Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **488,** 183–190 (2019)
Advance Access publication 2019 June 21

doi:10.1093/mnras/stz1627

The luminosity function of quasars by the Principle of Maximum Entropy

Alexandre Andrei,[1,2,3★] Bruno Coelho,[4★] Leandro L S Guedes[5★] and Alexandre Lyra[1]

[1]*Universidade Federal do Rio de Janeiro, Observatório do Valongo, Ladeira Pedro Antonio, 43, Rio de Janeiro, RJ, CEP 20080-090, Brazil*
[2]*Observatório Nacional, MCTIC, Rua Gal. José Cristino 77, Rio de Janeiro, CEP 20921-400, Brazil*
[3]*SYRTE, Observatoire de Paris, 61 Avenue de l'Observatoire, F-75014 Paris, France*
[4]*Instituto de Telecomunicações, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal*
[5]*Planetarium Foundation of the City of Rio de Janeiro, Rua Vice-Governador Rúbens Berardo 100 – Gávea, Rio de Janeiro – RJ, CEP 22451-070, Brazil*

Accepted 2019 June 7. Received 2019 May 31; in original form 2018 November 30

ABSTRACT

We propose a different way to obtain the distribution of the luminosity function of quasars by using the Principle of Maximum Entropy. The input data come from Richard et al 2006 quasar counts, extending up to redshift 5 and limited from apparent magnitude $i = 15$–19.1 at $z \lesssim 3$ to $i = 20.2$ for $z \gtrsim 3$. Using only few initial data points, the principle allows us to estimate probabilities and hence that luminosity curve. We carry out statistical tests to evaluate our results. The resulting luminosity function compares well to earlier determinations, and our results remain consistent either when the amount or choice of sampled sources is unbiasedly altered. Besides this, we estimate the distribution of the luminosity function for redshifts in which there is only observational data in the vicinity.

Key words: methods: data analysis, principle of maximum entropy – galaxies: luminosity function, mass function – quasars: general.

1 INTRODUCTION

The quasar luminosity function gives a measure for the bidimensional distribution of quasars in luminosity and redshift. Fundamentally, it indicates that the Universe is not in a stationary state. As consequence, it requires the due interpretation before using quasars to determine cosmological parameters, but at the same time it informs about the evolution of quasars themselves, and the changing content of the space intervening between distant quasars and the observer. The function usually describing the quasar luminosity function, as a function of redshift and absolute luminosity, basically starts from the modulus distance formulae and incorporates several corrections, to accommodate line emission, the expanding universe scale of distance, the intrinsic dependency of quasar light emission on wavelength, terms of self and media absorption, etc. The result is an empirical description, which exponents and coefficients are adjusted to each sample examined. It is interesting thus to build an independent function, able to describe the quasar luminosity function in a simpler form and from different physical principles. Although by necessity also incorporating the astrophysical and cosmological assumptions, an alternative, simpler form for the quasar luminosity function can be derived from the statistical mechanics methods.

The concept of entropy, since Clausius, became part of thermodynamics. In addition, it also became part of statistical mechanics. The study of systems in equilibrium and out of equilibrium is closely related to the notions of entropy as well as its production. There is a vast bibliography about it with warm discussions. We can cite three important related principles: Ziegler's *maximum entropy production principle* (e.g. Ziegler 1983; Ziegler 1987, see also Dewar 2005); Prigogine's *minimum entropy production principle* (Prigogine 1967; Prigogine 1978; Kondepudi & Prigogine 1999); [1] and the Maximum Entropy Principle (MaxEnt) (Jaynes 1957). This paper employs the last one.

Derivations of the first two principles from MaxEnt can be found in literature, as seen in Martyushev et al (2006). In that review, the authors make a very interesting description of the MaxEnt focusing on the production of entropy. Other authors emphasize that Jaynes's MaxEnt formulation of statistical mechanics provides a theoretical basis for Maximum Entropy Production Principle (Dewar & Maritan 2014). The applications of the MaxEnt are many. We'll see below related issues and discuss how they are connected to the focus of our treatment, that aims to find the distribution of the luminosity function of quasars.

Despite this vast reach there are authors who restrict the MaxEnt applications (see e.g. Shimony 1985, and references therein).

Some of these critiques were addressed by Jaynes himself (Jaynes 1989). In this paper, Jaynes also provides a fairly complete description of MaxEnt from its roots to its implications. On the other hand, we cannot fail to mention that there are also papers written exactly in defence of Jaynes's Principle as in Tikochinsky, Tishby &

★ E-mail: alexandr@astro.ufrj.br (AA); brunodfcoelho@gmail.com (BC); leandrolsguedes@gmail.com (LLSG)

[1]This principle is the subject of a specific work in Jaynes (1980).

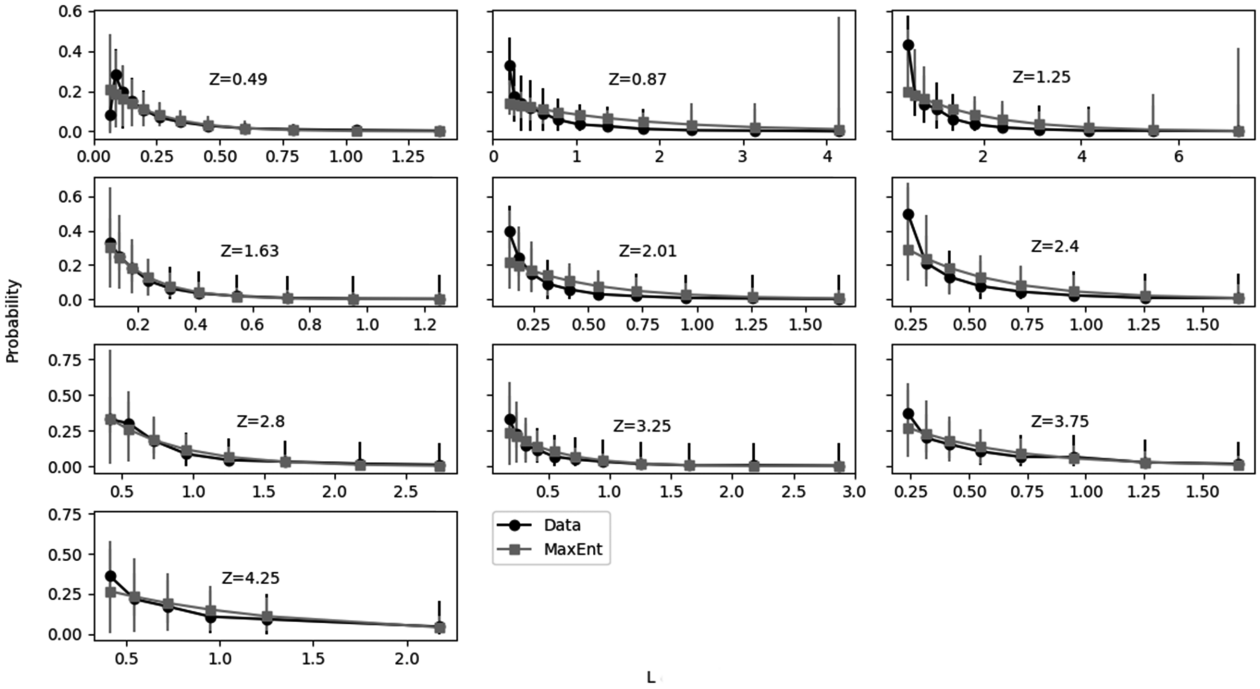

Figure 1. Comparison between probabilities calculated from the MaxEnt method and from the control results. Plots show luminosities on the horizontal axis and probabilities on the vertical ones. Values at horizontal axis should be multiplied by 10^{-31} (erg s^{-1} Hz^{-1}) at $Z = 0.49$, 0.87, 1.25 and by 10^{-32} (erg s^{-1} Hz^{-1}) at the others.

Table 1. Statistical tests comparing the MaxEnt and control results luminosity curves.

	Spearman		*F*-test	Student-*t*	
	ρ	*P*-value		*T*-status	*P*-value
0.49	0.93	1.17×10^{-5}	1.32	$\ll 10^{-99}$	1
0.87	1	$\ll 10^{-99}$	4.40	8.88×10^{-16}	0.99
1.25	1	$\ll 10^{-99}$	3.26	$\ll 10^{-99}$	1
1.63	0.99	6.65×10^{-64}	1.10	$\ll 10^{-99}$	1
2.01	0.99	6.65×10^{-64}	2.75	$\ll 10^{-99}$	1
2.4	1	$\ll 10^{-99}$	2.53	1.99×10^{-16}	0.99
2.8	1	$\ll 10^{-99}$	1.12	2.21×10^{-16}	0.99
3.25	0.99	3.76×10^{-9}	1.48	3.35×10^{-16}	0.99
3.75	1	$\ll 10^{-99}$	1.50	$\ll 10^{-99}$	1
4.25	1	$\ll 10^{-99}$	1.87	$\ll 10^{-99}$	1

Levine (1984), stating with: 'The only consistent algorithm is one that leads to the distribution of Maximum entropy subject to constraints given.'. There are other papers with very interesting critiques that bring out points for and against MaxEnt and provide quite compelling references on the subject, like in the appendix A of Pontzen & Governato (2013), where the authors sketch Jaynes's reasoning, 'that the maximization of entropy subject to certain constraints is equivalent to testing whether these constraints encapsulate later the physics of the situation...', and the use of the method to derive the phase-space distribution of a virialized dark-matter halo.

In addition, there are several other areas in physics and astrophysics where it can be applied. Some examples are, in spectral analysis (Ables 1974), where 'the method produces superior spectral representations when compared with more traditional methods...' as well as a powerful technique of image reconstruction (Skilling & Bryan 1984), in the same paper other applications of MaxEnt in astronomy can be found. In Gull & Daniell (1978), MaxEnt is applied in radio and X-ray astronomy. It is interesting that the method is also applied in X-ray tomographic image reconstruction and restoration (Mohammad & Djafari 1989). In the case of astrophysics and cosmology, we see papers where the dark energy equation of state $w(z)$ is reconstructed using the MaxEnt (Zunckel & Trotta 2007).

In gravitation, with the confirmation in 2016 of the existence of the gravitational waves predicted by A. Einstein, the study of the black holes assumes still greater importance. The earliest detections were precisely on collisions of black holes (Abbott et al. 2016). The traditional second law of thermodynamics was modified into a generalized second law for the study of black holes (Bekenstein 1974). The Jaynes's method of maximum entropy was also used by Bekenstein to determine the probability distribution for a system containing a Kerr black hole (Bekenstein 1975).

This paper presents a new approach to find the distribution of probabilities of the luminosity function using the MaxEnt technique. Even with some criticisms like those cited above, we believe that the MaxEnt is extremely useful to be applied when we have partial

Table 2. Statistical tests comparing the error budgets over the MaxEnt and control result luminosity curves.

	Spearman		*F*-test	Student-*t*	
	ρ	*P*-value		*T*-status	*P*-value
0.49	0.77	1.02×10^{-5}	0.70	1.48	0.15
0.87	0.48	0.02	0.19	0.71	0.48
1.25	0.63	0.00	0.22	0.11	0.91
1.63	0.60	0.00	0.29	0.44	0.67
2.01	0.53	0.02	0.48	0.03	0.97
2.4	0.46	0.07	0.29	0.50	0.62
2.8	0.60	0.01	0.16	0.55	0.58
3.25	0.56	0.01	0.41	0.45	0.65
3.75	0.65	0.01	0.41	0.45	0.65
4.25	0.58	0.05	0.24	1.49	0.15

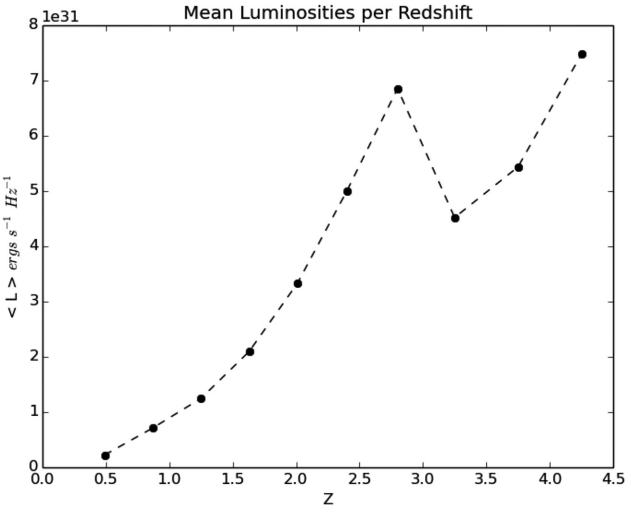

Figure 2. Mean luminosity in bins of redshift from $z = 0.5$ to $z = 4.25$ derived from SDSS-DR3 as in Richards et al. (2006), not taking into account any corrections for the Malmquist bias.

information about a certain system. So this principle allows us to know accurate probabilities (see formula 4) from a small data set. Although the number of known quasars is constantly increasing, to get perhaps to a million known objects in the next decade, small subsamples are useful and had not been yet designed by lack of elements. On one hand, the quasar zoo is also growing, different types of active galaxies conceivably exhibiting luminosity functions peculiar by a certain degree. On the other hand, the capability of mapping in detail particular thin slices of the universe in redshift is long sought, none the less to better define the complex form of the luminosity function. Finally, it is important to be able to drawn different samples of a large data set for sanity check control. Is this paper, we will explore such capability of the MaxEnt description of the luminosity distribution.

Since quasar discovery (Matthews & Sandage 1963; Schmidt 1963), their energy output and magnitude have been object of much observation and increasingly complex theories. Conversely, that information became much used for studies as surrounding host galaxies, gravitational lenses, *in situ* and intergalactic absorption, up to the cosmological scale of distances in an expanding universe. The so-called luminosity function is all important to make sense of such extraordinary energy output and to those astrophysical quantities from it derived. The evolution of the quasar luminosity function with redshift is an important observational tool that allows us to put constraints on the formation and growth history of supermassive black holes and their co-evolution with host galaxies. It also gives us a measure of the contribution of quasars in the cosmological reionization of the Universe. For all these, the study of the quasar luminosity function has received the attention of several works (e.g. Richards et al. 2006; Masters et al. 2012; Ross et al. 2013; Manti et al. 2017).

So, among some successful applications of MaxEnt in astrophysics, we are going now to explore a new one, in the study of the quasar luminosity function.

The Sloan Digital Sky Survey (SDSS)[2] provided observations of quasars in different redshifts, being responsible for the identification of the vast majority of the known quasi-stellar objects (QSOs; Pâris et al. 2018). However, there are observational limitations to the effect that one can ask: what would be the quasar distribution on each redshift slice if we could consider unobserved magnitudes? These observational limitations must be taken into account when computing the quasar luminosity function, however this does not constitute the aim of this work. So we are going to use already corrected counts of QSOs computed by Richards et al. (2006) in their study of the quasar luminosity function. We show here that the MaxEnt can provide a good distribution of probabilities for the luminosity function from few values of a limited sample in each redshift.

The luminosity function provides the density distribution of classes of objects, per unit volume and assuming a statistically complete sample. In the case of quasars, this indicates more or less probable scenarios for their formation and evolution, as well as their relationship with the host galaxy. Quasars have been found out by several projects, chiefly the SDSS, relying on different strategies to single them out from the more numerous contaminants of other celestial bodies. The European Space Agancy cornerstone mission *Gaia* combines the recognition of such known quasars, with microarcsecond determination of proper motions over 5 yr, therefore providing direct means to cleanse away the intruding false positives, as nearby red dwarfs. On top of it, *Gaia* will use a neural network strategy leading to autonomous recognition of quasars. Combined to the all-sky repeated sweeping of objects up to magnitude $i \sim 22$, it will produce an unprecedent complete sample of quasars. Therefore to establish an alternative, independent, and physics, robust method of tracing the quasar luminosity function affords a strong way of checking upon and getting feedback from the usual Schechter-based determination. In short, the motivations for these studies are threefold, an independent study of the luminosity

[2] http://www.sdss.org/

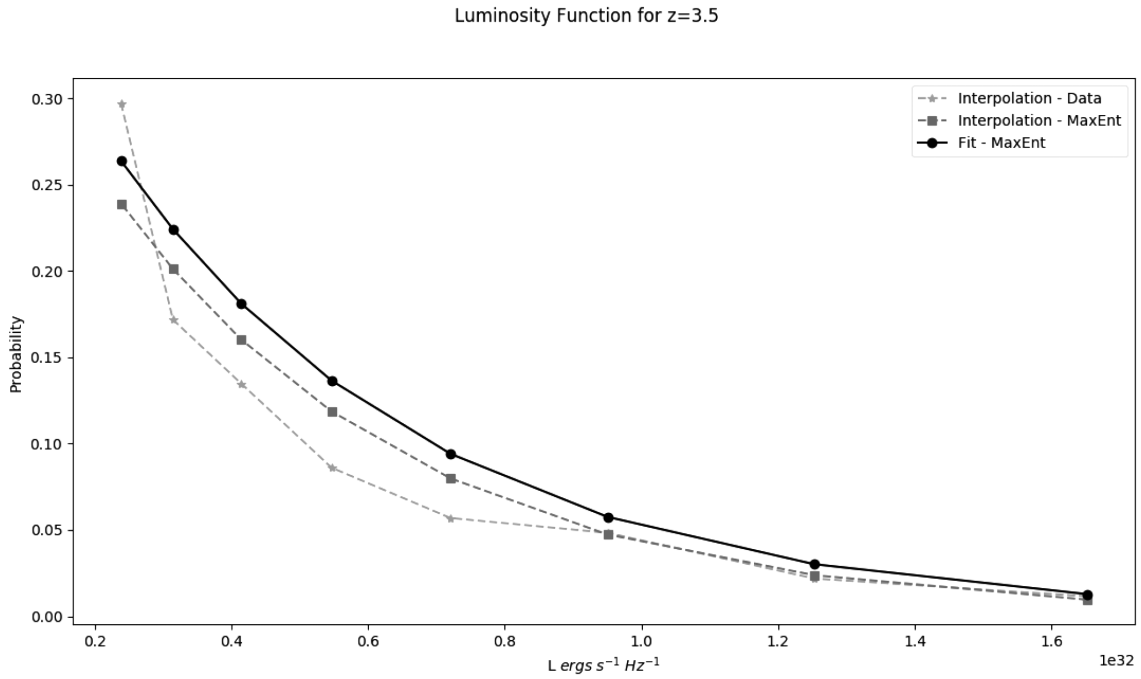

Figure 3. Luminosity function obtained from MaxEnt using the fitting curve to $z = 3.5$.

function on the quasar population in the Sloan Digital Sky Survey Data Release 3 (SDSS-DR3), the development of an independent tool for determining the luminosity function based on maximum entropy physics, and it is a comparative assessment on a limited sample with views to application on the all sky, statistically complete sample of quasars in final *Gaia* catalogue.

Elseways, the definition of the quasar luminosity function have so far been done using a modified template of the Schechter exponential for galaxies. Such a description although well adapted to the somehow simpler quasar case, since it is in practice free from the surface-brightness issue, limits the reliability of the astrophysical and cosmological interpretation of the luminosity function. To mention a few, it is known that the shape and turnover of the luminosity function would favour either models for the growth of the super massive black hole from mergers or by inflow and host-galaxy instabilities. The bright end of the luminosity function can favour intrinsic properties about which time black holes are increasing in mass rapidly, whereas the faintest end would indicate about the length of time quasars spend at relatively low accretion rates.

The remainder of this work is organized as follows. In Section 2, we will briefly review the MaxEnt method. With this, we establish our main formula, the equation (4), which defines from MaxEnt the probability of the luminosity function. In Section 3, we will summarize our technique to determine the luminosity function of quasars, and we show the details of how the Lagrange multipliers were calculated for the studied cases; in addition, we will show the comparison between our result by MaxEnt and the Schechter-based Richards et al. (2006) one.

In Section 4, we will describe the statistical tests we use. In Section 5, from few observational data in particular redshifts, we will make a prediction of the probability density function (PDF) for in-between redshifts; that is, we will estimate the distribution of the luminosity function. Finally, a summary discussion and conclusions are presented in the last section.

2 THE JAYNES APPROACH TO MAXIMUM ENTROPY PRINCIPLE

We can sum up the Maximum Entropy Principle, as we shall see in the sequel.[3] As there is a vast bibliography regarding this principle, we will only make a brief account.

Initially, we assume that a quantity x can have the discrete values $x_i (i = 1, 2, ..., n)$, but we do not know the corresponding probabilities p_i. All we have is the expectation value of the function $f(x)$,

$$\langle f(x) \rangle = \sum_{i=1}^{n} p_i f(x_i). \tag{1}$$

Based on this information, how can we obtain the expectation value of another function of the system $g(x)$? Jaynes responds to this apparently insoluble question. The given information is insufficient to determine the probabilities p_i. The equation (1) and the normalization condition

$$\sum p_i = 1 \tag{2}$$

would have to be supplemented by $(n - 2)$ more conditions before $g(x)$ could be found.

In order to find a solution to this problem, Jaynes's method uses the following expression for entropy:

$$H(p_1, p_2, ..., p_n) = -k \sum_i p_i \ln p_i, \tag{3}$$

where k is a positive constant. Since H is just the expression for entropy as found in statistical mechanics, it will be called the

[3]Here, we will follow Jaynes (1957).

'entropy of the probability distribution p_i'. The entropy H, given in (3) is maximized subject to the constraints (1) and (2).

In order to achieve a final expression for the probability of x_i, we use the method of Lagrangian multipliers, usually noted by λ and μ, where λ is associated with the normalization equation, i.e. the equation (2) and μ is associated with the equation of the expectation value (1). With this methodology, we obtain the probability

$$p_i = e^{-\lambda - \mu f(x_i)}. \quad (4)$$

This formula gives an important expression, which can be associated with the function of the luminosity distribution of the objects to which we wish to estimate the distribution, and the method used in its determination is called the Maximum Entropy Principle. See the complete development from data to Lagrange multipliers at Appendix B.

3 THE LUMINOSITY FUNCTION OF QSO(S) FROM MAXENT

To summarize what will be done next, from MaxEnt we will determine the distribution of the luminosity function of the quasars in a certain redshift z_k, by using the probability distribution (4). Notice that the strong energy released by quasars make possible to observe them from the nearby Universe at least up to redshifts greater than 7 (e.g. Bañados et al. 2018). This large range of distance, hence an evolving luminosity function, allowed us to inspect how consistent are the predictions from MaxEnt, and compare the results against those originally derived from the same observed data, used here as control result.

In the MaxEnt methodology, for the consistency of the principle, the strongest symmetry that we could have 'a priori' would be the uniform distribution, but this is not the case. We know that if we have a single constraint, that is associated with normalization $\sum_{i=1}^{n} p_i = 1$, we get exactly for $n = N \rightarrow p_i = 1/N$ or the uniform distribution. The other constraint, associated with equation (1), breaks this symmetry. Let us also remember that as it is well placed in Caticha & Preuss (2004): 'The method of maximum entropy (ME) is designed for updating from a prior probability distribution to a posterior distribution when the information to be processed takes the form of a constraint ... '. Then, we assume that we can extract a certain expected value obtained through some luminosity values provided by the system observations, which obviously have the uniformity between all values broken. These values are randomly chosen, and under these conditions we will apply MaxEnt with their two constraints: (1) and (2). This is the central point of the methodology, namely that from just some values a strong estimate of the luminosity function of the distribution of all values in this redshift can be made.[4]

For the present quasar luminosity function derivation by MaxEnt, we have tested different sets from the whole of the initial data, seeing in every case great accordance between the Luminosity curve from MaxEnt and the control result. In order to analyse the most realistic scenario, the one for which the sample is small and, thus, not necessarily containing a perfect representation of the data population, we choose to analyse here the results from random initial data. We have picked up just three luminosities in each redshift as initial data.

The starting point of using MaxEnt is the calculation of μ and λ from the equations (1) and (2) (see details in Appendix B). From a certain redshift, the mean value to be used in the Lagrange multiplier method is calculated from three luminosities randomly chosen, to each of which is assigned the corrected number of quasars in that luminosity bin after applying the selection function of Richards et al. (2006, table 6, p. 2782). Those values will be used to calculate the weighted mean luminosity $\langle L_z \rangle$, which is the value to be used in the equation (1). The other Lagrange multiplier comes from the normalization of the probability, or, $\sum_{i=1}^{n} p_i = 1$.

Errors have been calculated using a bootstrap method. In each case, three random luminosities were drawn 200 times and the mean value used to find a different λ and μ that, applied to original data, gave us a different set of points. The extreme values stand as the upper and lower limits of the error bars to the results from the principle. Likewise the errors on the control result were calculated using probabilities from bootstrap draws.

Verifying our assumptions, the calculated probabilities by MaxEnt and the ones of the control results show similar behaviour.

For each redshift, the complete table leading to the control result is in Appendix A, Table A1, and the three ones randomly chosen in each redshift are on the lines indicated in bold at the first column.

The conversion from calibrated magnitudes to luminosities was done using the following relation:

$$L = 10^{\frac{-(M_i+48,6)}{2,5}} 4\pi \left(3.0857 \times 10^{19}\right)^2, \quad (5)$$

where L (in erg s^{-1} Hz^{-1}) is the luminosity and M_i the magnitude.

The curves obtained for each redshift are shown in Fig. 1. We can see clearly that a correspondence is found at the sampled redshifts, within the error bars, between the MaxEnt results and those for the control.

4 STATISTICAL TESTS

As discussed in the previous paragraphs, and detailed in Appendix B, the MaxEnt approach, from robust yet simple physical principles and computational algorithms, delivers a statistical probability distribution of the luminosity function that is cosmologically plausible, vis-à-vis the literature on the subject. The magnitude and redshift data used for that is taken from the SDSS project. It is natural thus that the outcomes from the luminosity function here obtained shall be compared with those from the SDSS analysis.

At the start of the current application of the MaxEnt principle to derive the quasar luminosity function, several approaches were used. Choosing by hand representative data, choosing data from quartiles of the distribution, and picking up the extreme and mean values. The outcomes were always concurrent (they are available under request), what served as sanity check, as well as gave us ground to adopt the random draws finally used. The plots in Fig. 1 are compelling to show the agreement between the two luminosity function statistical probability distributions. Such agreement can be quantified. Table 1 shows results of the statistical tests comparing the two distributions of probabilities, the one from MaxEnt and the one from the control results, at each redshift.

As indicated in Section 3 a minimal number of points were randomly drawn from the data. Using only these few data points, MaxEnt can provide us an estimate luminosity function to be compared with the luminosity function obtained from the control results.

[4]One interesting question posed by Jaynes is 'generating paradoxes in the case of continuously variable random quantities, since intuitive notions of "equally possible" are altered by a change of variables' (Jaynes 1957, p. 622).

The results are compared by verifying the mutual correlation. The Spearman's correlation test is used because of the small number of chosen points, as well to not assume their normal distribution. The ρ is quite close to unity. Notice however that although the luminosity function is best represented as an exponential progression, the pair of points of the two compared distributions are not necessarily so, thus we have also used the *F*-test and the Student's *T*-test because these are non-parametric tests.

Since the shape of the curves is obviously similar but not the error bars, while the number of points is small, the *F*-test for variances is advisable. The table of the *F*-distribution indicates that the null hypothesis (no difference) must be accepted to a large degree of statistical certainty, with two exceptions, out of the limit redshift. Those exceptions lie at $z = 0.87$ and $z = 1.25$, for which the null hypothesis certainty is mediocre. In both cases, that befalls upon the large error bars seeing at the one brightest luminosities. The *F*-test without those points give us results 1.64 and 0.92, respectively, that take us back to a null hypothesis scenario.

On views of the outturn of the correlation and variance tests pointing to the agreement of the MaxEnt and control results, the two-samples Student's *T*-test is next justified. On this one, as Table 1 shows, in all cases – even for the troublesome redshifts as detected in the previous tests – the null-hypothesis on the means cannot be rejected for usual statistical standards.

Table 1 brings the three statistics for the distributions. On Table 2, instead, the same statistical tests are applied to compare their error budgets. Notice, at start, that the error bars are asymmetrical, and therefore up and down pairs are formed. The correlations are poor, though they undeniably exist. On the other hand, the *F*-test and *T*-test for the errors show the MaxEnt method and the control results faring quite alike also in this respect. We thus can further conclude for the independence of the methods, but similar efficiency.

5 ESTIMATION OF THE LUMINOSITY FUNCTION FOR OTHER REDSHIFTS

In this section, we use the MaxEnt luminosity function presented in this paper to investigate the outcomes for a redshift in which we suppose that data exist only in its vicinity.

For this simulation, the redshift $z = 3.5$ is chosen. As shown in Fig. 2 at this redshift the luminosity L seems to increase again after a drop between $z = 2.75$ and $z = 3.25$, at the same time there are enough input data and good results for the neighbour redshifts. From those the mean value $\langle L_{z=3.5} \rangle$ is interpolated, and next we will obtain by MaxEnt the distribution of L for the redshift 3.5.

In practice, we start from the same set of data used before, from Richards et al. (2006), plotting all the available redshifts with respective mean luminosities. Then the curve of best fitting to the observational data is obtained, and from this fitting curve we associate a mean luminosity with the redshift aimed at. Next, in order to procure the Lagrange multipliers μ and λ a set of observed luminosities is demanded. Those were picked up at random from the luminosities actually present for the neighbour redshifts.

The point now is to verify whether using this quite arbitrary choice the MaxEnt formulation is capable to issue a credible luminosity function. We thus compare the MaxEnt formulation results to a direct interpolation of the control results and of the MaxEnt results themselves (both depicted in Fig. 1). Fig. 3 shows these three results. It is seen that the MaxEnt formulation based on neighbour data gives a result comparable to the direct interpolation results, but at the same time, it delivers a smoother curve.

This type of situation occurs frequently in astrophysics, and MaxEnt demonstrates here to be a very useful tool to estimate values, what later can be tested later as more data become available.

6 CONCLUSIONS

The quasar luminosity function is intended as a measure of the actual distribution of quasars in luminosity and redshift. For that observational, astrophysical, and cosmological restricting factors must be accounted for and often different surveys must be combined, before a complete population is inferred. That satisfied, most quasar luminosity functions available in the literature are represented either by a double power-law regimen or by a modified Schechter function. The adjustments are semi-empirical, having as usual parameters a normalization factor, a break magnitude, a reference redshift, and bright and faint ends slopes.

By contrast, the MaxEnt method, on top of being quite simple to handle, offers three strong features. First, it represents a physically distinct approach, thus bringing the known benefits of different bias, limitations, and systematics. Secondly, because it is purely statistical, it depends of less astrophysical and cosmological assumptions, in special the key ones break magnitude and reference redshift. Thirdly, a hallmark of MaxEnt is to deliver trustful conclusions from small samples. This last quality is particularly suited to deal with limited dedicated surveys, as well as to piece off portions of the luminosity function without further requirements to the mathematical representation of the function itself. By the same token, it is suited to try out luminosity functions for putative new classes of quasars and their location, either within large clusters or relatively isolated.

In this pioneer derivation, we took the SDSS-DR3 quasar population and the normalization made by Richards et al. (2006) there in. The luminosity functions and corresponding curves were used here as control results. The comparisons hold very well, being practically immaterial whether the whole luminosity population or samples as small as three random elements were used.

As Jaynes has stated, that MaxEnt is the generalization of the Principle of Insufficient Reason. In our case, we show that little information of the system (quasar luminosities) gave us consistent results. In so, it is an effective way of practical generalization. As a result, the Lagrange multipliers behaved in a stable manner, enabling to use bootstrapping for determination of errors. The aspect of updating the knowledge when of the outcome of a much larger data set, as expected from *Gaia*, is foreseen to be coherently accommodated, as well as to investigate piecemeal the luminosity function.

ACKNOWLEDGEMENTS

AL thanks the colleagues at the Valongo Observatory, H. M. Boechat Roberty and M. Assafin, for suggestions in the beginning of this work. AA thanks CNPq Grant Bolsa de Produtividade em Pesquisa 306775/2018. BC acknowledges support from the Advanced EU Network of E-infrastructures for Astronomy with SKA (AENEAS), funded by the European Commission Framework Programme Horizon 2020 RIA under grant agreement no. 731016 and from the ENGAGE SKA RI, grant no. POCI-01-0145-FEDER-022217, funded by COMPETE 2020 and FCT, Portugal. We must also thank the anonymous referee for valuable suggestions and comments.

REFERENCES

Abbott B. P. et al., 2016, Phys. Rev. Lett., 116, 061102
Ables J. G., 1974, A&AS, 15, 383
Bañados E. et al., 2018, Nature, 553, 473
Bekenstein J. D., 1974, Phys. Rev. D, 9, 3292
Bekenstein J. D., 1975, Phys. Rev. D, 12, 3077
Caticha A., Preuss R., 2004, PhRvE, 70, 046127
Dewar R. C., 2005, JPhA, 38, L371
Dewar R. C., Maritan A., 2014, in Dewar R. C., Lineweaver C. H., Niven R. K., Regenauer-Lieb K., eds, Beyond the Second Law Entropy Production and Non-equilibrium Systems. Springer, New York, p. 49
Gull S. F., Daniell G. J., 1978, Nature, 272, 686
Jaynes E. T., 1957, Phys. Rev., 106, 620
Jaynes E. T., 1980, ARPC, 31, 579
Jaynes E. T., 1989, in Rosenkrantz R. D., ed., Papers on Probability, Statistics and Statistical Physics. Kluwer, London, p. 149
Kondepudi D., Prigogine I., 1998, Modern Thermodynamics. Wiley, New York, p. 392
Manti S., Gallerani S., Ferrara A., Greig B., Feruglio C., 2017, MNRAS, 466, 1160
Martyushev L. M., Seleznev V. D., 2006, Phys. Rep., 426, 1
Masters D. et al., 2012, ApJ, 755, 169
Matthews T. A., Sandage A. R., 1963, ApJ, 138, 30
Mohammad-Djafari A., Demoment G., 1988, in Skilling J., ed., Maximum Entropy and Bayesian Methods. Springer, Cambridge, p. 195
Pâris I. et al., 2018, A&A, 613, A51
Pontzen A., Governato F., 2013, MNRAS, 430, 121
Prigogine I., 1967, Introduction to Thermodynamics of Irreversible Processes, 3rd edn. John Wiley & Sons, New York
Prigogine I., 1978, Science, 201, 777
Richards G. T. et al., 2006, ApJ, 131, 2766
Ross N. P. et al., 2013, ApJ, 773, 14
Schmidt M., 1963, Nature, 197, 1040
Shimony A., 1985, Phys. Rev. Lett., 55, 1030
Skilling J., Bryan R. K., 1984, MNRAS, 211, 111
Tikochinsky Y., Tishby N. Z., Levine R. D., 1984, Phys. Rev. Lett., 52, 1357
Ziegler H., 1983, ZaMP, 34, 832
Ziegler H., 1987, An Introduction to Thermomechanics. North-Holland Publishing Company, Amsterdam, p. 229
Zunckel C., Trotta R., 2007, MNRAS, 380, 865

APPENDIX A

Table A1 was obtained from Richards et al. (2006), with the addition of the probability required to our objective and derived from their data, plus the probability we obtained for the comparison.

Table A1. The redshift, luminosities, and probability.

Z	L ($\times 10^{30}$) (erg s^{-1} Hz^{-1})	Probability	Probability MaxEnt
0.49	13.70	3.32×10^{-3}	3.84×10^{-4}
0.49	10.42	6.336×10^{-3}	1.89×10^{-3}
0.49	7.90	1.00×10^{-2}	6.36×10^{-3}
0.49	5.60	1.55×10^{-2}	1.59×10^{-2}
0.49	4.55	2.83×10^{-2}	3.20×10^{-2}
0.49	3.45	4.70×10^{-2}	5.42×10^{-2}
0.49	2.62	6.95×10^{-2}	8.09×10^{-2}
0.49	1.99	1.08×10^{-1}	1.10×10^{-1}
0.49	1.51	1.52×10^{-1}	1.38×10^{-1}
0.49	1.14	1.96×10^{-1}	1.65×10^{-1}
0.49	0.87	2.83×10^{-1}	1.88×10^{-1}
0.49	0.66	8.16×10^{-2}	2.08×10^{-1}
0.87	41.50	9.74×10^{-4}	1.07×10^{-2}
0.87	31.50	3.70×10^{-3}	2.07×10^{-2}
0.87	23.90	5.61×10^{-3}	3.40×10^{-2}
0.87	18.10	1.23×10^{-2}	4.96×10^{-2}
0.87	13.70	2.62×10^{-2}	6.61×10^{-2}
0.87	10.40	3.46×10^{-2}	8.21×10^{-2}
0.87	7.91	5.74×10^{-2}	9.69×10^{-2}
0.87	5.60	9.09×10^{-2}	1.10×10^{-1}
0.87	4.55	1.2×10^{-1}	1.21×10^{-1}
0.87	3.45	1.419×10^{-1}	1.30×10^{-1}
0.87	2.62	1.77×10^{-1}	1.37×10^{-1}
0.87	1.99	3.30×10^{-1}	1.43×10^{-1}
1.25	72.10	1.69×10^{-3}	2.79×10^{-3}
1.25	54.70	3.29×10^{-3}	8.38×10^{-3}
1.25	41.50	4.14×10^{-3}	1.93×10^{-2}
1.25	31.50	1.04×10^{-2}	3.63×10^{-2}
1.25	23.90	1.98×10^{-2}	5.87×10^{-2}
1.25	18.10	3.52×10^{-2}	8.45×10^{-2}
1.25	13.70	6.27×10^{-2}	1.11×10^{-1}
1.25	10.40	1.11×10^{-1}	1.37×10^{-1}
1.25	7.90	1.37×10^{-1}	1.61×10^{-1}
1.25	5.60	1.807×10^{-1}	1.82×10^{-1}
1.25	4.55	4.33×10^{-1}	1.99×10^{-1}
1.63	125.00	2.57×10^{-3}	1.55×10^{-4}
1.63	95.00	3.79×10^{-3}	1.14×10^{-3}
1.63	72.10	7.57×10^{-3}	5.20×10^{-3}
1.63	54.70	1.86×10^{-2}	1.64×10^{-2}
1.63	41.50	3.46×10^{-2}	3.93×10^{-2}
1.63	31.50	6.30×10^{-2}	7.61×10^{-2}
1.63	23.90	1.07×10^{-1}	1.26×10^{-1}
1.63	18.10	1.82×10^{-1}	1.84×10^{-1}
1.63	13.70	2.51×10^{-1}	2.46×10^{-1}
1.63	10.40	3.31×10^{-1}	3.06×10^{-1}
2.01	165.00	1.65×10^{-3}	4.48×10^{-3}
2.01	125.00	4.87×10^{-3}	1.25×10^{-2}
2.01	95.00	7.54×10^{-3}	2.71×10^{-2}
2.01	72.00	1.81×10^{-2}	4.89×10^{-2}
2.01	54.70	3.00×10^{-2}	7.65×10^{-2}
2.01	41.50	5.72×10^{-2}	1.07×10^{-1}
2.01	31.50	9.06×10^{-2}	1.39×10^{-1}
2.01	23.90	1.50×10^{-1}	1.69×10^{-1}
2.01	18.10	2.44×10^{-1}	1.96×10^{-1}
2.01	13.70	3.96×10^{-1}	2.19×10^{-1}
2.4	165.00	6.57×10^{-3}	7.35×10^{-3}
2.4	125.00	8.08×10^{-3}	2.07×10^{-2}
2.4	95.00	2.22×10^{-2}	4.56×10^{-2}
2.4	72.00	4.44×10^{-2}	8.28×10^{-2}
2.4	54.70	7.71×10^{-2}	1.30×10^{-1}
2.4	41.50	1.31×10^{-1}	1.84×10^{-1}
2.4	31.50	2.12×10^{-1}	2.39×10^{-1}
2.4	23.90	4.98×10^{-1}	2.91×10^{-1}
2.8	274.00	1.04×10^{-2}	3.64×10^{-3}
2.8	218.00	1.77×10^{-2}	1.09×10^{-2}
2.8	165.00	3.21×10^{-2}	3.01×10^{-2}
2.8	125.00	4.34×10^{-2}	6.53×10^{-2}
2.8	95.00	8.65×10^{-2}	1.17×10^{-1}
2.8	72.10	1.81×10^{-1}	1.83×10^{-1}
2.8	54.70	3.00×10^{-1}	2.57×10^{-1}
2.8	41.50	3.29×10^{-1}	3.32×10^{-1}
3.25	287.00	4.09×10^{-3}	4.64×10^{-4}
3.25	218.00	7.11×10^{-3}	2.32×10^{-3}
3.25	165.00	6.63×10^{-3}	7.85×10^{-3}
3.25	125.00	1.63×10^{-2}	2.00×10^{-2}

Table A1 – *continued*

Z	L ($\times 10^{30}$) (erg s^{-1} Hz^{-1})	Probability	Probability MaxEnt
3.25	95.00	3.25×10^{-2}	3.99×10^{-2}
3.25	72.10	4.80×10^{-2}	6.80×10^{-2}
3.25	54.70	6.79×10^{-2}	1.02×10^{-1}
3.25	41.50	1.15×10^{-1}	1.38×10^{-1}
3.25	31.50	1.45×10^{-1}	1.75×10^{-1}
3.25	23.90	2.25×10^{-1}	2.08×10^{-1}
3.25	18.10	3.32×10^{-1}	2.38×10^{-1}
3.75	165.00	1.65×10^{-2}	1.14×10^{-2}
3.75	125.00	2.74×10^{-2}	2.78×10^{-2}
3.75	95.00	6.42×10^{-2}	5.48×10^{-2}
3.75	72.10	6.57×10^{-2}	9.16×10^{-2}
3.75	54.70	1.04×10^{-1}	1.35×10^{-1}
3.75	41.49	1.54×10^{-1}	1.82×10^{-1}
3.75	31.50	1.98×10^{-1}	2.28×10^{-1}
3.75	23.90	3.69×10^{-1}	2.70×10^{-1}
4.25	218.00	4.65×10^{-2}	4.22×10^{-2}
4.25	125.00	9.28×10^{-2}	1.11×10^{-1}
4.25	95.00	1.09×10^{-1}	1.53×10^{-1}
4.25	72.10	1.73×10^{-1}	1.94×10^{-1}
4.25	54.70	2.18×10^{-1}	2.33×10^{-1}
4.25	41.50	3.61×10^{-1}	2.67×10^{-1}

APPENDIX B: LAGRANGE MULTIPLIERS METHOD: DATA, CONSTRAINTS, AND COMPUTATION

To develop the fundamentals of MaxEnt, consider the following set of data.

Object_1 A_1
Object_2 A_2
Object_3 A_3
... ...
Object_n A_n

Here, each Object_i is a quasar and A_i is its respective luminosity, with $i = (1, 2, 3, ..., n)$. From now on, we adapt Jaynes's notation to our work. Thus, we will call the luminosities by A_i, and their mean value by $\langle A\rangle$. Each A_i has a probability p_i to occur and we get from the data an average value $\langle A\rangle$ that can be obtained from arithmetic mean, weighted average, or from a more accurate form, using expression (1). This expression may be rewritten as

$$\langle A\rangle = \sum_{i=1}^{n} A_i\, p_i, \tag{B1}$$

where at one redshift z_k, the index n varies in the sum of $i = 1$, ..., n only on selected objects, that is, only in those three chosen luminosities in this redshift. Considering that the data set contains all possible values to occur, we have the bond condition that the summation of all probabilities must be equal to 1, see equation (2), or

$$\sum_{i=1}^{n} p_i = 1. \tag{B2}$$

The two Lagrange multipliers μ and λ are associated with these two equations, respectively, equations (B1) and (B2). Then, next they will be placed into a new form of the above equations.

From B1, we have

$$\mu\left[\sum_{i=1}^{n} A_i\, p_i - \langle A\rangle\right] = 0, \tag{B3}$$

and from B2,

$$\lambda\left[\sum_{i=1}^{n} p_i - 1\right] = 0. \tag{B4}$$

According to Jaynes, the method consists in the determination of the distribution function, p_i, by maximizing the so-called informational entropy,

$$H \equiv H(p_i, p_2, ...p_n) = -K\sum_{i=1}^{n} p_i \ln p_i;$$

this can be done by the standard method using the additional conditions (B1) and (B2) and the Lagrange multipliers λ and μ. The maximization procedure leads to the following result:

$$p_i = \mathrm{e}^{-\mu A_i - \lambda}. \tag{B5}$$

The two equations that we have to adjust to compute are obtained by taking (B5) into the equations of constraints (B1) and (B2), so we obtain the equations:

$$\mathrm{e}^{-\lambda}\sum_{i=1}^{n} A_i \mathrm{e}^{-\mu A_i} = \langle A\rangle, \tag{B6}$$

$$\mathrm{e}^{-\lambda}\sum_{i=1}^{n} \mathrm{e}^{-\mu A_i} = 1. \tag{B7}$$

The equation (B6) informs us that

$$\mathrm{e}^{\lambda} = \frac{\sum_{i=1}^{n} A_i \mathrm{e}^{-\mu A_i}}{\langle A\rangle}. \tag{B8}$$

Taking (B8) into equation (B7), we obtain an equation in μ to be solved:

$$\langle A\rangle \frac{\sum_{i=1}^{n} \mathrm{e}^{-\mu A_i}}{\sum_{i=1}^{n} A_i \mathrm{e}^{-\mu A_i}} = 1. \tag{B9}$$

To find λ, the obtained values of μ are taken into equation (B8). That is, the sequence of procedures to find μ from equation (B9) and substitute it into the equation (B8) to find λ.

With both Lagrange Multipliers found, we can get by MaxEnt the resulting probability (B5) and the average value (B1) for each redshift.

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