be improved by replacing n in equation (A8):

$$n \to n - \frac{1+\alpha-\alpha^2}{(\alpha-1)(\alpha-2)}, \qquad 1 < \alpha \lesssim 1.6. \tag{A9}$$

The most probable value of the total luminosity is given by

$$\frac{\tilde{L}_{\rm tot}}{\langle L_{\rm tot}\rangle} \approx \frac{1 - [(\alpha-1)n]^{(2-\alpha)/(\alpha-1)}}{1 - (L_2/L_1)^{2-\alpha}} \tag{A10}$$

with $\alpha > 1$, $\alpha \neq 2$ and $n \gg 1$. Using equations (11) and (12) it can be expressed via the normalization of the luminosity function A or transformed to the relations for $\tilde{L}_{\rm tot}$. As with equation (A8), the accuracy of equation (A10) can be significantly improved using equation (A9) for $\alpha \lesssim 1.6$.

[2] Note that in equations (A6) and (A7) $L_{\max}$ is a parameter rather than the most probable value of the maximum luminosity. The latter is defined by equation (A1), which is exact.

For $1 < \alpha < 2$ the $\tilde{L}_{\rm tot}$–n relation shows a sharp break between the non-linear and linear regimes (Figs 2 and 3). From equation (A10) one can obtain

$$\tilde{L}_{\rm tot} \propto \begin{cases} n^{1/(\alpha-1)} & \text{for } n < n_{\rm break,} \\ n & \text{for } n > n_{\rm break.} \end{cases} \tag{A11}$$

This is in an agreement with equation (8) based on simple qualitative arguments. The position of the break in the $\tilde{L}_{\rm tot}$–n relation can be obtained from equation (A7) by substituting $\xi = L_2/L_1$ and using the fact that $\xi \gg 1$:

$$n_{\rm break} \approx \frac{1}{\alpha - 1} \left(\frac{L_2}{L_1} \right)^{\alpha-1}. \tag{A12}$$

Expressed in terms of the normalization A of the luminosity function equation (25) it is

$$A_{\rm break} \approx L_2^{\alpha-1}. \tag{A13}$$

As intuitively expected, the break position expressed in terms of the normalization of the luminosity function does not depend on the low-luminosity cut-off L_1 and is defined only by the slope of the luminosity function and the high-luminosity cut-off (see discussion in Section 2.2.2). The total luminosity at the break, however, depends on the low-luminosity cut-off for a steep luminosity function with $\alpha > 2$:

$$L_{\rm tot,break} \approx \begin{cases} \dfrac{L_2}{2-\alpha} & \text{if } 1 < \alpha < 2, \\ \dfrac{L_2}{\alpha - 2} \left(\dfrac{L_2}{L_1} \right)^{\alpha-2} & \text{if } \alpha > 2, \end{cases} \tag{A14}$$

as the total luminosity for $\alpha > 2$ is defined by the sources near the low-luminosity cut-off.

APPENDIX B: VARIABILITY OF THE TOTAL EMISSION

In the linear regime of the $L_{\rm tot}$–n relation, the fractional rms^2 of the collective emission is inversely proportional to the number of sources:

$$\frac{{\rm rms}^2_{\rm tot}}{{\rm rms}^2_0} \approx \begin{cases} \dfrac{(2-\alpha)^2}{(1-\alpha)(3-\alpha)} \dfrac{1}{n} & \text{if } \alpha < 1, \\ \dfrac{(2-\alpha)^2}{(\alpha-1)(3-\alpha)} \left(\dfrac{L_2}{L_1} \right)^{\alpha-1} \dfrac{1}{n} & \text{if } 1 < \alpha < 2, \\ \dfrac{(2-\alpha)^2}{(\alpha-1)(3-\alpha)} \left(\dfrac{L_2}{L_1} \right)^{3-\alpha} \dfrac{1}{n} & \text{if } 2 < \alpha < 3, \\ \dfrac{(2-\alpha)^2}{(\alpha-1)(\alpha-3)} \dfrac{1}{n} & \text{if } \alpha > 3, \end{cases} \tag{B1}$$

or, equivalently, to their total luminosity:

$$\frac{{\rm rms}^2_{\rm tot}}{{\rm rms}^2_0} \approx \begin{cases} \dfrac{2-\alpha}{3-\alpha} \dfrac{L_2}{L_{\rm tot}} & \text{if } \alpha < 2, \\ \dfrac{\alpha-2}{3-\alpha} \dfrac{L_2}{L_{\rm tot}} \left(\dfrac{L_1}{L_2} \right)^{\alpha-2} & \text{if } 2 < \alpha < 3, \\ \dfrac{\alpha-2}{\alpha-3} \dfrac{L_1}{L_{\rm tot}} & \text{if } \alpha > 3. \end{cases} \tag{B2}$$

The above formulae are valid in the limit $L_1 \ll L_2$.

Similarly to the most probable value of the total luminosity (Appendix A), the fractional rms of the total emission can be approximately calculated substituting L_2 in equation (29) with some value $L_{\rm max} \leqslant L_2$. In principle, the probability distribution for the ${\rm rms}_{\rm tot}$ in equation (28) could be derived using the probability distribution for $L_{\rm max}$. The maximum of this probability distribution would give a sufficiently accurate approximation for ${\rm rms}_{\rm tot}/{\rm rms}_0$. However, for simplicity we use the value of $L_{\rm max}$ from equation (A7). In the limit $L_2 \gg L_1$ one finds for $1 < \alpha < 3$

$$\frac{{\rm rms}^2_{\rm tot}}{{\rm rms}^2_0} = \frac{(\alpha-2)^2}{(\alpha-1)(3-\alpha)} \frac{\xi^{3-\alpha}}{(\xi^{2-\alpha}-1)^2} \frac{1}{n} \tag{B3}$$

for $n < n_{\rm break}$, where ξ is defined by equation (A8) with substitution of equation (A9) for $1 < \alpha \lesssim 1.6$ and $n_{\rm break}$ is defined by equation (A12). For $\alpha < 1$ and $\alpha > 3$ there is no non-linear regime and the fractional ${\rm rms}_{\rm tot}$ obeys equations (B1) and (B2) for any n.

From equation (B3) one finds for $1 < \alpha < 2$

$$\frac{{\rm rms}^2_{\rm tot}}{{\rm rms}^2_0} = \frac{(\alpha-2)^2}{(3-\alpha)} \qquad n < n_{\rm break}, \tag{B4}$$

i.e. in the non-linear regime the fractional rms of the collective emission does not depend on the number of sources.

The accuracy of equations (B3) and (B4) is sufficiently good for $\alpha \gtrsim 1.5$ but deteriorates for smaller values of α. In the linear regime equations (B1) and (B2) are almost precise, their only approximation is in neglecting higher orders of L_1/L_2.

This paper has been typeset from a TEX/LATEX file prepared by the author.