Draft version April 28, 2024
Typeset using LaTeX **twocolumn** style in AASTeX631

Estimating the Likelihood of Hard X-ray Sources as LISA Progenitors

Sofía Martínez Fortis[1]

[1] *Vanderbilt University*

ABSTRACT

The rise of gravitational wave astronomy has provided astronomers the opportunity of studying a more complete history of merging compact objects, as such sources are able to be observed before, during, and after the merger by means of electromagnetic (EM) and gravitational waves (GWs). In this project, I study the likelihood of eROSITA's hard X-ray sources being LISA progenitors under the assumption that all of eROSITA's hard X-ray sources are binary AGN. I estimate a Gaussian PDF of the mass distribution assuming redshift, z=1 and estimate the probability of LISA observing eROSITA sources at varying SNRs for the cases where the mass ratio is q=0.1 and q=1. I find that it is highly probable that these AGN produce a minimum SNR of 8 and thus, are likely to be detectable by LISA. The case in which q=1 allows the sources to reach an SNR > 256 with a high probability while the q=0.1 scenario constrains the maximum plausible SNR threshold to around 64.

1. INTRODUCTION

The Laser Interferometer Space Antenna (LISA) is a space-based mission set to launch in 2034 that will observe gravitational waves of merging massive black hole binaries (MBHBs) with masses ranging from $\sim 10^4$ to $\sim 10^9$ (Amaro-Seoane et al. 2023). Prior to its launch, it is vital to find potential EM sources that could be LISA progenitors Bogdanović et al. (2022). The ROentgen Survey with an Imaging Telescope Array (eROSITA) is an all-sky survey of X-ray sources developed by the Max Planck Institute for Extraterrestrial Physics in Germany (Predehl et al. 2021). In this project, I claim that hard X-ray sources detected by eROSITA are plausible LISA progenitors. I support such claim by performing a statistical analysis of the mass distribution and how it compares to LISA's SNR sensibility curves. Specifically, I estimate the Gaussian probability density function (PDF) of the mass distribution at redshift $z = 1$ and approximate the likelihood of such sources reaching an SNR of 8, 64, 256, 1024, and 4096 for mass ratios, $q = 0.1$ and $q = 1$.

2. METHODOLOGY

2.1. *Data Selection*

The data used in this project is part of eROSITA's first data release, eRASS1. Specifically, I study the 719 hard X-ray (2.3 - 5 keV) point sources found in Merloni et al. (2024), which are assumed to be AGN binaries. This catalog only presents data of the western sky.

2.2. *Mass Estimation of X-Ray Sources*

The mass of each source is estimated at discrete redshifts, as the data only provides the X-ray flux. The redshifts used in this calculation range from $z = 0.01$ to $z = 10$ to encapsulate LISA's full redshift sensitivity.

First, the X-ray and bolometric luminosities (L_X and L_{bol}, respectively) are computed as done so in Runnoe et al. (2012),

$$L_X = \frac{4\pi d_l^2 f_X}{1+z} \tag{1}$$

where d_l is the luminosity distance, f_X is the X-ray flux, and z is the redshift.

Assuming the bolometric correction, $BC = 10$ as suggested by Vasudevan et al. (2009),

$$L_{bol} = \mathrm{BC} \cdot L_X \tag{2}$$

Assuming AGN accrete at the Eddington limit ($\lambda_{Edd} = 1$),

$$\begin{aligned} L_{Edd} &= \frac{L_{bol}}{\lambda_{Edd}} \\ M_{MBHB} &= \frac{L_{Edd}}{3.2 \times 10^4 L_\odot} \end{aligned} \tag{3}$$

where M_{MBHB} is in units of $\mathrm{M}_\odot$.

2.3. *Simulating LISA SNR Curves*

LISA-like waveforms, power spectral density functions, and signal-to-noise ratio are computed with the Python module, *pycbc*. The input parameters used to obtain the waveform are shown in the table below where f_{min} is the minimum frequency, f_{max} is the maximum

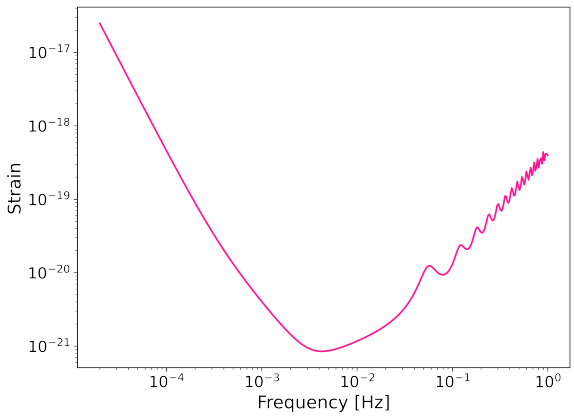


Figure 1. Noise curve for LISA

frequency, Δf is the step size in between measurements, and d is the distance at which $z = 0$. I assume the MBHBs have no spin.

f_{min} (Hz)	f_{max} (Hz)	Δf (Hz)	**d** (Mpc)
2×10^{-5}	0.5	10^{-4}	100

We calculate the horizon distance, d_h using the following formula,

$$d_h = d\frac{SNR}{SNR_{scale}} \tag{4}$$

where d is the distance at which z = 0, SNR is the signal-to-noise ratio, and SNR_{scale} is the target SNR. The corresponding redshift is computed using the python module, *astropy*, assuming cosmology parameters of those in Planck Collaboration et al. (2020). Finally, the source mass, M_{source}, is calculated as follows,

$$M_{source} = \frac{M_{det}}{z+1} \tag{5}$$

where M_{det} is the total mass detected.

3. RESULTS

The LISA SNR curves for mass ratios $q = 0.1$ and $q = 1$ as well as the distribution of eROSITA sources in redshift-mass space are plotted in Figure 2. Given that each of the 719 hard X-ray sources has a range of possible redshifts and masses as opposed to one value for each, I depicted them by means of a two-dimensional histogram to visualize the redshift-mass distribution instead of plotting 719 curves, which would have looked crowded and uninformative.

Each of the pink SNR curves in Figure 2 represents the redshift and mass boundaries necessary to achieve the desired SNR if observed by LISA. Therefore, any sources plotted below a particular SNR curve are able to produce and surpass the desired SNR. Therefore, it is not surprising that the $SNR = 8$ curves encapsulate more data than the $SNR = 4096$ curves since it is easier to survey sources that produce an SNR of at least 8 than an SNR of at least 4096.

Figure 3 depicts the mass distribution of eROSITA sources at $z = 1$ and marks the discrete masses that correspond to each LISA SNR curve at said discrete redshift. It is important to reiterate that the mass distribution is independent of the mass ratio and the mass threshold for each target SNR is the only parameter that varies with mass ratio.

4. DISCUSSION AND STATISTICAL ANALYSIS

The focus of the statistical analysis is to approximate a Gaussian distribution of the mass of eROSITA data at $z = 1$ and estimate the percentage of sources that are likely to be detected by LISA at SNRs greater than 8, 64, 256, 1024, or 4096 for the case of $q = 0.1$ and $q = 1$. Analyzing a one-dimensional distribution allows more intuitive interpretations of the data both visually and quantitatively. The bin-widths of all one-dimensional histograms in this project are optimized by "Freedman-Diaconis Rule",

$$\Delta b = \frac{2.7\sigma_G}{N^{1/3}} \tag{6}$$

where σ_G is the estimated uncertainty based on the interquartile range of the data and N is the number of data points.

At first glance, it is not clear whether or not the mass distribution of eROSITA data is Gaussian. Figures 2 and 3 show that the most populated bins are skewed right and the right tail falls off faster than the right tail. I computed the data's probability density function (PDF) as,

$$L(\mu, \sigma|x) = \prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) \tag{7}$$

where $\mu = 8.13$ is the sample mean and $\sigma = 23$ is the standard deviation. By visual speculation of Figure 4, the PDF fits the data mediocrely, as the curve's width is a bit wide and its peak is a bit short in comparison to the histogram.

4.1. *Optimizing Parameters*

The goal of this project is to compute a Gaussian PDF that accurately describes the data to further analyze the probability of observing an eROSITA source at $z = 1$ with LISA. I optimized the parameters μ, σ, and σ_G by means of 10,000 bootstrap samples with 500 draws each.

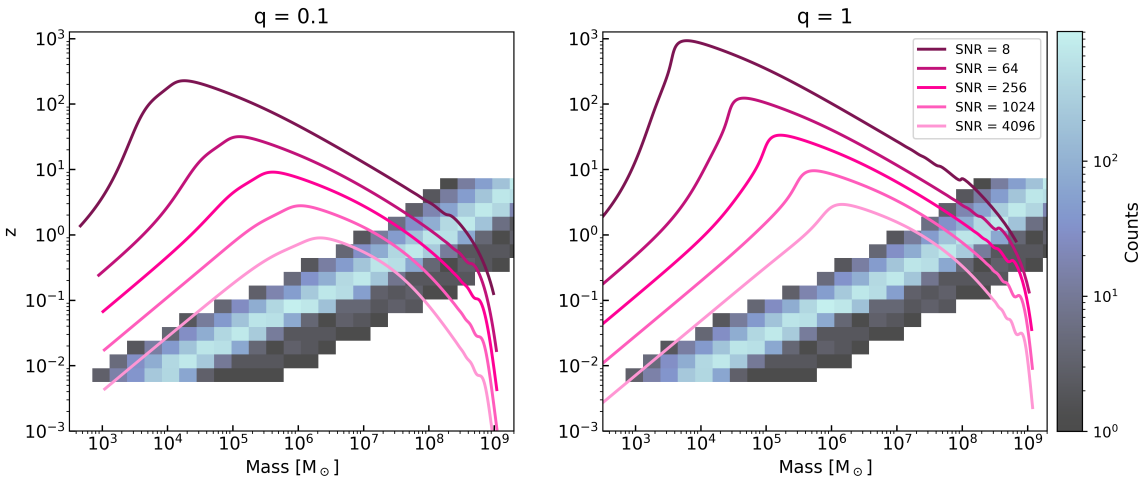


Figure 2. The two-dimensional histogram depicts the mass and redshift distribution of the eROSITA binaries. The pink curves illustrate the redshift necessary to achieve the desired SNR as a function of mass for a low mass ratio as shown in the left panel, and an equal mass ratio as depicted in the right panel. The SNR curves are the only differences between the panels, as the redshift-mass distribution of eROSITA sources is independent of mass ratio.

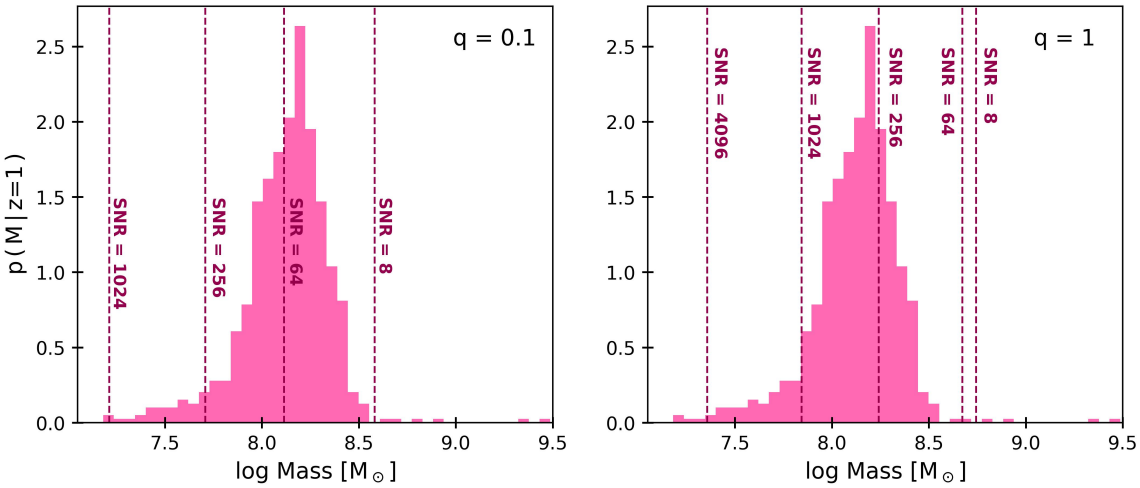


Figure 3. The histogram depicts the mass distribution at z=1. The burgundy dashed lines represent the mass threshold to achieve a given SNR. The case where the black holes have equal mass allows a higher SNR than in the low-mass ratio scenario.

The most probable values are $\mu = 8.13$, $\sigma = 0.23$, and σ_G.

The distribution of each parameter is depicted in Figure 5. Figure 6 compares two possible Gaussian distributions for the data. Both assume the mean μ value of the bootstrap distribution but differ in their uncertainties. The dashed function assumes the mean σ value as its uncertainty while the solid curve assumes the mean value of σ_G as its uncertainty. The latter PDF has a larger maximum likelihood and narrower spread that fits the steepness of the histogram better than the former since $\sigma_G < \sigma$ and σ_G is an appropriate estimate for the uncertainty of skewed data like the one in this project. Therefore, for the rest of the analysis, I assume the PDF with uncertainty σ_G to be the underlying distribution of the data.

4.2. *Testing Goodness of Fit*

I analyzed the assumed PDF to confirm that it fits the original data appropriately. First I resampled the data by randomly choosing 7,000 uniform draws and rejecting any that fall outside of the PDF (see panel (a) of Figure 7). To test that the distributions of the original and resampled data are similar, I performed a Kolmogorov-Smirnov (KS) test,

$$D = max|F_{og}(x) - F_{MC}(x)| \tag{8}$$

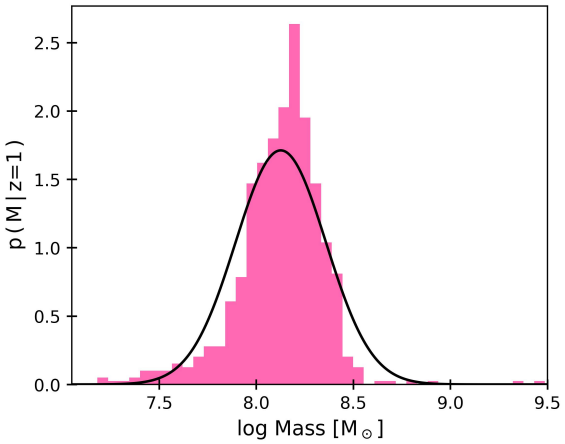


Figure 4. The pdf (black line) calculated from the sample mean and standard deviation overplotted on the mass distribution of the eROSITA binaries at z=1 (pink histogram) before resampling.

where $F_{og}(x)$ and $F_{MC}(x)$ are the empirical cumulative distribution functions (CDF) of the original 719 X-ray sources and the resampled data, respectively. From the KS test, the resulting p-value was 0.14. Assuming a null hypothesis in which there is no significant difference between distributions for $p \leq 0.05$, I accept the null. This means that there is not enough evidence between both samples that suggests that they were drawn from different distributions.

4.3. *Estimating LISA Detectability*

Finally, I assess the observability of eROSITA sources in LISA using the derived PDF. Figure 8 estimates the percentage of the mass distribution that can produce a minimum SNR of 8, 64, 256, 1024, and 4096 if observed by LISA. Such percentages are computed from the CDF value at each SNR threshold. In general, if the eROSITA sources are binary AGN of equal masses, they are likely to be observed by LISA at higher SNRs than if they have a mass ratio of $q = 0.1$.

4.3.1. *q = 0.1*

In the low mass ratio case, 99.4% of the mass distribution falls within the mass threshold at which an SNR of 8 can be achieved if observed by LISA. This means that if the eROSITA sources have mass ratio $q = 0.1$, accrete at the Eddington ratio, and are at redshift $z = 1$, there is about a 99.4% probability that they could produce the minimum SNR that LISA can observe (SNR=8), even if their exact mass is unknown. The probability of observing the AGN in this scenario at SNR=64 is about 47.2% and though it is less than half, it is large enough that the possibility of observing sources at such SNR should be considered. Conversely, it would be highly unlikely for a source in this scenario to produce an SNR of 256, as only 1.1% of the distribution falls within the range where a minimum SNR of 256 can be achieved. Therefore, SNR=64 would be an appropriate target to aim for if observing eROSITA sources in this scenario.

4.3.2. *q = 1*

In the case where the MBHs have equal masses, the detectability by LISA is a lot more optimistic than in the low mass ratio scenario. Assuming that the sources are at redshift $z = 1$ and are binary AGN accreting at the Eddington ratio, there is a 99.9% and a 73.7% probability that the sources can be observed at an SNR of 64 and 256, respectively. The SNR=1024 threshold does not enclose a significant portion of the distribution and therefore, an SNR of 256 would be an appropriate target SNR if observing this case with LISA.

4.4. *Reliability of Results*

The findings of this project are meant to provide astronomers with a rough estimate of whether or not hard X-ray sources like those in eROSITA make appropriate candidates. However, there is not enough evidence to support the specific quantitative findings of this project, given the amount of assumptions that were made. The mass-redshift relationship derived in this project assumes that all hard X-ray sources from eRASS1 are inspiraling binary AGN that accrete at the Eddington ratio. On the LISA portion, this work assumes that the MBHBs are non-spinning, are not inclined, and are 100 Mpc away. Finally, the underlying distribution of the mass at $z = 1$ is estimated under the assumption that it is Gaussian. Nevertheless, the results presented in this paper support the concept of X-ray sources like those detected by eROSITA being promising candidates for multi-messenger observations with LISA.

5. SUMMARY AND CONCLUSIONS

In this project, I aimed to show LISA's ability to detect eROSITA hard X-ray sources in a statistical manner. For simplicity, I focused on the one-dimensional mass distribution of eROSITA sources at $z = 1$ instead of the full two-dimensional mass-redshift distribution. Given the uncertain nature of the source mass and SNR calculations, I assumed a Gaussian distribution of the mass to estimate the corresponding PDF. I optimized the mean (μ) and uncertainty parameters (σ, σ_G) of the PDF using bootstrap sampling. I found that the most probable value of the mean was $\mu = 8.13$ and the mean σ_G value, $\sigma_G = 0.18$ described the spread of the distribution appropriately. Then, I determined the goodness-of-fit of the adopted PDF by resampling the

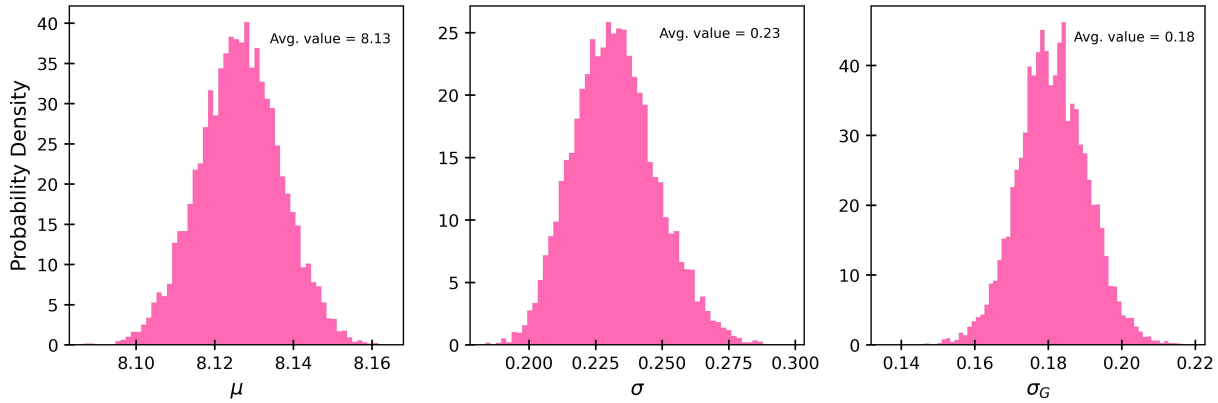


Figure 5. The distribution of μ, σ, and σ_G after bootstrap sampling of 10,000 samples with 500 draws each.

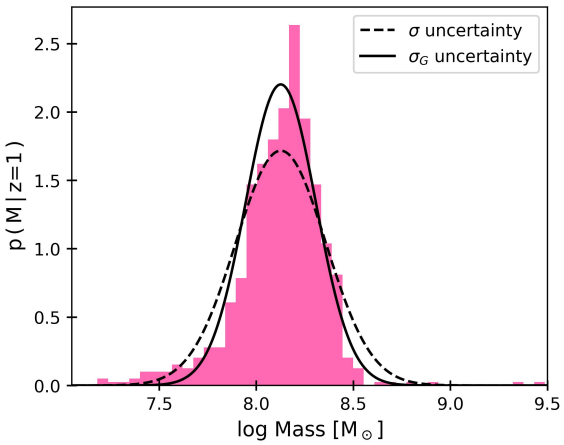


Figure 6. Two possible Gaussian PDFs for the MBHB mass at z=1 are shown–one with the mean σ (dashed line) and the other with the mean σ_G (solid line) as the uncertainty. Both distributions assume $\mu = 8.13$.

original data points, rejecting those that fell outside of the PDF, and comparing the distribution of both the old and resampled data by means of a KS test. I found that the p-value was 0.14 which did not follow the condition $p \leq 0.05$ required to reject the null hypothesis that the two samples were drawn from the same distribution. Thus, the adopted Gaussian PDF proved to be an appropriate estimate for the underlying distribution of the mass.

Finally, I studied the likelihood of observing such X-ray sources with LISA at varying SNR thresholds for the cases in which the mass ratio is $q = 0.1$ and $q = 1$. I found that in the case with equal BH masses, achieving a minimum SNR=256 was plausible, as 73.7% of the mass distribution is able to meet this requirement. The case where $q = 0.1$ showed less promising results, as only 47.2% of the mass distribution would be able to produce an SNR of at least 64. Nevertheless, in both scenarios over 99% of the data have the necessary mass to achieve

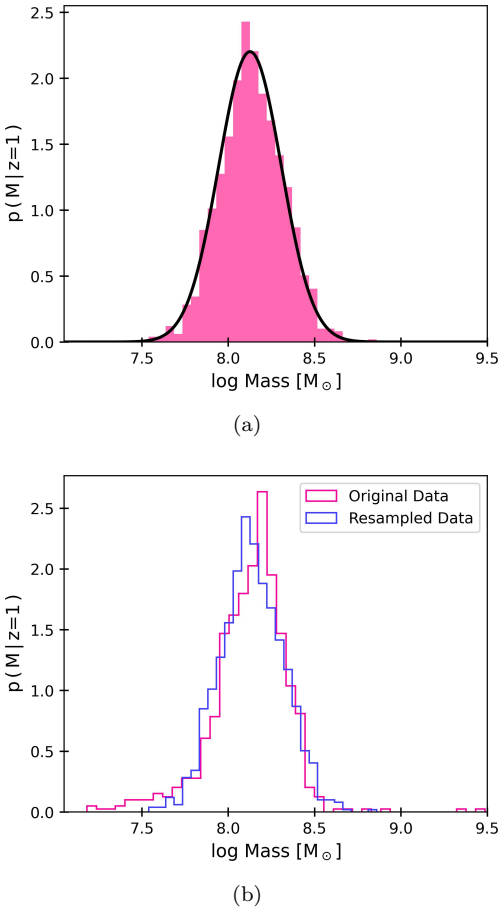
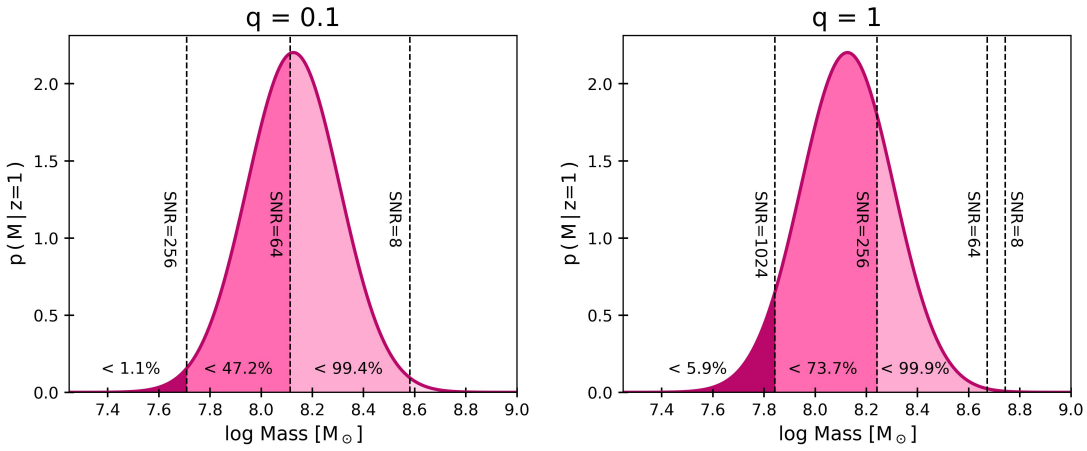


Figure 7. The data is resampled by means of Monte Carlo rejection sampling to better fit the adopted PDF. (a) compares the resampled histogram to the adopted PDF. (b) compares the original mass distribution for z=1 to the resampled one.

SNR=8, the minimum SNR that LISA may be able to observe.

6. ACKNOWLEDGEMENTS



(a) (b)

Figure 8. (a) and (b) depict the approximate mass percentage that can achieve each SNR threshold or higher.

I would like to thank Azeem Bari for filtering the eRASS1 catalog for hard X-ray sources and for computing their estimated mass. Additionally, I thank Professors Jessie Runnoe and Karan Jani for providing the necessary background in regard to MBHBs, AGN, eROSITA, and LISA which allowed me to develop this project. Finally, I am thankful for Professor Stephen Taylor's pedagogical enthusiasm, admirable patience regarding my late assignments, and for teaching me that everything is a Gaussian if you try hard enough.

REFERENCES

Amaro-Seoane, P., Andrews, J., Arca Sedda, M., et al. 2023, Living Reviews in Relativity, 26, doi: 10.1007/s41114-022-00041-y

Bogdanović, T., Miller, M. C., & Blecha, L. 2022, Living Reviews in Relativity, 25, doi: 10.1007/s41114-022-00037-8

Merloni, A., Lamer, G., Liu, T., et al. 2024, A&A, 682, A34, doi: 10.1051/0004-6361/202347165

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6, doi: 10.1051/0004-6361/201833910

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, Astronomy amp; Astrophysics, 647, A1, doi: 10.1051/0004-6361/202039313

Runnoe, J. C., Brotherton, M. S., & Shang, Z. 2012, Monthly Notices of the Royal Astronomical Society, 426, 2677, doi: 10.1111/j.1365-2966.2012.21644.x

Vasudevan, R. V., Mushotzky, R. F., Winter, L. M., & Fabian, A. C. 2009, Monthly Notices of the Royal Astronomical Society, 399, 1553, doi: 10.1111/j.1365-2966.2009.15371.x