A consistent approach to falsifying $\Lambda$CDM with rare galaxy clusters

Ian Harrison$^a$ and Shaun Hotchkiss$^b$

$^a$School of Physics & Astronomy, Cardiff University, 5 The Parade, Cardiff, CF24 3AA, United Kingdom
$^b$Department of Physics, University of Helsinki and Helsinki Institute of Physics, P.O. Box 64, FIN00014 University of Helsinki, Finland

E-mail: ian.harrison@astro.cf.ac.uk, shaun.hotchkiss@helsinki.fi

Received April 23, 2013
Revised June 10, 2013
Accepted June 15, 2013
Published July 15, 2013

Abstract. We consider methods with which to answer the question “is any observed galaxy cluster too unusual for $\Lambda$CDM?” After emphasising that many previous attempts to answer this question will overestimate the confidence level at which $\Lambda$CDM can be ruled out, we outline a consistent approach to these rare clusters, which allows the question to be answered. We define three statistical measures, each of which are sensitive to changes in cluster populations arising from different modifications to the cosmological model. We also use these properties to define the “equivalent mass at redshift zero” for a cluster — the mass of an equally unusual cluster today. This quantity is independent of the observational survey in which the cluster was found, which makes it an ideal proxy for ranking the relative unusualness of clusters detected by different surveys. These methods are then used on a comprehensive sample of observed galaxy clusters and we confirm that all are less than $2\sigma$ deviations from the $\Lambda$CDM expectation. Whereas we have only applied our method to galaxy clusters, it is applicable to any isolated, collapsed, halo. As motivation for future surveys, we also calculate where in the mass-redshift plane the rarest halo is most likely to be found, giving information as to which objects might be the most fruitful in the search for new physics.

Keywords: galaxy clusters, cluster counts
1 Introduction

The current concordance ΛCDM cosmology makes (via the halo mass function) definite predictions for the abundance of gravitationally bound dark matter haloes and how that abundance evolves with redshift. These haloes are visible to us as galaxy clusters, both through their baryonic matter content in the form of galaxies and hot, diffuse gas and directly via strong and weak gravitational lensing events. Because of the halo mass function’s sensitivity to cosmology, it has been useful in constraining parameters within ΛCDM, principally the power spectrum normalisation $\sigma_8$ and total matter fraction $\Omega_m$, [1, 2]. It has also been shown to be sensitive to many of the plausible modifications to concordance cosmology such as primordial non-Gaussianity [3], dark energy models [4, 5] and modifications to gravity [6–8].

Ongoing surveys are revealing what are expected to be the most massive galaxy clusters at ever increasing redshifts. These objects represent a tantalising possibility: because the high-mass tail of the halo mass function descends very steeply, the observation of even a single galaxy cluster which is massive enough at a given redshift has the potential to contradict the predictions of ΛCDM with high significance. Following early work [9–11], a number of statistical tests have been used to determine whether any of the clusters observed by recent experiments are in tension with the standard model of cosmology. These tests have fallen broadly into two groups: those which consider explicitly how likely a given cluster is to appear in a ΛCDM universe (which we will refer to as ‘rareness’ methods), and those which make...
use of Extreme Value Statistics (EVS). EVS methods predict the probability distribution function for the mass of the most-massive cluster expected within a given survey window and have generally found no tension with ΛCDM \cite{12-14} when applied correctly to an a priori defined survey window. In contrast, the rareness methods have chiefly considered the probability that a cluster could exist with both greater mass and redshift than those observed, either calculating tension explicitly \cite{9, 15-19} or by constructing ‘exclusion curves’ in the mass-redshift plane \cite{20-23}. However, it has been observed \cite{24} that these rareness analyses suffer from the (now continued) use of a biased statistical method, first introduced by \cite{9} and \cite{10}, which causes them to overestimate the amount of tension a particular observation may cause. Both \cite{24} and \cite{25} have analysed clusters using an un-biased rareness method, finding no tension with ΛCDM in accordance with the EVS methods.

Nonetheless, the idea that observations of individual clusters are capable of falsifying a given cosmological model remains a valid and intriguing possibility, even if no tension currently exists. Here we provide a consistent method to calculate the rareness of observed galaxy clusters, first quantifying how extreme each cluster is relative to others, before using the methods advocated in \cite{24} to determine the degree of tension they may cause with the predictions of ΛCDM. Our work here goes beyond that of \cite{24}, first by the inclusion of the effects of (cosmological) parameter uncertainty, secondly by the inclusion of the effects of (mass) measurement uncertainty and thirdly by examining a much more comprehensive list of clusters (including some discovered after \cite{24} was published). The code used to calculate the results in this paper is available at http://www.mv.helsinki.fi/home/hotchkis/rareness/, allowing our methods to be applied to new clusters observed in the future.

The paper is organised as follows. In section 2.1 we first define the methodology we use to calculate the expected number of clusters in a given region of the mass-redshift plane, this includes our choice of halo mass function and cosmological parameters. Then, in section 2.2 we review previous methods relating to galaxy clusters and reaffirm their problem of overestimating tension with the cosmological model. In section 3.1 we define three statistical measures of rareness which are sensitive to different modifications to ΛCDM. In section 3.2 we show where the rarest clusters are most likely to be observed, and in sections 3.3 and 3.4 we explain how observational uncertainties will be dealt with. Next, in section 4 we define the ‘equivalent mass at redshift zero’, by which clusters from different surveys may be compared and ranked. Finally, we apply these methods to a large sample of observed clusters in section 5. We both rank all the clusters according to their equivalent mass at redshift zero and, by making conservative assumptions about selection functions, we also set upper limits on the amount of tension these clusters provide with ΛCDM predictions. In this section we also discuss the correspondence between our result and those gained using Extreme Value Statistics. Finally, in section 6 we discuss our results and conclude.

2 Cosmology with rare galaxy clusters

2.1 Galaxy cluster abundance and cosmology

In the standard cosmological model with Gaussian initial conditions and hierarchical structure growth, high-mass galaxy clusters are expected to evolve from high peaks in the initial cold dark matter (CDM) density fluctuations. The smallest scales collapse first, before merging over time to form ever more massive CDM haloes, into which baryons fall to form galaxy clusters. Consequently, high mass clusters are expected to be very rare at early times, as reflected in the exponential steepness of the halo mass function \( n(m, z) \). The steepness of this
tail is also highly sensitive to the physical assumptions which go into the initial conditions and dynamical evolution of the dark matter over-density field, meaning the observation of even a single sufficiently extreme (in terms of both its mass and redshift) cluster has the potential to provide strong evidence against a particular cosmological model.

The number of galaxy clusters expected to occur in a survey window covering fraction of the sky $f_{\text{sky}}$ and sensitive to clusters with masses between $m_{\text{min}}$ and $m_{\text{max}}$ at redshifts between $z_{\text{min}}$ and $z_{\text{max}}$ is given by the integrated product of the halo mass function and volume element within this region:

$$\langle N \rangle = f_{\text{sky}} \int_{z_{\text{min}}}^{z_{\text{max}}} \int_{m_{\text{min}}}^{m_{\text{max}}} dz dM \frac{dV}{dM} \frac{dn(M,z)}{dz}.$$  \hspace{1cm} (2.1)

In real surveys the mass of a halo is not measured directly, but via proxies such as X-ray gas temperature $T_X$, galaxy velocity dispersion $\sigma_v$ or the thermal Sunyaev-Zel’dovich (tSZ) Compton-y. The realities of detecting these proxies mean that real surveys are not typically mass limited (although tSZ surveys approach this) and the use of absolute mass and redshift limits is a crude approximation to the real selection function. However, in this paper we will endeavour to be conservative with our approximate selection functions, providing lower limits on cluster detection probabilities. The methodology presented here can still be applied in the advantageous situation where the full selection function is known, and our conclusions are expected to be stable.

Throughout this work, the cosmology assumed is that described by the WMAP7+BAO+H0 ML parameters [26]. From these parameters we calculate the linear matter power spectrum $P(k)$ using the numerical Einstein-Boltzmann code CAMB\(^1\) and in turn the variance $\sigma^2(m,z)$, smoothed with a top hat window function $W(k;m)$ and evolved to a redshift of $z$ with the normalised linear growth function $D(z)$

$$\sigma^2(m,z) = D^2(z) \int_0^\infty \frac{dk}{2\pi} k^2 P(k) W^2(k;R).$$  \hspace{1cm} (2.2)

The calculated $\sigma(m,z)$ is then used in the version of the Tinker halo mass function [27]:

$$\frac{dn(m,z)}{dm} = A \left[ \left( \frac{\sigma}{b} \right)^{-a} + 1 \right] e^{-c/\sigma^2 \bar{\rho}_m} \frac{d\ln(\bar{\rho}_m)}{dm}.$$  \hspace{1cm} (2.3)

which includes parameters which evolve with redshift: $A = 0.186(1+z)^{-0.14}$, $a = 1.47(1+z)^{-0.06}$, $b = 2.57(1+z)^{-0.011}$, $c = 1.19$. This mass function has been well tested against large, high-resolution N-body simulations and has become the most frequently used in cosmological analyses.

### 2.2 Comparison with previous analyses

Many observable quantities are potentially available to classify galaxy clusters: halo mass, profile and concentration; redshift; population of galaxies (and their type, colour etc); gas temperature and many others. Values of these observables can be combined to define a statistic and, when an observation of a cluster is made, the value of the statistic for that observation can be calculated. If we then wish to use this statistic to do inference on our cosmological model then we need to calculate the probability distribution for this statistic. It

\(^1\)http://camb.info.
is then straightforward to determine how unlikely/rare a particular cluster would be in ΛCDM (and a given survey) by calculating the probability that any cluster could be observed with a value that exceeds the measured value of the statistic. This probability to exceed (PTE) is a direct measure of the tension an observation provides with ΛCDM. Here we summarise the previous work of [24] considering correct and incorrect ways in which to calculate this tension.

Many previous analyses wished to quantify whether some observed clusters were too massive or formed too early for ΛCDM. The statistic typically used in these analyses [9, 15–19] is the Poisson probability of observing at least one cluster (with observed mass $\hat{m}$ and redshift $\hat{z}$ denoted by hats) with both greater mass and redshift than the one which has been observed:

$$\hat{R}_{\hat{m}, \hat{z}} = 1 - \exp (-\langle N_{\hat{m}, \hat{z}} \rangle),$$

In these analyses the value of $\hat{R}_{\hat{m}, \hat{z}}$ was taken, directly, as the degree of tension a cluster provides with ΛCDM. However, as first pointed out by Fergus Simpson and later expounded in [24], using $\hat{R}_{\hat{m}, \hat{z}}$ as a PTE will lead to incorrect conclusions because it ignores the fact that (observable) clusters at lower redshift and higher mass or higher redshift and lower mass would have values of this $R_{\hat{m}, \hat{z}}$ statistic equal to or lower than what was observed. As explained in [24] the true probability of an observation exceeding $R_{\hat{m}, \hat{z}}$ is necessarily greater than the value of $R_{\hat{m}, \hat{z}}$, meaning a low value of $R_{\hat{m}, \hat{z}}$ is not an uncommon property for the most extreme galaxy clusters expected in a ΛCDM universe. The correct probability can be found by finding the line in the mass-redshift plane of clusters which have an equal $\langle N_{\hat{m}, \hat{z}} \rangle$ and calculating the probability of observing a galaxy cluster anywhere above this line.

This flaw in calibration also exists in the exclusion curves calculated by Mortonson et al [28] and hence in subsequent uses of these curves in the literature [20–23]. The defining property of an ‘exclusion curve’ is that observation of a single cluster above the curve will rule out a ΛCDM cosmology at the corresponding confidence level, meaning for an 100α% exclusion curve we should expect to observe a cluster above the line only 100(1 − α)% of the time (i.e. because of a random fluctuation caused by sample variance). The curves from [28] do not obey this property. Figure 1 shows 100 Monte-Carlo realisations of halo masses within

**Figure 1.** High-mass clusters from 100 Monte-Carlo realisations of the WMAP7 cosmology, plotted along with exclusion curves from Mortonson et al (MHH) [28] and this paper. As can be seen, significantly more than the expected 5 clusters lie above the Mortonson et al 95% exclusion curve.
Figure 2. Fraction of WMAP7 cosmologies ruled out against confidence level, for both the Mortonson et al. (MHH) [28] exclusion curves and those presented in this paper. Joint-CL refers to the choice $s = p$ in [28] and Mean-CL $s = 0.95, p = 0.5$ (corresponding to sample uncertainty only, with $\sigma_8$ and $\Omega_m$ at their maximum likelihood value). As can be seen, Mortonson et al. curves rule out the underlying cosmology to an erroneously high confidence level, with the correct behaviour represented by the solid straight line and followed closely by the results of this work.

A WMAP7 cosmology, along with a 95% confidence level (CL) exclusion curve from [28]. As can be seen, whilst there are $\sim 5$ clusters in the region $> m > z$ of each point on the line (as is to be expected from their construction), there are significantly more than the expected five clusters lying above the curve in total, a number which increases as more of the mass-redshift-$f_{\text{sky}}$ region is probed. Figure 2 further emphasises this; plotted is $\alpha$ against the fraction of Monte Carlo realisations of a $\Lambda$CDM cosmology that contain a cluster that lies above a 100$\alpha$% exclusion curve. It can clearly be seen that Mortonson et al. curves do not follow the behaviour required of correctly calibrated exclusion curves, represented by the solid straight line (i.e. that an 100$\alpha$% CL-breaking cluster is found in 100(1 - $\alpha$)% of realisations). They instead show a significant hump, ruling out $\Lambda$CDM at a high confidence level in a high fraction of realisations. Also displayed in figures 1 and 2 are the results gained using the analysis in this work, which do behave correctly as exclusion curves.

3 Calculating the rareness of an observed cluster

As discussed in section 2.2, we can correctly estimate the tension a galaxy cluster may be in with a given cosmological model, and define the related exclusion curves, by defining a rareness statistic, finding the contour of constant rareness and then calculating the probability of making an observation of a cluster anywhere above this line. As well as being correctly calibrated it is necessary to draw such curves in a physically motivated manner, as discussed in [25]. Here, we identify three separate physically-motivated statistics which will later be used to calculate the PTE for observed clusters in a given cosmology and construct correctly-calibrated exclusion curves.

3.1 Three statistics to measure extremeness

3.1.1 Expected number with greater mass and redshift $> m > z$

Even though it has been used incorrectly in previous works, the statistic defined by the expected number of clusters in a region with both greater mass and redshift:

$$\langle N_{>m>z} \rangle = \int_z^\infty \int_m^\infty dz \ dM \ \frac{dV \ dn(M,z)}{dM}.$$

(3.1)
is intuitively physical and may be used in a correctly calibrated way, by finding the probability of observing a cluster anywhere above a line of constant $\langle N_{>m>z} \rangle$. However, $\langle N_{>m>z} \rangle$ is sensitive to modifications in background expansion, growth and initial conditions, meaning well-motivated modifications to $\Lambda$CDM are not separable.

### 3.1.2 Expected number with greater initial peak height $> \nu$

Galaxy clusters are expected to form at the location of high peaks in the distribution of primordial density perturbations, seeded by inflation. For a given fixed background expansion and growth law, changes in the CDM initial conditions, such as the widely-considered introduction of primordial non-Gaussianity (often parameterised by positive $f_{\text{NL}}$), would produce more rare clusters from higher peaks. We thus also consider the peak height from which a cluster is expected to have formed:

$$\nu(m,z) \propto \frac{1}{D_+(z)\sigma(m)},$$

as a physically-motivated rareness statistic.

### 3.1.3 Expected number with greater mass, per unit volume $> m dV$

Finally, we also use the statistic defined by the expected number of more massive clusters per unit volume at a given redshift:

$$\langle N_{>m dV} \rangle = \left[ \int_m^{\infty} dM \frac{dn(M,z)}{dM} \right].$$

This definition has the advantage that it fairly weights all clusters at high masses, even those which come from low-volume regions in the redshift dimension.

### 3.2 Expected masses and redshifts of the rarest clusters

We may also consider where in the mass-redshift plane we expect the rarest observed cluster to be found. Answering this question can give information about where cluster surveys can be most productively targeted, or indeed what kind of objects may be most sensitive probes of the tail of the halo mass function. The plots in figure 3 show the probability distribution for the location in the mass-redshift plane of the rarest observed cluster, for each statistic. The rarest cluster according to the $> \nu$ measure is always most likely to appear at the highest specified redshift ($z = 4$ for these plots), whilst the rarest cluster according to the $> m > z$ and $> m dV$ measures are most likely to be observed at $z \approx 1$ and $z \approx 2.5$ respectively.

An interesting inference can be made from the $> \nu$ plot with regards to attempts to constrain primordial non-Gaussianity with rare objects. The modification to the halo mass function caused by primordial non-Gaussianity depends almost entirely on $\nu$. The tendency of surveys to be most likely to find their rarest objects, according to the $\nu$ definitions, at the highest possible redshift (and lower absolute masses) indicates that it is perhaps not galaxy clusters but higher redshift events such as lensing arcs and quasars which may prove the most sensitive probes of non-Gaussianity.

---

1We thank Raul Angulo (private correspondence) for motivating this definition.
3.3 Dealing with parameter uncertainty

If we are seeking to test a cosmological model, it is necessary to take into account the uncertainties on the values of the parameters within the model. As long as we do not introduce biases or make poor assumptions, we wish to be as sensitive to new physics as possible. A statistically robust way to treat parameter uncertainty is to simply marginalise the probability to exceed \( \hat{R} \) over available prior constraints on the cosmological parameters:

\[
\hat{R} = \int d\vec{\Lambda} \, \hat{R}(\vec{\Lambda}) \Pi(\vec{\Lambda}),
\]

where \( \vec{\Lambda} \) is the full set of cosmological parameters and \( \Pi(\vec{\Lambda}) \) is the available prior probability distribution for those parameters. Of the standard model’s cosmological parameters, it is the normalisation of the linear matter power spectrum \( \sigma_8 \) which has by far the most significant influence on cluster abundance. For the analysis below we use a Gaussian prior on \( \sigma_8 \) from [26], with a mean of 0.811 and standard deviation of 0.03.

3.4 Dealing with measurement uncertainty

A final consideration to be made when examining high-mass galaxy clusters is the expected posterior distribution for the cluster mass \( P(m|\hat{m}) \), for which we follow the treatment of [29]. Here, \( \hat{m} \) is to be understood as the full set of observable parameters relating to the measurement of a cluster’s mass. In Bayesian reasoning, the posterior probability distribution function for the cluster mass \( m \) in terms of an observable \( \hat{m} \) is proportional to the product of the likelihood of the observation \( L(\hat{m}) \) and a prior probability distribution for mass \( \Pi(m) \).

Here, \( L(\hat{m}) \) is taken to be the observed cluster mass and error region, with either a normal or log-normal form. Because the prior distribution on cluster mass (the halo mass function) varies significantly over the width of this likelihood, its effect must be taken into account. This effect constitutes the classical Eddington bias for number counts: because there are significantly more clusters in lower mass bins which may upscatter into higher bins than there are high mass clusters to scatter downwards, we must adjust our number counts accordingly. This allows us to calculate the posterior distribution for the true mass of a galaxy cluster:

\[
P(m|\hat{m})dm \propto \frac{dN(m,z)}{dmdz} L(\hat{m})dm.
\]
The PTE values given here are then calculated by marginalising over their values for the
support of this distribution. This method will give the correct posterior mass uncertainty for
a ΛCDM prior if and only if the original quoted observable uncertainties are the statistically
correct posterior mass uncertainties obtained assuming a uniform/no prior on cluster mass.

4 Ranking clusters with equivalent mass at redshift zero

Once a statistic has been defined, we can (as suggested in [24]) gain an intuitive understanding
of how extreme an observed cluster is by calculating how massive a cluster at redshift zero
would need to be in order to have the same value of this statistic. We will denote this by
$m|_0$, the ‘equivalent mass at redshift zero’. Unlike the probability that a cluster could be
detected in a given survey, $m|_0$ is an intrinsic property of each cluster and does not depend
in any way on the depth, region or any other property of the survey it was selected from.
This allows for a comparison (or even a ranking) of the extremeness of objects detected in
different surveys and at different redshifts.

Figure 4 shows contours in the mass-redshift plane on which clusters will have equal
values of the three statistics defined in section 3.1. Where these contours intersect with
the mass axis is $m|_0$. As can be seen, the different definitions do not map points onto $m|_0$ in the
same way. For instance, the $\nu$ definition will assign the largest $m|_0$ to the deepest fluctuation
in the initial density field, irrespective of how the volume expansion proceeds between that
epoch and $z = 0$, meaning they appear as steeper contours on the mass-redshift plane. In
contrast, the tendency of the $\langle N_{>m|_0} \rangle$ measure to downweight very low redshift clusters
because of the larger volume element at $z \lesssim 0.3$ can be seen in the flattening of the contours
at these low redshifts.

5 Rareness and ranking of currently observed clusters

In this section we consider a large number of cluster observations and apply our methodology;
we first calculate $m|_0$ for each to find which are the most extreme objects before finding the
tension each observations represents with the standard cosmological model.
In order to calculate this tension we are required to find (as described in section 2.2) the probability for a particular observational survey to make any observation at least as rare as the detected galaxy clusters: the PTE. This requires knowledge of the survey selection function — as a survey covers more of the sky and more of the mass-redshift plane it surveys more objects, increasing the number of objects which may be found with a given rareness. Here, we choose to conservatively set lower limits on the PTE (which correspond to upper limits on tension with $\Lambda$CDM) by choosing suitable approximate selection functions. We do this by choosing the minimal survey window in mass-redshift space in which the cluster may have been found, as defined in table 1. We do this by considering only the complete (i.e. where the probability of detection $\rightarrow 1$) region of the survey, choosing high values of $m_{\text{min}}$ and low values of $z_{\text{max}}$ for each survey and only considering $f_{\text{sky}}$ for that particular survey. We also assume that the probability a cluster could exist in a region of the mass-redshift plane to be Poisson distributed (a good approximation for very high-mass galaxy clusters). A more sophisticated analysis would be possible on a per-survey basis, taking into account the full selection functions, such as the one which has been carried out by [30], who perform a correctly-calibrated rareness analysis using simulations of their observational survey to compare with the observed cluster.

In addition to this, estimation of cluster masses is a procedure fraught with uncertainty. It has been found both observationally (see [31] and references therein) and in N-body simulations [32] that masses (and ordering of most-massive clusters) estimated using different proxies are frequently inconsistent with each other. Further uncertainty occurs when converting between mass definitions for comparison with halo mass functions: both a halo profile (frequently NFW) and a mass-concentration relation must be assumed, both of which must be calibrated using N-body simulations. Such considerations are outside the scope of this paper. Here we choose to search the literature for published estimations of cluster masses and take them ‘at face value’. This choice naively ignores differences between survey mass proxies and sensitivities, which may in reality widen published error estimates, and all of our conclusions are predicated on this naivety. However, where robust estimates on cluster mass and uncertainty are available our methods will remain robust.

5.1 Cluster catalogue

Table 2 shows the list of papers used to construct our cluster catalogue. In total 2334 cluster mass estimations were included, where measurements in multiple proxies were allowed. As mentioned above, the mass uncertainties on each method were taken to be those given by each paper and were assumed to be normally distributed where error regions were symmetric and log-normally distributed when asymmetric. For the MCXC catalogue [33], where no error estimates are given, a log-normal error distribution with $\sigma_{\ln m} = 0.2$ was assumed, as is fairly typical for X-ray observations of clusters.

All cluster masses are converted to $m_{200m}$ (the mass which is within the cluster region 200 times the average density of the Universe) assuming an NFW halo profile, with a single concentration parameter $c$, which is calculated using the concentration-mass relation of [34] and WMAP7+BAO+H0 ML parameters from [26]. Where multiple cluster mass estimations appeared in the top-ten of $m_{|0}$, the observation with the smallest error region was used.

5.2 Rarest and most-massive clusters

Tables 3–5 show the clusters with the ten highest values of $m_{|0}$ calculated using each of the three statistics defined in section 3.1 and the PTE for that cluster in the relevant survey.
Table 1. The approximate survey selection functions used to calculate PTE values for the cluster catalogue.

| Survey   | $A_{sky}[\text{deg}^2]$ | $m_{\text{min}}[M_{\odot}h^{-1}]$ | $z_{\text{min}}$ | $z_{\text{max}}$ |
|----------|--------------------------|----------------------------------|------------------|------------------|
| ACT      | 755                      | $8 \times 10^{14}$               | 0.3              | 6.0              |
| SPT      | 2500                     | $8 \times 10^{14}$               | 0.3              | 6.0              |
| XMM      | 80                       | $4 \times 10^{14}$               | 0.9              | 1.5              |
| MACS     | 22735                    | $8 \times 10^{14}$               | 0.3              | 0.7              |
| WARPS    | 72                       | $8 \times 10^{14}$               | 0.0              | 0.6              |
| PLCK     | 41253                    | $1 \times 10^{15}$               | 0.3              | 6.0              |
| RDCS     | 50                       | $8 \times 10^{14}$               | 0.05             | 0.8              |
| LoCuSS   | 32085                    | $3 \times 10^{14}$               | 0.15             | 0.3              |

Table 2. Papers used to compile cluster catalogue. $N_{Cl}$ is the number of clusters contained within each paper and observable mass proxies are WL, Weak Lensing; $N_{200}$, cluster richness; $T_X$, X-ray gas temperature; $Y_X$, integrated X-ray flux; $Y_{SZ}$, integrated Compton-$y$; $\sigma_v$, velocity dispersion; $L_X$, X-ray luminosity and $\zeta$, SPT matched filter signal-to-noise.

Even with our conservative treatment of selection functions, the lowest PTE value is found to be as large as 0.07, for the cluster CLJ1226+3332. Note, however, that we have examined eight independent surveys. The probability that the smallest PTE in all eight surveys is greater than or equal to 0.07 is given by $(1 - 0.07)^8 = 0.56$. Therefore, if we live in a $\Lambda$CDM
Figure 5. Currently observed clusters in tables 3–5 and exclusion lines (using the $mdV$ measure described in the text) corresponding to the ACT and ACT+SPT survey areas and sets of clusters.

| Cluster            | $m_{200m}$       | $z$  | ref. (Proxy) | $m_{16}$ | Survey | PTE |
|--------------------|------------------|------|--------------|----------|--------|-----|
| ACT-CLJ0102-4915   | 1.51 ± 0.22      | 0.87 | [22] (Combined) | 6.53     | ACT    | 0.48|
| ACT-CLJ2317-0204   | 1.82 ± 0.28      | 0.705| [46] (Y$_{SZ}$) | 6.02     | ACT    | 0.60|
| SPT-CLJ2106-5844   | 0.893 ± 0.148    | 1.13 | [41] (Combined) | 6.01     | SPT    | 0.85|
| SPT-CLJ0205-5829   | 0.617 ± 0.096    | 1.32 | [30] (Combined) | 5.70     | SPT    | 0.90|
| ACT-CLJ0012-0046   | 0.534 ± 0.137    | 1.36 | [46] (Y$_{SZ}$) | 5.32     | ACT    | 0.73|
| XMMUJ2235-2557     | 0.558$^{+0.129}_{-0.10}$ | 1.39 | [17] (WL)   | 5.08     | XMM    | 0.40|
| MACSJ0417.5-1154   | 2.86$^{+0.62}_{-0.50}$ | 0.44 | [33] (L$_{X}$) | 4.85     | MACS   | 0.97|
| CLJ1226+3332       | 1.12$^{+0.19}_{-0.16}$ | 0.89 | [17] (WL)   | 4.55     | WARPS  | 0.07|
| ACT-CLJ0546-5345   | 0.709 ± 0.121    | 1.07 | [46] (Y$_{SZ}$) | 4.54     | ACT    | 0.89|
| PLCKG266.6-27.3    | 0.899 ± 0.067    | 0.94 | [42] (Y$_{X}$) | 4.49     | Planck | > 0.99|

Table 3. The 10 clusters with highest $m_{16}$ for the $\nu$ method described in the text. Masses marked with a * are uniform prior masses and were corrected using the procedure described in section 3.4 before being used. All masses are in $10^{15} M_{\odot} h^{-1}$.

In order to demonstrate how these PTE relate to exclusion curves in figure 5 we show the relevant plot for the ACT and SPT surveys. These correctly-calibrated exclusion curves were calculated using the ACT only and ACT+SPT survey regions and the clusters appearing in the top-ten tables are plotted. As can be seen, none of the clusters breaks the 66% exclusion curve.

5.3 Extreme Value Statistics of $m_{16}$

Extreme Value Statistics (EVS) make predictions for the probability distribution function of sample extrema and have been used in the context of high-mass clusters by predicting the distribution for the most-massive cluster in a given survey region. The PTEs calculated in the universe, there is at least a 44% chance that the smallest PTE in our tables will be less than or equal to 0.07. Even with our very conservative treatment of selection functions, designed to make clusters appear rarer than they actually are, none of the clusters or surveys we have considered indicate any tension with $\Lambda$CDM.

In order to demonstrate how these PTE relate to exclusion curves in figure 5 we show the relevant plot for the ACT and SPT surveys. These correctly-calibrated exclusion curves were calculated using the ACT only and ACT+SPT survey regions and the clusters appearing in the top-ten tables are plotted. As can be seen, none of the clusters breaks the 66% exclusion curve.
section 5 represent the probability that at least one galaxy cluster exists in a survey region above a line of constant \(m|_0\). This is \(1 - P_0\), where \(P_0\) is the void probability that no clusters exist in the region above the line of constant rareness (constant \(m|_0\)). As emphasised by [55], this void probability is the same distribution as the EVS cumulative distribution function, the probability that the highest \(m|_0\) in the survey region is less than or equal to the observed value. Forming the EVS from \(m|_0\) contours in this way has the advantage of considering the entire survey region within a single distribution, unlike [13] which predicts the distribution for the most-massive cluster in narrow redshift bins with \(f_{\text{sky}} = 1\) or [12] which considers smaller \(f_{\text{sky}}\) but large redshift bins. However, the EVS distribution for \(m|_0\) cannot be written down directly as it can be for the cluster mass \(m\) only; here we obtain the distribution numerically by simulating \(10^5\) highest \(m|_0\)s and fitting a Generalised Extreme Value (GEV, the limiting value for all extreme value distributions, see e.g. [56]) distribution to the results. Figure 6 shows this procedure performed for the ACT survey selection function defined in table 1, with the location of \(m|_0\) for the most extreme object in the survey, ACT-CLJ0102-4915, also shown. As expected the probability for \(m|_0\) to exceed the observed value on the EVS plot matches the PTE calculated in the rareness approach.

---

| Cluster       | \(m_{200m}\) | \(z\) | ref. (Proxy) | \(m|_0^{>m>z}\)          | Survey | PTE |
|---------------|-------------|------|-------------|--------------------------|--------|-----|
| ACT-CLJ0102-4915 | 1.51 ± 0.22 | 0.87 | [22] (Combined) | 3.43 ACT 0.48          |        |     |
| ACT-CLJ2317-0204 | 1.82 ± 0.28 | 0.705 | [46] (YSZ) | 3.34 ACT 0.50          |        |     |
| MACSJ0417.5-1154 | 2.86^{+0.62}_{-0.50} | 0.44 | [33] (LX) | 3.03 MACS 0.95          |        |     |
| SPT-CLJ2106-5844 | 0.893 ± 0.148 | 1.13 | [41] (Combined) | 2.76 SPT 0.92         |        |     |
| MACSJ2243.3-0935 | 2.23^{+0.48}_{-0.39} | 0.45 | [33] (LX) | 2.45 MACS 0.99          |        |     |
| MACSJ2211.7-0349 | 2.36^{+0.51}_{-0.41} | 0.40 | [33] (LX) | 2.42 MACS 0.99          |        |     |
| SPT-CLJ0205-5829 | 0.617 ± 0.096 | 1.32 | [30] (Combined) | 2.38 SPT 0.97         |        |     |
| MACSJ0308.9+2645 | 1.59^{+0.56}_{-0.47} | 0.36 | [33] (LX) | 2.23 MACS > 0.99        |        |     |
| ACT-CLJ0012-0046 | 0.534 ± 0.137 | 1.36 | [46] (YSZ) | 2.14 ACT 0.89          |        |     |
| CLJ1226+3332 | 1.12^{+0.19}_{-0.16} | 0.89 | [17] (WL) | 2.13 WARPS 0.26         |        |     |

Table 4. The 10 clusters with highest \(m|_0\) for the \(> m > z\) method described in the text. Masses marked with a * are uniform prior masses and were corrected using the procedure described in section 3.4 before being used. All masses are in \(10^{15} M_\odot h^{-1}\).

| Cluster       | \(m_{200m}\) | \(z\) | ref. (Proxy) | \(m|_0^{>mdV}\)          | Survey | PTE |
|---------------|-------------|------|-------------|--------------------------|--------|-----|
| ACT-CLJ0102-4915 | 1.51 ± 0.22 | 0.87 | [22] (Combined) | 5.49 ACT 0.46          |        |     |
| ACT-CLJ2317-0204 | 1.82 ± 0.28 | 0.705 | [46] (YSZ) | 5.20 ACT 0.53          |        |     |
| SPT-CLJ2106-5844 | 0.893 ± 0.148 | 1.13 | [41] (Combined) | 4.70 SPT 0.89         |        |     |
| MACSJ0417.5-1154 | 2.86^{+0.64}_{-0.59} | 0.44 | [33] (LX) | 4.37 MACS 0.96          |        |     |
| SPT-CLJ0205-5829 | 0.617 ± 0.096 | 1.32 | [30] (Combined) | 4.21 SPT 0.95         |        |     |
| ACT-CLJ0012-0046 | 0.534 ± 0.137 | 1.36 | [46] (YSZ) | 3.84 ACT 0.83          |        |     |
| MACSJ2243.3-0935 | 2.23^{+0.48}_{-0.39} | 0.45 | [33] (LX) | 3.60 MACS 0.99          |        |     |
| CLJ1226+3332 | 1.12^{+0.19}_{-0.16} | 0.89 | [17] (WL) | 3.60 WARPS 0.13         |        |     |
| XMMUJ2235-2557 | 0.558^{+0.012}_{-0.017} | 1.39 | [17] (WL) | 3.58 XMM 0.47         |        |     |
| MACSJ2211.7-0349 | 2.36^{+0.51}_{-0.41} | 0.40 | [33] (LX) | 3.48 MACS > 0.99        |        |     |

Table 5. The 10 clusters with highest \(m|_0\) for the \(> mdV\) method described in the text. Masses marked with a * are uniform prior masses and were corrected using the procedure described in section 3.4 before being used. All masses are in \(10^{15} M_\odot h^{-1}\).
6 Discussion and Conclusions

In this paper we have considered an unbiased, consistent treatment of rare galaxy clusters. Because previous considerations of cluster rareness have frequently fallen foul of uncalibrated statistics that overestimate the amount of tension a given observation is in with the ΛCDM theory, we have been careful in defining the probabilities we are calculating to avoid a posteriori effects. We defined three statistics by considering three physically motivated properties of a cluster which may be sensitive to modifications in the underlying cosmology: the expected number of clusters at greater mass and redshift; the peak height $\nu$ in the CDM over-density from which the cluster grew; and the expected number of clusters with a greater mass, per unit volume. Using these statistics we calculated the probability that a defined survey would have observed a cluster as rare as an observed cluster or rarer, anywhere in the mass-redshift plane, i.e. the probability to exceed (PTE) the observed value of the statistic. This is a crucial difference to most earlier methods, wherein only clusters which had greater mass and redshift were considered as more extreme than the one which had been observed.

We have also considered where in the mass-redshift plane the most extreme clusters in a survey are most likely to reside. This provided us with an interesting result: for the $\nu$ statistic, which is sensitive to primordial non-Gaussianity, the most unusual cluster is always found at the highest redshift available to the survey, meaning that, in principle, higher-redshift objects (i.e. quasars, lensing arcs or gamma-ray bursts as opposed to galaxy clusters) are potentially the more sensitive probes of non-Gaussianity in large scale structure.

We also discussed a method to rank clusters between different surveys. This is $m|_0$, the equivalent mass at redshift zero. That is, the mass of the notional cluster at $z = 0$ which has the same value of the chosen statistic, for an observed cluster. The value of $m|_0$ is an intrinsic property of each cluster and does not depend at all on the survey in which it was found, meaning it is an ideal proxy for categorising and ranking clusters according to their extremeness, even when they have been detected in different surveys and at different redshifts. In fact, this method is immediately generalisable to any isolated and collapsed halo for which a reliable mass measurement can be obtained, which would even allow us to compare and rank the relative “extremeness” of entirely different objects in a self-consistent way.

Finally, we have conducted a systematic review of cluster mass estimations in the literature. Using conservative approximations to survey selection functions and an ‘at face value’
approach to published error estimates, we have calculated the expected PTE for each cluster in its observational survey, finding that none are rarer than the rarest cluster expected in some 7% of surveys in $\Lambda$CDM universes. As we have examined eight separate surveys, this value is entirely unremarkable.

To facilitate future estimates of galaxy cluster rareness we have made a numerical code available at: http://www.mv.helsinki.fi/home/hotchkis/rareness/. This code will calculate $m_{j0}$, PTE and a set of exclusion curves for any sets of clusters that are subsequently observed.

Acknowledgments

Ian Harrison receives an STFC studentship, thanks Peter Coles and Chris Messenger for useful discussions and pays tribute to Leonid Grishchuk for inspiring clear and rigorous science. Shaun Hotchkiss acknowledges support from the Academy of Finland grant 131454. Both authors thank Raul Angulo for motivating the $>mdV$ definition of extremeness used in this paper.

References

[1] A. Mantz, S.W. Allen, D. Rapetti and H. Ebeling, The Observed Growth of Massive Galaxy Clusters I: Statistical Methods and Cosmological Constraints, Mon. Not. Roy. Astron. Soc. 406 (2010) 1759 [arXiv:0909.3098] [nSPIRE].
[2] PLANCK collaboration, P. Ade et al., Planck 2013 results. XX. Cosmology from Sunyaev-Zel’dovich cluster counts, arXiv:1303.5080 [nSPIRE].
[3] S. Matarrese, L. Verde and R. Jimenez, The Abundance of high-redshift objects as a probe of non-Gaussian initial conditions, Astrophys. J. 541 (2000) 10 [astro-ph/0001366] [nSPIRE].
[4] J. Weller, R. Battye and R. Kneissl, Constraining dark energy with Sunyaev-Zel’dovich cluster surveys, Phys. Rev. Lett. 88 (2002) 231301 [astro-ph/0110353] [nSPIRE].
[5] M. Baldi, Early massive clusters and the bouncing coupled dark energy, Mon. Not. Roy. Astron. Soc. 420 (2012) 430 [arXiv:1107.5049] [nSPIRE].
[6] F. Schmidt, M.V. Lima, H. Oyaizu and W. Hu, Non-linear Evolution of $f(R)$ Cosmologies III: Halo Statistics, Phys. Rev. D 79 (2009) 083518 [arXiv:0812.0545] [nSPIRE].
[7] S. Ferraro, F. Schmidt and W. Hu, Cluster Abundance in $f(R)$ Gravity Models, Phys. Rev. D 83 (2011) 063503 [arXiv:1011.0992] [nSPIRE].
[8] L. Lombriser, A. Slosar, U. Seljak and W. Hu, Constraints on $f(R)$ gravity from probing the large-scale structure, Phys. Rev. D 85 (2012) 124038 [arXiv:1003.3009] [nSPIRE].
[9] R. Jimenez and L. Verde, Implications for Primordial Non-Gaussianity ($f_{NL}$) from weak lensing masses of high-z galaxy clusters, Phys. Rev. D 80 (2009) 127302 [arXiv:0909.0403] [nSPIRE].
[10] D.E. Holz and S. Perlmutter, The most massive objects in the Universe, Astrophys. J. 755 (2012) L36 [arXiv:1004.5349] [nSPIRE].
[11] S. Colombi, O. Davis, J. Devriendt, S. Prunet and J. Silk, Extreme value statistics of smooth random Gaussian fields, Mon. Not. Roy. Astron. Soc. 414 (2011) 2436 [arXiv:1102.5707] [nSPIRE].
[12] J.-C. Waizmann, S. Ettori and L. Moscardini, An application of extreme value statistics to the most massive galaxy clusters at low and high redshifts, Mon. Not. Roy. Astron. Soc. 420 (2012) 1754 [arXiv:1109.4820] [nSPIRE].
[13] I. Harrison and P. Coles, "Testing Cosmology with Extreme Galaxy Clusters," Mon. Not. Roy. Astron. Soc. 421 (2012) L19 [arXiv:1111.1184] [SPIRE].

[14] S. Chongchitnan and J. Silk, "Primordial Non-Gaussianity and Extreme-Value Statistics of Galaxy Clusters," Phys. Rev. D 85 (2012) 063508 [arXiv:1107.5617] [SPIRE].

[15] M. Jee, P. Rosati, H. Ford, K. Dawson, C. Lidman et al., "Hubble Space Telescope Weak-lensing Study of the Galaxy Cluster XMMU J2235.3-2557 at z=1.4: A Surprisingly Massive Galaxy Cluster when the Universe is One-third of its Current Age," Astrophys. J. 704 (2009) 672 [arXiv:0908.3897] [SPIRE].

[16] L. Cayon, C. Gordon and J. Silk, "Probability of the most massive cluster under non-Gaussian initial conditions," Mon. Not. Roy. Astron. Soc. 415 (2011) 849 [arXiv:1006.1950] [SPIRE].

[17] M. Jee, K. Dawson, H. Hoekstra, S. Perlmutter, P. Rosati et al., "Scaling Relations and Overabundance of Massive Clusters at z > 1 from Weak-Lensing Studies with HST," Astrophys. J. 704 (2009) 672 [arXiv:0908.3897] [SPIRE].

[18] B. Hoyle, R. Jimenez and L. Verde, "Implications of multiple high-redshift galaxy clusters," Phys. Rev. D 83 (2011) 103502 [arXiv:1009.3884] [SPIRE].

[19] K. Enqvist, S. Hotchkiss and O. Taanila, "Estimating $f_{NL}$ and $g_{NL}$ from Massive High-Redshift Galaxy Clusters," JCAP 04 (2011) 017 [arXiv:1012.2732] [SPIRE].

[20] R. Williamson, B. Benson, F. High, K. Vanderlinde, P. Ade et al., "An SZ-selected sample of the most massive galaxy clusters in the 2500-square-degree South Pole Telescope survey," Astrophys. J. 738 (2011) 139 [arXiv:1105.3186] [SPIRE].

[21] M. Brodwin, A. Gonzalez, S. Stanford, T. Plagge, D. Marrone et al., "IDCS J1426.5+3508: Sunyaev-Zel’dovich Measurement of a Massive IR-selected Cluster at z=1.75," Astrophys. J. 753 (2012) 162 [arXiv:1105.3630] [SPIRE].

[22] F. Menanteau, J.P. Hughes, C. Sifon, M. Hilton, J. Gonzalez et al., "The Atacama Cosmology Telescope: ACT-CL J0102-4215 ‘El Gordo’, a Massive Merging Cluster at Redshift 0.87," Astrophys. J. 748 (2012) 7 [arXiv:1101.1290] [SPIRE].

[23] F. Menanteau, C. Sifon, L.F. Barrientos, N. Battaglia, J.R. Bond et al., "The Atacama Cosmology Telescope: Physical Properties of Sunyaev-Zel’dovich Effect Clusters on the Celestial Equator," Astrophys. J. 765 (2013) 67 [arXiv:1210.4048] [SPIRE].

[24] B. Stalder, J. Ruel, R. Suhada, M. Brodwin, K. Aird et al., "SPT-CL J0205-5829: A z = 1.32 evolved massive galaxy cluster in the south pole telescope Sunyaev-Zel’dovich effect survey," Astrophys. J. 763 (2013) 93 [arXiv:1205.6478] [SPIRE].

[25] B. Stalder, J. Ruel, R. Suhada, M. Brodwin, K. Aird et al., "SPT-CL J0205-5829: A z = 1.32 evolved massive galaxy cluster in the south pole telescope Sunyaev-Zel’dovich effect survey," Astrophys. J. 763 (2013) 93 [arXiv:1205.6478] [SPIRE].
[31] E. Rozo, J.G. Bartlett, A.E. Evrard and E.S. Rykoff, Closing the Loop: A Self-Consistent Model of Optical, X-ray and SZ Scaling Relations for Clusters of Galaxies, arXiv:1204.6305 [SPIRE].

[32] R. Angulo, V. Springel, S. White, A. Jenkins, C. Baugh et al., Scaling relations for galaxy clusters in the Millennium-XXL simulation, Mon. Not. Roy. Astron. Soc. 426 (2012) 2046 [arXiv:1203.3216] [SPIRE].

[33] R. Piffaretti, M. Arnaud, G. Pratt, E. Pointecouteau and J.-B. Melin, The MCXC: a Meta-Catalogue of X-ray detected Clusters of galaxies, Astron. Astrophys. 534 (2011) A109 [arXiv:1007.1916] [SPIRE].

[34] A.R. Duffy, J. Schaye, S.T. Kay and C. Dalla Vecchia, Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology, Mon. Not. Roy. Astron. Soc. 390 (2008) L64 [arXiv:0804.2486] [SPIRE].

[35] R.N. McInnes, F. Menanteau, A.F. Heavens, J.P. Hughes, R. Jimenez et al., First lensing measurements of SZ-discovered clusters, Mon. Not. Roy. Astron. Soc. 399 (2009) 84 [arXiv:0903.4410] [SPIRE].

[36] F. High, B. Stalder, J. Song, P. Ade, K. Aird et al., Optical Redshift and Richness Estimates for Galaxy Clusters Selected with the Sunyaev-Zel’dovich Effect from 2008 South Pole Telescope Observations, Astrophys. J. 723 (2010) 1736 [arXiv:1003.0005] [SPIRE].

[37] R. Suhada, J. Song, H. Boehringer, B. Benson, J. Mohr et al., XMM-Newton detection of two clusters of galaxies with strong SPT Sunyaev-Zel’dovich effect signatures, Astron. Astrophys. 514 (2010) L3 [arXiv:1003.3020] [SPIRE].

[38] K. Andersson, B. Benson, P. Ade, K. Aird, B. Armstrong et al., X-ray Properties of the First SZE-selected Galaxy Cluster Sample from the South Pole Telescope, Astrophys. J. 738 (2011) 48 [arXiv:1006.3068] [SPIRE].

[39] M. Brodwin, J. Ruel, P. Ade, K. Aird, K. Andersson et al., SPT-CL J0546-5345: A Massive z > 1 Galaxy Cluster Selected Via the Sunyaev-Zel’dovich Effect with the South Pole Telescope, Astrophys. J. 721 (2010) 90 [arXiv:1006.5639] [SPIRE].

[40] T. Marriage, V. Acquaviva, P. Ade, P. Aguierre, M. Amiri et al., The Atacama Cosmology Telescope: Sunyaev-Zel’dovich Selected Galaxy Clusters at 148 GHz in the 2008 Survey, Astrophys. J. 737 (2011) 61 [arXiv:1010.1065] [SPIRE].

[41] R. Foley, K. Andersson, G. Bazin, T. de Haan, J. Ruel et al., Discovery and Cosmological Implications of SPT-CL J2106-5844, the Most Massive Known Cluster at z > 1, Astrophys. J. 731 (2011) 86 [arXiv:1101.1286] [SPIRE].

[42] PLANCK collaboration, N. Aghanim et al., Planck Intermediate Results. I. Further validation of new Planck clusters with XMM-Newton, Astron. Astrophys. 543 (2012) A102 [arXiv:1112.5595] [SPIRE].

[43] C. Reichardt, B. Stalder, L. Bleem, T. Montroy, K. Aird et al., Galaxy clusters discovered via the Sunyaev-Zel’dovich effect in the first 720 square degrees of the South Pole Telescope survey, Astrophys. J. 763 (2013) 127 [arXiv:1203.5775] [SPIRE].

[44] F. High, H. Hoekstra, N. Leethochawalit, T. de Haan, L. Abramson et al., Weak-Lensing Mass Measurements of Five Galaxy Clusters in the South Pole Telescope Survey Using Magellan/Megacam, Astrophys. J. 758 (2012) 68 [arXiv:1205.3103] [SPIRE].

[45] AMI consortium, M.P. Schammel et al., AMI SZ observations and Bayesian analysis of a sample of six redshift-one clusters of galaxies, arXiv:1210.7771 [SPIRE].

[46] M. Hasselfield, M. Hilton, T.A. Marriage, G.E. Addison, L.F. Barrientos et al., The Atacama Cosmology Telescope: Sunyaev-Zel’dovich Selected Galaxy Clusters at 148 GHz from Three Seasons of Data, arXiv:1301.0816 [SPIRE].
[47] P. Rosati, P. Tozzi, R. Gobat, J. Santos, M. Nonino et al., *Multi-wavelength study of XMMU J2235.3-2557: the most massive galaxy cluster at z > 1*, Astron. Astrophys. 508 (2009) 583 [arXiv:0910.1716] [SPIRE].

[48] R. Gobat, E. Daddi, M. Onodera, A. Finoguenov, A. Renzini et al., *A mature cluster with X-ray emission at z=2.07*, Astron. Astrophys. 526 (2010) A133 [arXiv:1011.1837] [SPIRE].

[49] R. Fassbender, H. Boehringer, A. Nastasi, R. Suhada, M. Muehlegger et al., *The X-ray luminous galaxy cluster population at 0.9<z<1.6 as revealed by the XMM-Newton Distant Cluster Project*, New J. Phys. 13 (2011) 125014 [arXiv:1111.0009] [SPIRE].

[50] R. Suhada, J. Song, H. Bohringer, J. Mohr, G. Chon et al., *The XMM-BCS galaxy cluster survey: I. The X-ray selected cluster catalog from the initial 6 deg²*, Astron. Astrophys. 537 (2012) A39 [arXiv:1111.0141] [SPIRE].

[51] N. Okabe, M. Takada, K. Umetsu, T. Futamase and G.P. Smith, *LoCuSS: Subaru Weak Lensing Study of 30 Galaxy Clusters*, Publ. Astron. Soc. Jap. 62 (2010) 811 [arXiv:0903.1103] [SPIRE].

[52] R. Demarco, G. Wilson, A. Muzzin, M. Lacy, J. Surace et al., *Spectroscopic Confirmation of Three Red-Sequence Selected Galaxy Clusters at z=0.87, 1.16 and 1.21 from the SpARCS Survey*, Astrophys. J. 711 (2010) 1185 [arXiv:1002.0160] [SPIRE].

[53] F. Menanteau, J.P. Hughes, L.F. Barrientos, A.J. Deshpande, M. Hilton et al., *Southern Cosmology Survey II: Massive Optically-Selected Clusters from 70 square degrees of the SZE Common Survey Area*, Astrophys. J. Suppl. 191 (2010) 340 [arXiv:1002.2226] [SPIRE].

[54] B. Vulcani, A. Aragon-Salamanca, B.M. Poggianti, B. Milvang-Jensen, A. von der Linden et al., *Cl 1103.7-1245 at z=0.96: the highest redshift galaxy cluster in the EDisCS survey*, Astron. Astrophys. 544 (2012) A104 [arXiv:1207.1530] [SPIRE].

[55] O. Davis, J. Devriendt, S. Colombi, J. Silk and C. Pichon, *Most massive halos with Gumbel Statistics*, Mon. Not. Roy. Astron. Soc. 413 (2011) 2087 [arXiv:1101.2896] [SPIRE].

[56] E.J. Gumbel, *Statistics of Extremes*, Columbia University Press, New York (1958).