This currently addressed through section three. I've glanced at the rest and things look pretty good. I have extensive comments and questions for the ridgeline likelihood section.

Overall notes:

There sometimes seems to be mismatching between the referenced object and the actual number of the object. For example, just after equation 15, you refer to figure 5, but it is actually figure 6 you are referring to. Are you not using labels?

1 Introduction

p.1

Some of these statements are not yet proven, such as inflation.

Cosmology isn't an entity, so it cannot contribute :) I would say something about how the evolution of structure is predicted by the current models, which of course are parametrized by some numbers (which is probably what you meant by cosmology).

I would generally rework this first paragraph. This is the toughest paragraph to write and is also very important.

p.2

Individual galaxies are also observable counterparts of dark matter halos.

p.3

"However they suffer...." I think you mean the method suffers. "....which can mimic in projection..." Could be reworded. The last sentence needs to be reworded and expanded, since it isn't clear from context how you would be estimating the mass.

p.4

"X–ray methods are..." Maybe you should be more clear about what is not as sensitive. I assume you mean "the identification of potential wells through their X–ray emmision..." or something. Note, there are are point sources, so

“..the hot glow...” can come from other things. You need an extended source in the right energy range, etc. Its enough to say the catalogs are flux limited without saying they are flux limited at low redshift. Generally people use capital X and two dashes in X–ray.

p.5

“...across wide redshifts..” no plurality on redshift ranges. Your comment about flux preservation is a little confusing. It implies that somehow preservation of flux is a problem for the other methods, when really the point is that the energy boost of the CMB photons just depends on the physics of the gas not the redshift. Note, there is still that angular size -redshift relation to contend with, so small clusters are still limited. It allows a construction of volume, pressure, and size limited samples, the last two being correlated.

p.6

Weak lensing doesn’t really give you a mass limited sample, since it is highly dependend on redshift. Its a shear-size limited sample. The light isn’t distorted; you might say the shapes are distorted.

p.7

Here is a good place to say what you will do in the article in addition to whose work you will build upon.

2 Algorithm

2.1 Outline

p.1

No need to say it is an alternative, its just a method. You might call it a new method, since this is the first paper on the algorithm! You say it labels peaks in the galaxy density field, but then talk about color too.

You should say 1/r in projection. CMD, is this the first time it’s used? Maybe you should also first explain what the E/S0 ridgeline is.

Not all the galaxies on a given E/S0 ridgeline are very luminous, since it has been seen to extend to a fraction of L*.

p.2

Might want to talk about the evidence for this, and the fact that it isn't always the case that the brightest galaxy lives at the center.

p.3

"Individual objects..." This is a very difficult sentence. Needs to be made clear.

p.5

You wrote S0 (S-zero) when abbreviating Spherical Overdensity.

2.2 Likelihood Framework

p.1

The sentence "..center of cluster..." should be "..center of a cluster...".

This first sentence is a little misleading. You are really using the galaxies as discrete tracers of the "density field" in ra/dec/z, then looking at the maximum likelihood of the objects being a cluster center in three dimensions. Varying the redshift tells you nothing about its proximity to the center in ra/dec.

2.3 Ridgeline Likelihood

p.2 right after "Spatial Filter"

Do you mean truncated at x=0.1 or given a core? The galaxies are still counted right? I think you should show the plot as a log-log.

This is a long paragraph.

p.6

LRG hasn't been defined I don't think.

"...LRGs in this range must be..." I think you mean LRG's with $z < 0.15$ must be selected in alternative ways.

Define eclass and fracDev.

Another long paragraph.

p.7 right after "Luminosity Filter"

First sentence needs reworking. I find this paragrph confusingly worded. You want a cutoff in absolute magnitude that results in a volume limited catalog to moderate redshift with high S/N detections of the members at that redshift limit.

p.12

I don't understand equation 8. Let me try to reverse engineer it.

I think you are comparing a model plus background $M+b$ to the measured counts D. A simple model for the measured counts, ignoring filters and redshift,

$$N(< R) = N_{bg} + P_{bg} + \int_0^R d^2R \times NFW(R) + P_{NFW} \tag{1}$$

where $N(< R)$ is the total number in the aperture of radius R, N_{bg} is the mean number expected from the field, P_{bg} is the a poisson random number for the background, the integral is the integral over an NFW profile assuming the galaxies follow an NFW, and P_{NFW} is the poisson noise on this term. The noise is not background dominated beyond some intermediate redshift, but I'll need to run some numbers to know: later.

We normalize the NFW w.r.t to the maximum radius which will correspond to our aperture:

$$\begin{aligned} NFW(R) &= N_{gal}F(R) \\ \int_0^{Rmax} d^2RF(R) &= 1 \qquad (2) \\ & \qquad\qquad (3) \end{aligned}$$

where $F(R)$ is just the normalized, projected NFW profile. This give's a definition to N_{gal}, which can therefore be fit for.

I presume your model is essentially this, but with the filters added. The new normalization condition involves integrals over color as well as space.

The filters must then be included in the model, but the fit parameters are still just N_{gal} and z. The model including the color filters will suppress the background by integrating the field color distribution with the filters.

$$N_{bg}^{eff} = N_{bg} \int_0^{Rmax} d^2 R F(R) \int_{-\infty}^{\infty} d(g{-}r) n_{bg}(g{-}r) G_{gr} \int_{-\infty}^{\infty} d(r{-}i) n_{bg}(r{-}i) G_{ri} \tag{4}$$

Which can be measured, along with the noise term, from data.

The NFW model term then becomes an integral over colors and an integral of the spatial distribution $n_{cl}(r)$ times the NFW.

$$NFW(< R) = \int_0^{Rmax} d^2 R\, n_{cl}(R) F(R) \int_{-\infty}^{\infty} d(g{-}r) G_{gr} n_{cl}(g{-}r) \int_{-\infty}^{\infty} d(r{-}i) G_{ri} n_{cl}(r{-}i) \tag{5}$$

In the case where we are sitting at the center of a cluster at the right redshift this becomes

$$\begin{aligned} NFW(< R) &= N_{gal} \int_0^{Rmax} d^2 R\ F^2(R) \int_{-\infty}^{\infty} d(g-r) G^2(g-r) \int_{-\infty}^{\infty} d(r-i) G^2(r-i) \\ &= \frac{N_{gal}}{2} \int_0^{Rmax} d^2 R\ F^2(R) \equiv N_{gal} \frac{F_2}{2} \end{aligned} \tag{6}$$

you will need to keep the factor of 1/2 in mind. I don't know the answer to the integral F_2 but it must also be accounted for in order to recover the correct N_{gal}

One can then test the χ^2 for a grid of N_{gal} and z to find the maximum likelihood. I'm not sure where equation 12 comes from, however.

Also, you are assuming background noise dominated, but this won't be true at all redshifts. What do we know about the validity of this?