Scientific Research Publishing

International Journal of Astronomy and Astrophysics, 2025, 15(4), 372-389
https://www.scirp.org/journal/ijaa
ISSN Online: 2161-4725
ISSN Print: 2161-4717

New Probability Distributions in Astrophysics: XVI. Truncation for the Kiang Distribution and the VAST Catalog for Cosmic Voids

Lorenzo Zaninetti

Physics Department (Retired), University of Turin, Turin, Italy
Email: l.zaninetti@alice.it

How to cite this paper: Zaninetti, L. (2025) New Probability Distributions in Astrophysics: XVI. Truncation for the Kiang Distribution and the VAST Catalog for Cosmic Voids. *International Journal of Astronomy and Astrophysics*, **15**, 372-389.
https://doi.org/10.4236/ijaa.2025.154023

Received: September 17, 2025
Accepted: November 22, 2025
Published: November 25, 2025

Abstract

The release of the VAST catalog for the effective radius of cosmic voids has produced a large quantity of data that ask to be analysed with old and new probability distributions. We tested two families of distributions: the lognormal and the Kiang function. In more detail, we have analysed the standard lognormal distribution, the truncated lognormal distribution, the exponentiated lognormal distribution and the exponentiated lognormal distribution with truncation. About the Kiang function, we easily derived the distribution in radius. We introduced the effect of truncation for the general Kiang function and for the distribution in radius. All the new and old distributions were compared with the six different catalogs of cosmic voids for VAST.

Keywords

Methods: Statistical, Cosmology: Observations, (Cosmology): Large-Scale Structure of the Universe

1. Introduction

The VAST void catalog for SDSS DR7, see [1], computes the effective radii and centres of cosmic voids for two cosmologies with three different methods. This catalog has been used to explain the gamma-ray dark matter emission from halos of galaxies, see [2], and the calibration of the background universe, see [3] [4]. A useful quantity to model the cosmic voids is the spherically averaged stacked density profile

$$\delta(r)=\frac{n(r)}{\overline{n}}-1, \tag{1}$$

where r is the radius of the void, $n(r)$ is the spherical density of galaxies at r and $\overline{n}$ is the average density of galaxies of the catalog, see [1]. The ratio $\frac{n(r)}{\overline{n}}$ is zero at $r=0$ and reaches the maximum value of 1 at $r \approx 20$ Mpc. The availability of reliable data for the cosmic voids requires a precise analysis of the data. There are, roughly speaking, two methods of fitting data. One method adopts a generic probability distribution ad hoc, e.g. the lognormal distribution. The other one explores the probability distributions which arise from the particular problem analysed. In the case of cosmic voids, a useful tool is the distribution connected with the Voronoi diagrams, *i.e.* the Kiang distribution [5], which dates back to 1961. Two targets of the research on cosmic voids are the determination of the radius of the maximum sphere, which is ≈ 50 Mpc, and the evaluation of the radius of the minimum sphere which can be detected, which is ≈ 8 Mpc. The maximum radius is connected with the physical process which produces the voids, but the minimum radius is connected with the observational techniques. The presence of the two boundaries in the distribution of radii requires the introduction of the truncated PDFs.

In order to explore the two above distributions, Section 2 reviews the lognormal distribution and the other three modified lognormal distributions. Section 3 applies the lognormal family to the VAST catalog of voids. Section 4 reviews the Kiang function, its particularization to the radius of the Voronoi diagrams in 3D and introduces the truncation for both the distributions.

2. The Lognormal Family

In the following, PDF means probability density function and DF distribution function. We now review the lognormal distribution, the truncated lognormal distribution, the exponentiated lognormal distribution and the exponentiated lognormal distribution with truncation.

2.1. The Lognormal Distribution

Let X be a random variable taking values x in the interval $[0,\infty]$; the first definition for the *lognormal* PDF, following [6] or formula (14.2) in [7], is

$$f(x:m,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\frac{-\left[\ln(x/m)\right]^2}{2\sigma^2}. \tag{2}$$

The average value, $E(m,\sigma)$, is

$$E(m,\sigma)=m\mathrm{e}^{\frac{\sigma^2}{2}}, \tag{3}$$

and the distribution function, $F(x:m,\sigma)$, is given by

$$F(x:m,\sigma)=\frac{1}{2}-\frac{\operatorname{erf}\left(\frac{\sqrt{2}\left(\ln(m)-\ln(x)\right)}{2\sigma}\right)}{2}. \tag{4}$$

The second definition is

$$f\left(x:\mu,\sigma\right)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\frac{-\left(\ln x-\mu\right)^2}{2\sigma^2}, \tag{5}$$

where $m=\exp\mu$ and $\mu=\ln m$.

2.2. The Truncated Lognormal Distribution

Let X be a random variable defined in $\left[x_l,x_u\right]$; the *truncated lognormal* PDF (f_T) is based on the first definition of the lognormal as given by Equation (2)

$$f_T\left(x:m,\sigma,x_l,x_u\right)=\frac{\sqrt{2}\mathrm{e}^{-\frac{1}{2}\frac{1}{\sigma^2}\left(\ln\left(\frac{x}{m}\right)\right)^2}}{-\sqrt{\pi}\sigma\left(\operatorname{erf}\left(\frac{1}{2}\frac{\sqrt{2}}{\sigma}\ln\left(\frac{x_l}{m}\right)\right)-\operatorname{erf}\left(\frac{1}{2}\frac{\sqrt{2}}{\sigma}\ln\left(\frac{x_u}{m}\right)\right)\right)x}, \tag{6}$$

where m is the scale parameter, σ is the shape parameter, x_l denotes the minimal value, and x_u denotes the maximal value. The introduction of the following coefficients allows a compact notation:

$$a_1=\frac{1}{2}\frac{\sqrt{2}\left(-\sigma^2+\ln\left(x_l\right)-\ln\left(m\right)\right)}{\sigma},$$

$$a_2=\frac{1}{2}\frac{\sqrt{2}\left(\sigma^2+\ln\left(m\right)-\ln\left(x_u\right)\right)}{\sigma},$$

$$a_3=\frac{1}{2}\frac{\sqrt{2}\left(\ln\left(x_l\right)-\ln\left(m\right)\right)}{\sigma},$$

$$a_4=\frac{1}{2}\frac{\sqrt{2}\left(-\ln\left(x_u\right)+\ln\left(m\right)\right)}{\sigma},$$

$$a_5=\frac{1}{2}\frac{\sqrt{2}\left(-2\sigma^2+\ln\left(x_l\right)-\ln\left(m\right)\right)}{\sigma},$$

$$a_6=\frac{1}{2}\frac{\sqrt{2}\left(2\sigma^2+\ln\left(m\right)-\ln\left(x_u\right)\right)}{\sigma},$$

$$a_7=\frac{1}{2}\frac{\sqrt{2}\left(-2\sigma^2+\ln\left(x_u\right)-\ln\left(m\right)\right)}{\sigma},$$

$$a_8=\frac{1}{2}\frac{\sqrt{2}\left(\ln\left(x_u\right)-\ln\left(m\right)\right)}{\sigma}.$$

In this compact notation, the PDF is

$$f_T\left(x:m,\sigma,x_l,x_u\right)=\frac{-\sqrt{2}\mathrm{e}^{-\frac{1}{2}\frac{1}{\sigma^2}\left(\ln\left(\frac{x}{m}\right)\right)^2}}{\sqrt{\pi}\sigma\left(\operatorname{erf}\left(a_3\right)-\operatorname{erf}\left(a_8\right)\right)x}, \tag{7}$$

the DF is

$$F_T\left(x:m,\sigma,x_l,x_u\right)=\frac{-\operatorname{erf}\left(\frac{1}{2}\frac{\sqrt{2}}{\sigma}\ln\left(\frac{x}{m}\right)\right)+\operatorname{erf}\left(a_3\right)}{\operatorname{erf}\left(a_3\right)-\operatorname{erf}\left(a_8\right)}, \tag{8}$$

and the mean, $E_T\left(m,\sigma,x_l,x_u\right)$, is

$$E_T\left(m,\sigma,x_l,x_u\right)=\frac{\mathrm{e}^{\frac{1}{2}\sigma^2}m\left(\mathrm{erf}\left(a_1\right)+\mathrm{erf}\left(a_2\right)\right)}{\mathrm{erf}\left(a_3\right)+\mathrm{erf}\left(a_4\right)}. \tag{9}$$

More details can be found in [8].

2.3. The Exponentiated Lognormal Distribution

The modified lognormal distribution with a power-law (MLP) defined in the interval $\left[0,\infty\right]$ has the PDF

$$f\left(x:\alpha,\sigma,\mu\right)=\frac{x^{-1-\alpha}\mathrm{erfc}\left(\frac{\sqrt{2}\left(\alpha\sigma-\frac{\ln\left(x\right)-\mu}{\sigma}\right)}{2}\right)\alpha}{2\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu}}, \tag{10}$$

where $\mathrm{erfc}\left(x\right)$ is the complementary error function

$$\mathrm{erfc}\left(x\right)x=\frac{2}{\sqrt{\pi}}\int_x^\infty \mathrm{e}^{-t^2}\mathrm{d}t=1-\mathrm{erf}\left(x\right), \tag{11}$$

see the handbook [9]. Its DF is

$$F\left(x:\alpha,\sigma,\mu\right)=-\frac{\mathrm{erfc}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln\left(x\right)+\mu\right)}{2\sigma}\right)x^{-\alpha}\mathrm{e}^{\frac{1}{2}\alpha^2\sigma^2+\alpha\mu}}{2}+1-\frac{\mathrm{erfc}\left(\frac{\sqrt{2}\left(\ln\left(x\right)-\mu\right)}{2\sigma}\right)}{2}, \tag{12}$$

see formula (16) in [10]. The first moment, or mean, $E\left(\alpha,\sigma,\mu\right)$, is defined for $\alpha>1$:

$$E\left(\alpha,\sigma,\mu\right)=\frac{\alpha\mathrm{e}^{\frac{\sigma^2}{2}+\mu}}{\alpha-1}, \tag{13}$$

see formula (19) in [10].

The variance, $Var\left(\alpha,\sigma,\mu\right)$, is defined for $\alpha>2$:

$$Var\left(\alpha,\sigma,\mu\right)=-\frac{\alpha^2\left(\mathrm{e}^{\frac{\sigma^2}{2}+\mu}\right)^2}{\left(\alpha-1\right)^2}+\frac{\alpha\mathrm{e}^{2\sigma^2+2\mu}}{\alpha-2}, \tag{14}$$

see formula (21) in [10]. The mode should be evaluated numerically. The random generation of the variate X of the truncated MLP is accomplished by solving the following nonlinear equation in x:

$$F\left(x:\alpha,\sigma,\mu\right)=R, \tag{15}$$

where R is the unit rectangular variate. More details can be found in [11].

2.4. Truncated MLP

The truncated version of the MLP defined in the interval $\left[x_l, x_u\right]$ has the following PDF:

$$f_T\left(x:\alpha,\sigma,\mu,x_l,x_u\right)=\frac{x^{-1-\alpha}\operatorname{erfc}\left(\frac{\sqrt{2}\left(\alpha\sigma-\frac{\ln(x)-\mu}{\sigma}\right)}{2}\right)}{A}, \tag{16}$$

where

$$\begin{aligned}A=\frac{1}{\alpha}\times\Bigg[&\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu}\operatorname{erf}\left(\frac{\sqrt{2}\left(\ln(x_u)-\mu\right)}{2\sigma}\right)\\&-\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu}\operatorname{erf}\left(\frac{\sqrt{2}\left(\ln(x_l)-\mu\right)}{2\sigma}\right)\\&-x_l^{-\alpha}\operatorname{erf}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln(x_l)+\mu\right)}{2\sigma}\right)\\&+x_u^{-\alpha}\operatorname{erf}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln(x_u)+\mu\right)}{2\sigma}\right)+x_l^{-\alpha}-x_u^{-\alpha}\Bigg].\end{aligned} \tag{17}$$

The DF is

$$\begin{aligned}F_T\left(x:\alpha,\sigma,\mu,x_l,x_u\right)=\frac{1}{\alpha A}\times\Bigg[&-x_l^{-\alpha}\operatorname{erf}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln(x_l)+\mu\right)}{2\sigma}\right)\\&+\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu}\operatorname{erf}\left(\frac{\sqrt{2}\left(\ln(x)-\mu\right)}{2\sigma}\right)\\&-\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu}\operatorname{erf}\left(\frac{\sqrt{2}\left(\ln(x_l)-\mu\right)}{2\sigma}\right)\\&+x^{-\alpha}\operatorname{erf}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln(x)+\mu\right)}{2\sigma}\right)-x^{-\alpha}+x_l^{-\alpha}\Bigg].\end{aligned} \tag{18}$$

The first moment, or average value, is

$$\begin{aligned}&E\left(x:\alpha,\sigma,\mu,x_l,x_u\right)\\&=\frac{1}{A(-1+\alpha)}\times\Bigg[\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu+\frac{1}{2}\sigma^2+\mu}\operatorname{erf}\left(\frac{\sqrt{2}\left(\sigma^2-\ln(x_l)+\mu\right)}{2\sigma}\right)\\&-\mathrm{e}^{-\frac{1}{2}\alpha^2\sigma^2-\alpha\mu+\frac{1}{2}\sigma^2+\mu}\operatorname{erf}\left(\frac{\sqrt{2}\left(\sigma^2-\ln(x_u)+\mu\right)}{2\sigma}\right)\\&-x_l^{1-\alpha}\operatorname{erf}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln(x_l)+\mu\right)}{2\sigma}\right)\\&+x_u^{1-\alpha}\operatorname{erf}\left(\frac{\sqrt{2}\left(\alpha\sigma^2-\ln(x_u)+\mu\right)}{2\sigma}\right)+x_l^{1-\alpha}-x_u^{1-\alpha}\Bigg].\end{aligned} \tag{19}$$

More details can be found in [11].

3. Astrophysical Applications

This section reviews the adopted statistics and the various catalogs for cosmic voids. The application of four types of lognormal to the last and more complete catalog of voids is reported.

3.1. Adopted Statistics

The Kolmogorov-Smirnov test (K-S), see [12]-[14], does not require the data to be binned. The K-S test, as implemented by the FORTRAN subroutine KSONE in [15], given ordered statistics x_i, $i=1,\cdots,N$, where N defines the number of elements in the sample, finds the maximum distance, D, between the theoretical and the astronomical DFs. The significance is

$$Q_{KS}(\lambda)=2\left(\sum_{j=1}^{\infty}(-1)^{j-1}\,\mathrm{e}^{-2j^2\lambda^2}\right), \tag{20}$$

and the significance level to have an observed D is

$$P_{KS}=Probability\left(D>D_{observed}\right)=Q_{KS}\left(Q_{KS}\left(\sqrt{N}+0.12+\frac{0.11}{\sqrt{N}}\right)D\right), \tag{21}$$

see formulas 14.3.5 and 14.3.9 in [15]. If $P_{KS}\geq 0.1$, then the goodness of the fit is believable.

3.2. Catalogs of Voids

The *first* catalog of cosmic voids can be found in [16], where the effective radius of the voids, R_{eff}, has been derived to be

$$R_{eff}=18.23h^{-1}\ \text{Mpc}\quad \text{Pam } et\ al.\ 2012. \tag{22}$$

The *second* catalog is that with radii up to redshift $0.12h^{-1}$ Mpc in (SDSS-DR7), see [17],

$$R_{eff}=11.85h^{-1}\ \text{Mpc}\quad \text{Varela } et\ al.\ 2012. \tag{23}$$

The *third* catalog is that of the Baryon Oscillation Spectroscopic Survey, see [18],

$$R_{eff}=57.53h^{-1}\ \text{Mpc}\quad \text{Mao } et\ al.\ 2017. \tag{24}$$

The *fourth* catalog is that of the VAST void catalog for SDSS DR7 which uses three algorithms, VoidFinder, V^2/VIDE and V^2/REVOLVER in the framework of two cosmologies, Planck2018 and WMAP5, see [1]. The data with Cartesian coordinates and effective radius R_{eff} in Mpc are available at the following address https://zenodo.org/records/11043278. A graphical display of the voids is reported in **Figure 1**.

3.3. Statistics of the Voids

The statistics for the lognormal distribution of the VAST catalog are given in **Ta-**

ble 1, for the truncated lognormal distribution in Table 2, for the MLP distribution in Table 3 and for the truncated MLP distribution in Table 4.

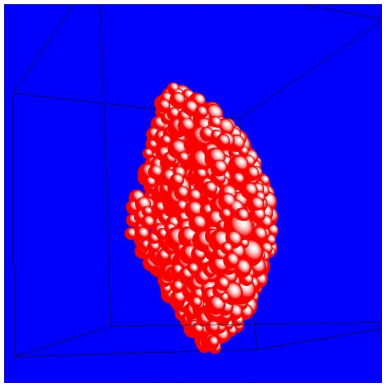

Figure 1. 3D display through spheres of the 1184 voids derived with the VoidFinder method in WMAP5 cosmology.

Table 1. Numerical values of D, the maximum distance between theoretical and observed DFs, and P_{KS}, significance level, in the K-S test for the lognormal distribution, see Equation (5), for different cosmologies and methods.

Cosmology	Method	Parameters	D	P_{KS}
WMAP5	VoidFinder, $N=1184$	$\sigma=0.177$, $\mu=2.765$	0.0652	7.84×10^{-5}
Planck 2018	V^2/VIDE, $N=531$	$\sigma=0.336$, $\mu=2.841$	0.0669	0.0163
Planck 2018	V^2/REVOLVER, $N=518$	$\sigma=0.229$, $\mu=2.988$	0.0677	0.0162

Table 2. Numerical values of D, the maximum distance between theoretical and observed DFs, and P_{KS}, significance level, in the K-S test for the truncated lognormal distribution, see Equation (7), for different cosmologies and methods.

Cosmology	Method	Parameters	D	P_{KS}
WMAP5	VoidFinder, $N=1184$	$\sigma=0.1875$, $\mu=2.755$	0.0551	1.05×10^{-21}
Planck 2018	V^2/VIDE, $N=531$	$\sigma=0.44$, $\mu=2.701$	0.0193	0.988
Planck 2018	V^2/REVOLVER, $N=518$	$\sigma=0.335$, $\mu=2.833$	0.021	0.972

Table 3. Numerical values of D, the maximum distance between theoretical and observed DFs, and P_{KS}, significance level, in the K-S test for the exponentiated lognormal distribution, see Equation (10), for different cosmologies and methods.

Cosmology	Method	Parameters	D	P_{KS}
WMAP5	VoidFinder, $N=1184$	$\sigma=0.0982$, $\mu=2.6$, $\alpha=6.32$	0.0416	3.16×10^{-2}
Planck 2018	V^2/VIDE, $N=531$	$\sigma=0.31$, $\mu=2.73$, $\alpha=9.23$	0.063	2.78×10^{-2}
Planck 2018	V^2/REVOLVER, $N=518$	$\sigma=0.214$, $\mu=2.91$, $\alpha=13.96$	0.0665	1.94×10^{-2}

Table 4. Numerical values of D, the maximum distance between theoretical and observed DFs, and P_{KS}, significance level, in the K-S test for the exponentiated and truncated lognormal distribution or truncated MLP see Equation (16), for different cosmologies and methods.

Cosmology	Method	Parameters	D	P_{KS}
WMAP5	VoidFinder, $N=1184$	$\sigma=0.0931$, $\mu=2.59$, $\alpha=5.61$	0.0363	8.57×10^{-2}
Planck 2018	V^2/VIDE, $N=531$	$\sigma=0.387$, $\mu=2.56$, $\alpha=5.94$	0.0211	0.969
Planck 2018	V^2/REVOLVER, $N=518$	$\sigma=0.284$, $\mu=2.74$, $\alpha=7.2$	0.022	0.961

We now report the best results for the three cases here analysed as empirical and theoretical PDFs: **Figure 2** reports the WMAP5 and VoidFinder data in the case of the truncated MLP distribution, **Figure 3** reports the Planck 2018 and V^2/VIDE data in the case of the truncated lognormal distribution and **Figure 4** reports the Planck 2018 and V^2/REVOLVER data in the case of the truncated lognormal distribution.

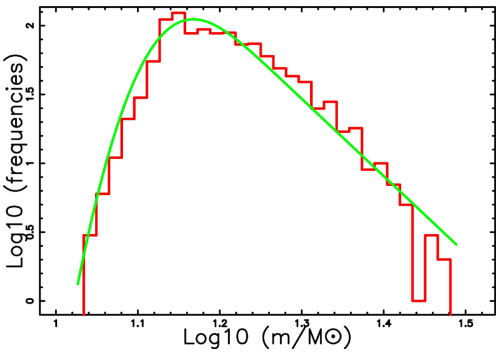

Figure 2. Logarithmic histogram of distribution of voids as given by WMAP5 and VoidFinder data (red) with a superposition of the truncated MLP distribution when the number of bins, m, is 30 (green line). Parameters as in **Table 4**. Vertical and horizontal axes have logarithmic scales.

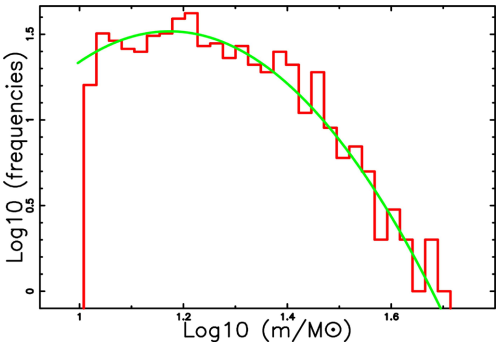

Figure 3. Logarithmic histogram of distribution of voids as given by Planck 2018 and V^2/VIDE data (red) with a superposition of the truncated lognormal distribution when the number of bins, m, is 30 (green line). Parameters as in **Table 2**. Vertical and horizontal axes have logarithmic scales.

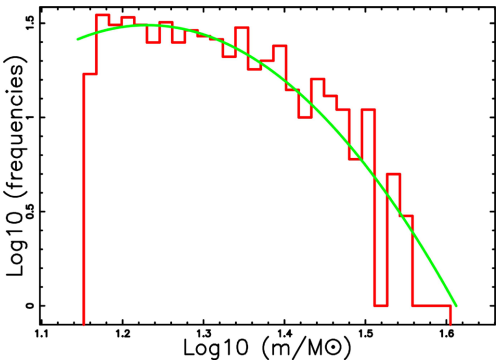

Figure 4. Logarithmic histogram of distribution of voids as given by Planck 2018 and V^2/REVOLVER data (red) with a superposition of the truncated lognormal distribution when the number of bins, m, is 30 (green line). Parameters as in **Table 2**. Vertical and horizontal axes have logarithmic scales.

4. Voronoi Diagrams

This section first deals with the generic Kiang distribution and then particularizes it to the distribution in radius. The effect of truncation on both of these two distributions is reported.

4.1. Generic Distribution

Let X be a random variable taking values x in the interval $[0,\infty]$; the *gamma* PDF is

$$f(x;b,c)=\frac{\left(\frac{x}{b}\right)^{c-1}\mathrm{e}^{-\frac{x}{b}}}{b\Gamma(c)} \tag{25}$$

where

$$\Gamma(z)=\int_0^{\infty}\mathrm{e}^{-t}t^{z-1}\mathrm{d}t, \tag{26}$$

is the gamma function, $b>0$ is the scale parameter and $c>0$ is the shape parameter, see formula (17.23) in [7]. Its expected value is

$$\mu(x;b,c)=bc. \tag{27}$$

We now re-scale the above PDF in such a way that the expected value is 1 and the gamma variate, $H(x;c)$, ([5]) is obtained

$$H(x;c)=\frac{c}{\Gamma(c)}(cx)^{c-1}\exp(-cx), \tag{28}$$

where $0\le x<\infty$, $c>0$. The DF of the Kiang function is

$$F(x;c)=1-\frac{\Gamma(c,cx)}{\Gamma(c)}, \tag{29}$$

where $\Gamma(a,z)$ is the incomplete Gamma function defined as

$$\Gamma(a,z)=\int_z^{\infty}t^{a-1}\mathrm{e}^{-t}\mathrm{d}t, \tag{30}$$

see [9]. The Kiang PDF has a mean, μ, of

$$\mu = 1, \tag{31}$$

the variance is

$$\sigma^2 = \frac{1}{c}, \tag{32}$$

the skewness is

$$\text{skewness} = \frac{2}{\sqrt{c}}, \tag{33}$$

the kurtosis is

$$\text{kurtosis} = \frac{3c+6}{c}, \tag{34}$$

and the mode is at

$$\text{mode} = \frac{c-1}{c}. \tag{35}$$

An approximate expression for the median can be obtained by an order 3 Taylor series for the DF about the mean

$$\text{median} = -\frac{-2c^c e^{-c} + \sqrt{\left(c^c\right)^2 \left(e^{-c}\right)^2 - 2e^{-c}c^c\Gamma(c,c) + e^{-c}c^c\Gamma(c)}}{c^c e^{-c}} \tag{36}$$

The percent error of the above median is 0.013% at $c = 2$.

In the case of a 1D Poissonian Voronoi tessellation (PVT), $c = 2$ is an exact analytical result, but c is supposed to be 4 or 6 for 2D or 3D PVTs, respectively, the so called Kiang conjecture [5]. The value $c = 6$ in 3D was successively refined to 5.5 due to a change in the generator of random numbers [19], or to 5.78 as deduced by [20]. A *first method* to derive the numerical value of c is to equalize the variance of the sample, var, with the theoretical variance

$$\text{var} = \sigma^2. \tag{37}$$

A *second method* to derive the numerical value of c is the maximum likelihood (MLE) method, which maximizes

$$\Lambda = nc\ln(c) - n\ln(\Gamma(c)) + \sum_{j=1}^{n}\left(\ln(x_j)c - cx_j - \ln(x_j)\right), \tag{38}$$

where n is the number of elements in the sample x_j. The value of c is found solving the following non-linear equation

$$\frac{\partial\Lambda}{\partial c} = n\ln(c) + n - n\Psi(c) + \sum_{j=1}^{n}\left(\ln(x_j) - x_j\right) = 0, \tag{39}$$

where $\Psi(z)$ is the digamma or Psi function defined as

$$\Psi(z) = \Gamma'(z)/\Gamma(z), \tag{40}$$

where $\mathcal{R}z > 0$, see [9].

Figure 5 reports a typical evaluation of the volume distribution of the Voronoi diagrams in the presence of Poissonian seeds.

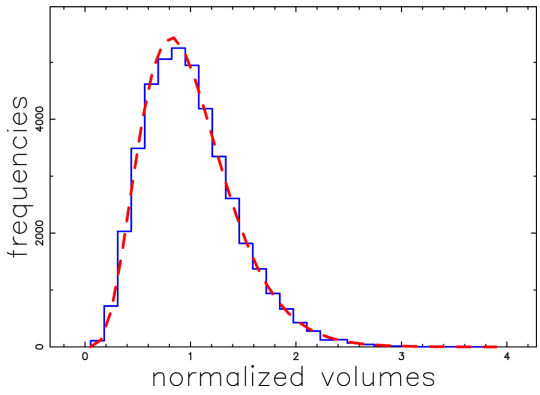

Figure 5. Histogram of volume distribution for the Voronoi diagrams generated by $N=1122$. Poissonian seeds (blue line) with a superposition of a red dashed line representing the Kiang PDF as given by Equation (28) when $c=5.378$ as deduced with the first method. The two parameters of the K-S test are $D=9.361\times10^{-3}$ and $P_{KS}=1.2\times10^{-3}$.

4.2. Radius Distribution

We approximate the volume of each cell of the Voronoi diagrams, v, with

$$v=\frac{4}{3}\pi R^3. \tag{41}$$

The Kiang function in volumes is

$$H(v;c)=\frac{c(cv)^{c-1}\,\mathrm{e}^{-cv}}{\Gamma(c)}, \tag{42}$$

and that in radius, R,

$$H(R;c)=\frac{4c\left(\frac{4c\pi R^3}{3}\right)^{c-1}\mathrm{e}^{-\frac{4c\pi R^3}{3}}\pi R^2}{\Gamma(c)}, \tag{43}$$

where c is the variable to be found from the data of the volumes. We now introduce the scale b: the scaled PDF is

$$f(R;b,c)=\frac{4c\left(\frac{4c\pi R^3}{3}\right)^{c-1}\mathrm{e}^{-\frac{4c\pi R^3}{3}}\pi R^2}{\Gamma(c)}, \tag{44}$$

and the scaled DF is

$$F(R;b,c)=1+\frac{4^c\pi^c 3^{-c}\left(R^3\right)^c b^{-3c}c^c\mathrm{e}^{-\frac{4c\pi R^3}{3b^3}}-\Gamma\left(c+1,\frac{4c\pi R^3}{3b^3}\right)}{c\Gamma(c)}. \tag{45}$$

The above distribution has a mean

$$\mu(b,c)=b\Gamma\left(c+\frac{1}{3}\right)6^{1/3}, \tag{46}$$

variance

$$\sigma^2(b,c)=b^2\left(\Gamma\left(c+\frac{2}{3}\right)\Gamma(c)-\Gamma\left(c+\frac{1}{3}\right)^2\right)6^{2/3}, \tag{47}$$

skewness

$$\text{skewness}=\Gamma(c)^3\,c-3\Gamma\left(c+\frac{2}{3}\right)\Gamma\left(c+\frac{1}{3}\right)\Gamma(c)+2\Gamma\left(c+\frac{1}{3}\right)^3, \tag{48}$$

and kurtosis

$$\text{kurtosis}=\frac{N}{D}, \tag{49}$$

where

$$\begin{aligned}N=3\Gamma\left(c+\frac{4}{3}\right)\times\Bigg(&-2\left(c+\frac{1}{3}\right)^2\Gamma\left(c+\frac{4}{3}\right)\Gamma(c)\Gamma\left(c+\frac{2}{3}\right)\\&+\Gamma\left(c+\frac{4}{3}\right)^3+\left(c+\frac{1}{3}\right)^3\left(c-\frac{1}{9}\right)\Gamma(c)^3\Bigg),\end{aligned} \tag{50}$$

and

$$D=\left(\left(c+\frac{1}{3}\right)^2\Gamma(c)\Gamma\left(c+\frac{2}{3}\right)-\Gamma\left(c+\frac{4}{3}\right)^2\right)^2. \tag{51}$$

An approximate expression for the median can be obtained by an order two Taylor series for the DF about the mean:

$$\begin{aligned}\text{median}=\frac{1}{12\pi^{1/3}c^{4/3}\Gamma(c)}\times\Bigg(&-b\Gamma\left(c+\frac{1}{3}\right)\Bigg(c\Gamma(c)^{3c+1}\Gamma\left(c+\frac{1}{3}\right)^{-3c}\mathrm{e}^{\frac{\Gamma\left(c+\frac{1}{3}\right)^3}{\Gamma(c)^3}}\\&-2\Gamma\left(c+1,\frac{\Gamma\left(c+\frac{1}{3}\right)^3}{\Gamma(c)^3}\right)\mathrm{e}^{\frac{\Gamma\left(c+\frac{1}{3}\right)^3}{\Gamma(c)^3}}\Gamma(c)^{3c}\Gamma\left(c+\frac{1}{3}\right)^{-3c}-6c+2\Bigg)6^{1/3}\Bigg).\end{aligned} \tag{52}$$

One way to deduce the parameters associates b with the maximum of the sample; c is found by solving the non-linear Equation (37) with the theoretical variance as given by Equation (47). Another way, the MLE method, maximizes

$$\begin{aligned}\Lambda=&\,2nc\ln(2)-nc\ln(3)+nc\ln(\pi)-3nc\ln(b)+nc\ln(c)+n\ln(3)\\&-n\ln\left(\Gamma(c)\right)+\sum_{j=1}^{n}\left(3c\ln\left(x_j\right)-\ln\left(x_j\right)-\frac{4c\pi x_j^3}{3b^3}\right)\\=&\,0.\end{aligned} \tag{53}$$

The values of b and c are found by solving the two following non-linear equations:

$$\frac{\partial\Lambda}{\partial b}=\frac{c\left(-3nb^3+4\pi\left(\sum_{j=1}^{n}x_j^3\right)\right)}{b^4}=0, \tag{54}$$

and

$$\frac{\partial \Lambda}{\partial c} = 2n\ln(2) - n\ln(3) + n\ln(\pi) - 3n\ln(b) + n\ln(c) + n - n\Psi(c) - \frac{\sum_{j=1}^{n}\left(4\pi x_j^3 - 9\ln(x_j)b^3\right)}{3b^3} = 0. \tag{55}$$

The statistics for the Kiang function in radius of the VAST catalog are given in Table 5.

Table 5. Numerical values of D, the maximum distance between theoretical and observed DFs, P_{KS}, significance level, in the K-S test for the Kiang distribution in radius, see Equation (44), and percent error for the approximated median, see Equation (52), for different cosmologies and methods.

Cosmology	Method	Parameters	D	P_{KS}	% error median
WMAP5	VoidFinder, $N=1184$	$b=26.99$ Mpc $c=3.03$	0.103	1.73×10^{-11}	4.37%
Planck 2018	V^2/VIDE, $N=531$	$b=33.56$ Mpc $c=0.99$	0.129	3.41×10^{-8}	11.57%
Planck 2018	V^2/REVOLVER, $N=518$	$b=34.87$ Mpc $c=2.08$	0.019	6.3×10^{-5}	5.87%

As a visual example, Figure 6 reports the Planck 2018 and V^2/REVOLVER data in the case of the Kiang distribution in radius.

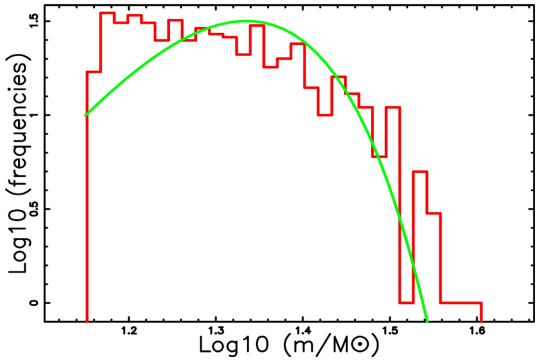

Figure 6. Logarithmic histogram of distribution of voids as given by Planck 2018 and V^2/REVOLVER data (red) with a superposition of the Kiang distribution in radius when the number of bins, m, is 30 (green line). Parameters as in Table 5. Vertical and horizontal axes have logarithmic scales.

4.3. The Truncated Generic Distribution

The truncated gamma variate $H_T(x;c,x_l,x_u)$ is

$$H_T(x;c,x_l,x_u) = \frac{c(cx)^{c-1}\,\mathrm{e}^{-cx}}{-\Gamma(c,cx_u)+\Gamma(c,cx_l)}, \tag{56}$$

where $x_l \le x < x_u$, $c > 0$.

The truncated distribution function is

$$F_T\left(x;c,x_l,x_u\right)=\frac{\mathrm{e}^{-cx_l}c^c x_l^c - c^c x^c \mathrm{e}^{-cx} + \Gamma\left(c+1,cx\right)-\Gamma\left(c+1,cx_l\right)}{c\left(\Gamma\left(c,cx_u\right)-\Gamma\left(c,cx_l\right)\right)}. \tag{57}$$

The truncated Kiang PDF has a mean, μ_T, given by

$$\mu_T\left(c,x_l,x_u\right)=\frac{-\Gamma\left(c+1,cx_l\right)+\Gamma\left(c+1,cx_u\right)}{c\left(\Gamma\left(c,cx_u\right)-\Gamma\left(c,cx_l\right)\right)}, \tag{58}$$

and the r th moment about the origin is, μ_r', is

$$\begin{aligned}&\mu_r'\left(c,x_l,x_u\right)\\&=\frac{-x_u^{c+r}\mathrm{e}^{-cx_u}c^c + x_l^{c+r}\mathrm{e}^{-cx_l}c^c - c^{-r}\Gamma\left(1+c+r,cx_l\right)+c^{-r}\Gamma\left(1+c+r,cx_u\right)}{\left(\Gamma\left(c,cx_u\right)-\Gamma\left(c,cx_l\right)\right)\left(c+r\right)}.\end{aligned} \tag{59}$$

The variance is

$$\begin{aligned}\sigma^2 =& -\frac{\left(-c^c\mathrm{e}^{-cx_u}x_u^{c+1}+c^c\mathrm{e}^{-cx_l}x_l^{c+1}-\frac{\Gamma\left(c+2,cx_l\right)}{c}+\frac{\Gamma\left(c+2,cx_u\right)}{c}\right)^2}{\left(\Gamma\left(c,cx_u\right)-\Gamma\left(c,cx_l\right)\right)^2\left(c+1\right)^2}\\&+\frac{-c^c\mathrm{e}^{-cx_u}x_u^{c+2}+c^c\mathrm{e}^{-cx_l}x_l^{c+2}-\frac{\Gamma\left(3+c,cx_l\right)}{c^2}+\frac{\Gamma\left(3+c,cx_u\right)}{c^2}}{\left(\Gamma\left(c,cx_u\right)-\Gamma\left(c,cx_l\right)\right)\left(c+2\right)}.\end{aligned} \tag{60}$$

One way to find c uses Equation (37) in which σ^2 is given by the previous equation. Another way to find c is the MLE method, which maximizes

$$\Lambda = nc\ln\left(c\right)+n\ln\left(\frac{1}{-\Gamma\left(c,cx_u\right)+\Gamma\left(c,cx_l\right)}\right)+\sum_{i=1}^{n}\left(\ln\left(x_i\right)c-cx_i-\ln\left(x_i\right)\right). \tag{61}$$

In this case, $\frac{\partial\Lambda}{\partial c}$ has a complicated expression which is not reported. As an application, Figure 7 reports a typical evaluation of the volume distribution of the Voronoi diagrams in the presence of Poissonian seeds.

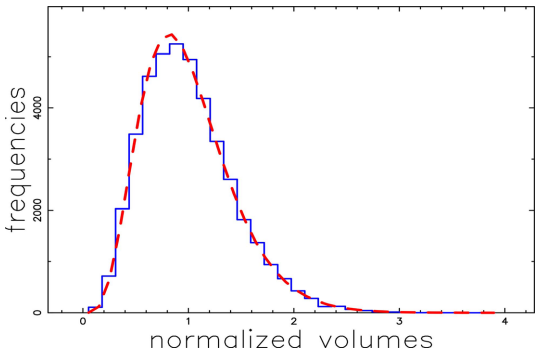

Figure 7. Histogram of volume distribution for the Voronoi diagrams generated by $N=1122$. Poissonian seeds (blue line) with a superposition of a red dashed line representing the truncated Kiang PDF as given by Equation (56) when $c=5.385$ as deduced with the second method. The two parameters of the K–S test are $D=9.296\times10^{-3}$ and $P_{KS}=1.336\times10^{-3}$.

4.4. Radius Distribution with Truncation

We now introduce the truncation in the radius distribution with scale, see Equation (44). The truncated PDF is

$$f_T\left(R;b,c,R_l,R_u\right)=\frac{4c^2\left(\frac{4c\pi R^3}{3b^3}\right)^{c-1}\mathrm{e}^{-\frac{4c\pi R^3}{3b^3}}\pi R^2}{k}, \tag{62}$$

where $R_l \le R < R_u$, $c>0$, $b>0$ and

$$\begin{aligned} k = {} & 4^c 3^{-c}\pi^c c^c b^{-3c}\left(R_u^c\right)^3 \mathrm{e}^{-\frac{4c\pi R_u^3}{3b^3}} b^3 - 4^c 3^{-c}\pi^c c^c b^{-3c}\left(R_l^c\right)^3 \mathrm{e}^{-\frac{4c\pi R_l^3}{3b^3}} b^3 \\ & + \Gamma\left(c+1,\frac{4c\pi R_l^3}{3b^3}\right)b^3 - \Gamma\left(c+1,\frac{4c\pi R_u^3}{3b^3}\right)b^3. \end{aligned} \tag{63}$$

The distribution function is

$$\begin{aligned} & F_T\left(R;b,c,R_l,R_u\right) \\ & = \frac{b^3\left(4^c\pi^c 3^{-c}\left(R^3\right)^c b^{-3c}c^c \mathrm{e}^{-\frac{4c\pi R^3}{3b^3}} + \Gamma(c+1) - \Gamma\left(c+1,\frac{4c\pi R^3}{3b^3}\right)\right)}{k} \\ & - \frac{b^3\left(4^c\pi^c 3^{-c}\left(R_l^3\right)^c b^{-3c}c^c \mathrm{e}^{-\frac{4c\pi R_l^3}{3b^3}} + \Gamma(c+1) - \Gamma\left(c+1,\frac{4c\pi R_l^3}{3b^3}\right)\right)}{k}, \end{aligned} \tag{64}$$

and the average value

$$\begin{aligned} \mu_T\left(b,c,R_l,R_u\right) = {} & \frac{1}{(6ck+2k)\pi^{1/3}} \times \left(3b^4c^{2/3}\Gamma\left(\frac{4}{3}+c,\frac{4c\pi R_l^3}{3b^3}\right)6^{1/3}\right. \\ & - 3b^4c^{2/3}\Gamma\left(\frac{4}{3}+c,\frac{4c\pi R_u^3}{3b^3}\right)6^{1/3} \\ & \left. - 6b^{-3c+3}c^{c+1}\pi^{c+\frac{1}{3}}4^c 3^{-c}\left(-\mathrm{e}^{-\frac{4c\pi R_u^3}{3b^3}}R_u^{3c+1} + \mathrm{e}^{-\frac{4c\pi R_l^3}{3b^3}}R_l^{3c+1}\right)\right). \end{aligned} \tag{65}$$

The median can be found by numerically solving the equation

$$F_T\left(R;b,c,R_l,R_u\right)=\frac{1}{2}. \tag{66}$$

The statistics for the truncated Kiang function in radius of the VAST catalog are given in Table 6, the high values of the probability do not mean overfitting or limitations in the data-model fit. Figure 8 reports the Planck 2018 and V^2/REVOLVER data.

5. Conclusions

Lognormal family

We tested the progressive increase in the number of parameters for the lognormal family here represented by the lognormal, the truncated lognormal, the mod-

ified lognormal and the modified lognormal with truncation on the radius distribution for cosmic voids as given by the VAST catalog. This increase produces progressive best fits, see the last column in Tables 1-4.

Table 6. Numerical values of D, the maximum distance between theoretical and observed DFs, P_{KS}, significance level, in the K-S test for the truncated Kiang distribution in radius, see Equation (62), and percent error for the numerical median, see Equation (66), for different cosmologies and methods.

Cosmology	Method	Parameters	D	P_{KS}	% error median
WMAP5	VoidFinder, $N=1184$	$b=26.35$ Mpc $c=2.33$	0.083	1.08×10^{-7}	3.025%
Planck 2018	V^2/VIDE, $N=531$	$b=24.27$ Mpc $c=0.173$	0.051	0.121	4.27%
Planck 2018	V^2/REVOLVER, $N=518$	$b=26.04$ Mpc $c=0.361$	0.0192	0.99	0.623%

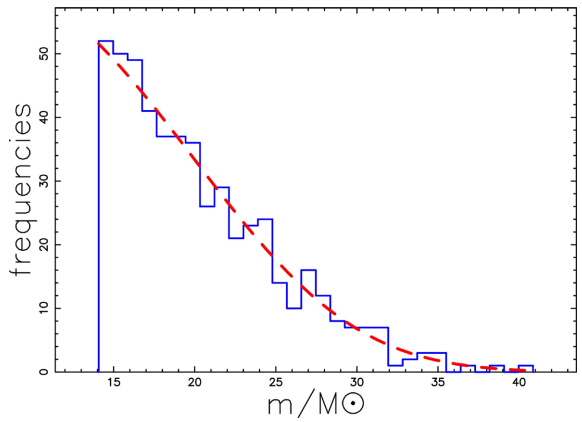

Figure 8. Histogram of distribution of voids as given by Planck 2018 and V^2/REVOLVER data (blue) with a superposition of the truncated Kiang distribution in radius when the number of bins, m, is 30 (red dashed line). Parameters as in Table 6. Vertical and horizontal axes have linear scales.

Kiang distribution

The introduction of a truncation in the Kiang distribution in radius increases the reliability of the fits, see the last column in Table 5 and Table 6. We also derived an approximation for the median of the Kiang distribution, see Equation (36), and for the Kiang distribution in radius, see Equation (52). The results for the VAST catalog of the median in terms of percent error are reported in the last column of Table 6. The positive results for the Kiang distribution require a refinement of the model, *i.e.* the stereological properties of the Voronoi diagrams.

Comparison of the two distributions

The modified lognormal distribution with truncation for the radius distribution in cosmic voids produces better results for two of the three cases analysed. The Kiang distribution in radius with truncation reports a surprising $P_{KS}=0.99$

which is the highest level of significance for the VAST catalog, see Table 4 and Table 6.

Maximum void

The Boötes void, sometimes called 'the great nothing' or great void, was discovered in 1987. Its radius is 62 Mpc [21], which is the maximum at the moment of writing. In our framework, this can be identified with the upper radius, R_u, of the truncated Kiang distribution in radius, see formula 63.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Douglass, K.A., Veyrat, D. and BenZvi, S. (2023) Updated Void Catalogs of the SDSS DR7 Main Sample. *The Astrophysical Journal Supplement Series*, **265**, Article 7. https://doi.org/10.3847/1538-4365/acabcf

[2] Arcari, S., Pinetti, E. and Fornengo, N. (2022) Got Plenty of Nothing: Cosmic Voids as a Probe of Particle Dark Matter. *Journal of Cosmology and Astroparticle Physics*, **2022**, Article 011. https://doi.org/10.1088/1475-7516/2022/11/011

[3] Wang, B.Y. and Pisani, A. (2024) Cosmology from One Galaxy in a Void? *The Astrophysical Journal Letters*, **970**, L32. https://doi.org/10.3847/2041-8213/ad5ffe

[4] Bromley, B.C. and Geller, M.J. (2025) Cosmology with Voids. *Journal of Cosmology and Astroparticle Physics*, **2025**, Article 011. https://doi.org/10.1088/1475-7516/2025/05/011

[5] Kiang, T. (1966) Random Fragmentation in Two and Three Dimensions. *Zeitschrift für Astrophysik*, **64**, Article 433.

[6] Evans, M., Hastings, N. and Peacock, B. (2000) Statistical Distributions.3rd Edition, John Wiley & Sons Inc.

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994) Continuous Univariate Distributions. Vol. 1, 2nd Edition, Wiley.

[8] Zaninetti, L. (2017) A Left and Right Truncated Lognormal Distribution for the Stars. *Advances in Astrophysics*, **2**, 197-213.

[9] Olver, F.W.J., Lozier, D.W., Boisvert, R.F. and Clark, C.W. (2010) NIST Handbook of Mathematical Functions. Cambridge University Press.

[10] Basu, S., Gil, M. and Auddy, S. (2015) The MLP Distribution: A Modified Lognormal Power-Law Model for the Stellar Initial Mass Function. *Monthly Notices of the Royal Astronomical Society*, **449**, 2413-2420. https://doi.org/10.1093/mnras/stv445

[11] Zaninetti, L. (2025) New Probability Distributions in Astrophysics: XIV. Truncation of the Modified Lognormal Distribution. *International Journal of Astronomy and Astrophysics*, **15**, 19-42. https://doi.org/10.4236/ijaa.2025.151003

[12] Kolmogoroff, A. (1941) Confidence Limits for an Unknown Distribution Function. *The Annals of Mathematical Statistics*, **12**, 461-463. https://doi.org/10.1214/aoms/1177731684

[13] Smirnov, N. (1948) Table for Estimating the Goodness of Fit of Empirical Distributions. *The Annals of Mathematical Statistics*, **19**, 279-281. https://doi.org/10.1214/aoms/1177730256

[14] Massey, F.J. (1951) The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of*

the American Statistical Association, **46**, Article 68. https://doi.org/10.2307/2280095

[15] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) Numerical Recipes in Fortran. The Art of Scientific Computing. Cambridge University Press.

[16] Pan, D.C., Vogeley, M.S., Hoyle, F., Choi, Y. and Park, C. (2012) Cosmic Voids in Sloan Digital Sky Survey Data Release 7. *Monthly Notices of the Royal Astronomical Society*, **421**, 926-934. https://doi.org/10.1111/j.1365-2966.2011.20197.x

[17] Varela, J., Betancort-Rijo, J., Trujillo, I. and Ricciardelli, E. (2012) The Orientation of Disk Galaxies around Large Cosmic Voids. *The Astrophysical Journal*, **744**, Article 82. https://doi.org/10.1088/0004-637x/744/2/82

[18] Mao, Q., Berlind, A.A., Scherrer, R.J., Neyrinck, M.C., Scoccimarro, R., Tinker, J.L., *et al.* (2017) A Cosmic Void Catalog of SDSS DR12 BOSS Galaxies. *The Astrophysical Journal*, **835**, Article 161. https://doi.org/10.3847/1538-4357/835/2/161

[19] Kiang, T. (1990) Private Communication.

[20] Kumar, S., Kurtz, S.K., Banavar, J.R. and Sharma, M.G. (1992) Properties of a Three-Dimensional Poisson-Voronoi Tesselation: A Monte Carlo Study. *Journal of Statistical Physics*, **67**, 523-551. https://doi.org/10.1007/bf01049719

[21] Kirshner, R.P., Oemler Jr., A., Schechter, P.L. and Shectman, S.A. (1987) A Survey of the Bootes Void. *The Astrophysical Journal*, **314**, Article 493. https://doi.org/10.1086/165080