Counts of galaxy clusters as cosmological probes: the impact of baryonic physics

Andrés Balaguera-Antolínez, Cristiano Porciani
Argelander Institute für Astronomie, Auf dem Hügel 71, D-53121 Bonn, Germany
E-mail: abalan@astro.uni-bonn.de, porciani@astro.uni-bonn.de

Abstract. The halo mass function from N-body simulations of collisionless matter is generally used to retrieve cosmological parameters from observed counts of galaxy clusters. This neglects the observational fact that the baryonic mass fraction in clusters is a random variable that, on average, increases with the total mass (within an overdensity of 500). Considering a mock catalog that includes tens of thousands of galaxy clusters, as expected from the forthcoming generation of surveys, we show that the effect of a varying baryonic mass fraction will be observable with high statistical significance. The net effect is a change in the overall normalization of the cluster mass function and a milder modification of its shape. Our results indicate the necessity of taking into account baryonic corrections to the mass function if one wants to obtain unbiased estimates of the cosmological parameters from data of this quality. We introduce the formalism necessary to accomplish this goal. Our discussion is based on the conditional probability of finding a given value of the baryonic mass fraction for clusters of fixed total mass. Finally, we show that combining information from the cluster counts with measurements of the baryonic mass fraction in a small subsample of clusters (including only a few tens of objects) will nearly optimally constrain the cosmological parameters.

Keywords: cosmology: cosmological parameters, large-scale structure of universe, theory, observations, galaxies: clusters: general, methods: statistical
1 Introduction

Galaxy clusters are powerful cosmological probes [1–4] which are expected to be associated with massive haloes of dark matter. To date, two different procedures have been employed to extract cosmological information from large samples of galaxy clusters. The simplest one is to measure their observed number density as a function of mass and/or redshift and compare it with theoretical models for the halo abundance which are calibrated against N-body simulations of collisionless matter [1]. The second option is to use their baryonic content within the virial radius as a sort of standard ruler [5, 6]. This assumes that the baryon fraction in clusters coincides with the universal value [7] and does not scatter from object to object. The recent analysis by [8] has shown that this is a good assumption for relaxed, cool-core clusters. In this paper, we revisit the first approach and show that some modifications to the standard procedure will be needed when the forthcoming generation of cluster surveys will become available. Therefore, the second method will not be discussed any further.

The key theoretical quantity to analyze cluster-count experiments is the halo mass function $n_h(M, z)$ which is defined so that $n_h(M, z)\, dM$, gives the number of dark-matter haloes (at redshift $z$) with mass between $M$ and $M + dM$ per unit comoving volume. The calibration of the halo mass function using N-body simulations has been pushed to a precision of $\sim 5\%$ [9–12] and can be even further improved [13]. However, this precision is somewhat illusory for direct applications to galaxy clusters. In fact, numerical simulations that, more realistically, also consider a baryonic component on top of the dark matter [14–18] have shown that gas physics can alter cluster masses in a systematic way (i.e. in one specific direction and not randomly) by up to $\sim 10\%$ [16] with respect to pure N-body simulations. Consequently, given the steep slope of the halo mass function at cluster scales, the cluster abundance at fixed mass changes by $\sim 10 − 20\%$ [14]. This is larger than the forecasted statistical uncertainties ($\sim 1 − 10\%$) for the forthcoming generation of cluster surveys.

Regrettably, due to the large uncertainties in modeling baryon physics (i.e. radiative cooling, non-thermal pressure support, star formation, and different feedback mechanisms in the presence of stars and active galactic nuclei), current gas simulations can only provide order-of-magnitude approximations of these effects (see Section 2 for an extended discussion) and cannot provide robust estimates for the cluster mass function.

State-of-the-art techniques to constrain cosmological parameters from cluster counts take into consideration the effects of baryons by artificially enlarging the uncertainties in the parameters of the halo mass function extracted from N-body simulations and marginalizing over them [3]. This is a statistical trick to account for possible differences between the shapes of the halo and cluster mass functions. Such an approach is adequate for the current samples of clusters which probe relatively small volumes and have large error bars [19]. Forthcoming cluster surveys will probe much larger volumes with higher sensitivities. In this paper, we show that, in order to use the new data to derive cosmological constraints that are not only precise (i.e. small statistical errors) but also accurate (i.e. unbiased estimates), it will be imperative to model the baryonic effects on the cluster mass function.

As routinely done in X-ray and Sunyaev-Zel’dovich studies, we define cluster and halo masses as spherical overdensities bounded by the radius $R_{500}$ within which the mean density is $\Delta = 500$ times the critical density of the Universe. Note the $R_{500}$ is typically 1.5-2 times smaller than the virial radius (defined as in [20]) and the baryon fraction within this smaller scale, $f_b$, is observed to be below the cosmic value (e.g. [8, 21]) and vary systematically with the cluster mass [22–25]. Following a statistical approach, we describe the effects of gas physics in two steps. We first use the outcome of recent numerical studies, to relate the dark-matter content of galaxy clusters with halo masses in N-body simulations. Then we reason in terms of the mean baryonic mass fraction as a function of cluster mass, $f_b(M)$, and of the corresponding dispersion around the mean $\sigma_b$. We focus our attention on two cosmological parameters: the matter density parameter, $\Omega_m$, and the linear rms matter fluctuation within a spherical top-hat window of radius $8 \, h^{-1} \, \text{Mpc}$, $\sigma_8$. Considering observationally motivated guesses for the shape and amplitude of the function $f_b(M)$ and of the scatter $\sigma_b$, we first quantify the bias in the estimates for $\Omega_m$ and $\sigma_8$ caused by using the halo mass function as a proxy for the cluster mass function. Subsequently, we explore the option of using extra free parameters in the models to simultaneously determine from the cluster counts both the cosmological parameters and the relation...
We find that this procedure increases the uncertainties on $\Omega_m$ and $\sigma_8$. Finally, we show that considering additional information on the baryonic mass fraction from follow-up studies of a small subsample of clusters is pivotal to recover unbiased estimates for $\Omega_m$ and $\sigma_8$ with small uncertainties.

We conclude that taking into account the effect of baryons on the cluster mass function is key to fully exploit the potential of forthcoming large-volume cluster surveys as cosmological probes.

We adopt a fiducial cosmological model based on a flat $\Lambda$CDM Universe with a matter density parameter of $\Omega_m = 0.258$, a baryon density parameter of $\Omega_b = 0.044$, a dimensionless Hubble parameter $h = 0.735$ (in units of $100\,\text{km}\,\text{s}^{-1}\text{Mpc}^{-1}$), a linear rms mass fluctuation within $8\,h^{-1}\text{Mpc}$ of $\sigma_8 = 0.773$ and a scalar spectral index $n_s = 0.954$. We use the transfer function of [26] to compute the linear matter power spectrum. The halo mass function is calculated using the fitting formulae of [11]. Logarithms are always taken after measuring the mass in units of $h^{-1}M_\odot$.

2 Matching cluster masses to N-body simulations

Many authors have compared the mass density profiles of clusters generated from hydrodynamic and collisionless simulations with identical initial conditions. The presence of the baryons generates not only a change in the total mass (defined within a fixed overdensity) of a cluster compared to the corresponding halo in a N-body simulation, but also a re-distribution of the dark-matter component [27–29]. When the gas is not allowed to radiate, infalling baryons are heated up in shocks when they cross the virial radius and subsequently exchange energy with the dark matter while the cluster undergoes dynamical relaxation. They end up having a more extended distribution than the dark matter [28, 30–34]. As a matter of fact, the baryon fraction within the virial radius decreases with the cluster mass and lies below the cosmic value, $f_c \equiv \Omega_b/\Omega_m$, even for the most massive clusters (see Fig. 5 in [28]). The dark-matter profile in spherical shells becomes slightly more concentrated than in N-body simulations (see Fig. 6 in [28]) and the total dark-matter mass within $\Delta = 500$ increases by $\sim 1\%$ for an object of $10^{14}h^{-1}M_\odot$. When, instead, gas cooling and star formation are included in simulations, baryons dissipate energy radiatively and condense in the central regions pulling the dark matter along. Orbits of substructures are also altered [16]. If the energy feedback due to star formation is weak, the net effect is that the concentration of the dark matter increases [16, 28]. In this case, clusters exhibit dark-matter (and total) masses within $R_{500}$ which are a few per cent higher than in the corresponding N-body simulations. Note, however, that the baryon fractions of the simulated clusters within the virial radius systematically lie 1-5% above the universal value. Stellar profiles also deviate substantially from the observed ones [16, 29]. Observations [22–25] have shown that the baryon fraction within $R_{500}$ lies below $f_c$ and varies with the cluster mass. Recent studies favour the interpretation that the “missing” baryons are located in the cluster outskirts so that the baryon fraction reaches the cosmic value near the virial radius [8, 35]. This behaviour is qualitatively reproduced only by the numerical simulations that include either strong stellar feedback or active galactic nuclei (see Fig. 2 in [29]). In these simulations, the dark-matter concentration (and therefore the dark-matter mass) of the clusters coincides with what is measured in the corresponding N-body simulations to very good accuracy (see Fig. 8 in [29]).

Let us define a cluster as the spherical region centered on a density peak and enclosing the overdensity $\Delta = 500$ times the critical density of the universe. Each cluster will be characterized by a well defined total mass $M_{\text{tot}}$ (such that $M_{\text{tot}} = M_b + M_{\text{dm}}$, where $M_{\text{dm}}$ is the dark matter component of a cluster) and a baryon fraction $f_b \equiv M_b/M_{\text{tot}}$. The same definition can be applied to identify haloes in a collisionless simulation. Note, however, that particles in N-body simulations are assigned masses proportional to the total matter content of the Universe, i.e. $m_p \propto \Omega_m = \Omega_b + \Omega_{\text{dm}}$ (in terms of the baryonic and dark-matter components). Thus, by construction, the mass of simulated dark-matter haloes includes a baryonic fraction which coincides with $f_c$. The total mass thus reads, $M_{\text{nb}} = M_{\text{dm}}/(1-f_c)$, where $M_{\text{dm}}$ is the corresponding dark matter mass. As discussed above, hydro simulations with strong feedback are the only ones which are able to reproduce the observed baryon distribution. Based on the fact that in these simulations the dark-matter mass of the clusters is hardly modified with respect to N-body simulations, in what follows we assume that $M_{\text{dm}} = M_{\text{dm}}$. The total mass of a cluster can be then linked with the total mass of the same halo in a N-body simulation $M_{\text{nb}}$. 


via

\[ M_{\text{tot}} = \left( \frac{1 - f_c}{1 - f_b} \right) M_{\text{nb}}. \]  

(2.1)

Our assumption that \( M_{\text{dm}} = \tilde{M}_{\text{dm}} \) should be seen as a working hypothesis supported by the match between hydro simulations and observations. If future numerical simulations will evidence the need of a more complicated treatment, our analysis can be, of course, generalized by introducing a probability density function \( P(M_{\text{dm}}|M_{\text{dm}}) \) at the expenses of introducing a number of extra parameters.

3 The observed mass function

Consider now a population of galaxy clusters and a set of haloes from a N-body simulation both at the same redshift. Our goal is to write the mass function of the clusters in terms of that of the haloes (at a fixed \( \Delta = 500 \)). We need to keep into account that the baryon fraction of the real clusters is a stochastic quantity that varies from object to object. Based on observational studies, we assume that the conditional probability density \( P_b(f_b|M_{\text{tot}}) \), is well approximated by a Gaussian distribution with mean

\[ \langle f_b|M_{\text{tot}} \rangle = A f_c \left( \frac{M_{\text{tot}}}{2 \times 10^{14} h^{-1} M_\odot} \right)^B, \]

(3.1)

and scatter \( \sigma_b \). We adopt the results by [22], \( A = 0.74 \pm 0.01 \), \( B = 0.09 \pm 0.03 \) and \( \sigma_b \approx 2 \times 10^{-3} \), as fiducial values but it is worth stressing that other studies favour slightly different values for these parameters. We want to write the conditional probability density \( P(M_{\text{tot}}|M_{\text{nb}}) \) which gives the distribution of the total cluster mass for a given \( M_{\text{nb}} \). This quantity satisfies the integral equation

\[ P(M_{\text{tot}}|M_{\text{nb}}) = \frac{P(M_{\text{nb}}|M_{\text{tot}})}{n_b(M_{\text{nb}})} \int_0^\infty n_h(x) P(M_{\text{tot}}|x) \, dx, \]

(3.2)

where \( P(M_{\text{nb}}|M_{\text{tot}}) = (1 - f_c)P_b(f_h = f_*|M_{\text{tot}})/M_{\text{tot}} \) with \( f_* \equiv 1 - (1 - f_c)M_{\text{nb}}/M_{\text{tot}} \). Since \( \sigma_b \ll \langle f_b|M_{\text{tot}} \rangle \), then \( P(M_{\text{tot}}|M_{\text{nb}}) \) is a narrow function centered at \( \langle M_{\text{tot}}|M_{\text{nb}} \rangle \approx M_{\text{nb}}(1 - f_c)/(1 - \langle f_b|M_{\text{tot}} \rangle) \), so that we can finally write

\[ P(M_{\text{tot}}|M_{\text{nb}}) = \frac{1 - f_c}{M_{\text{tot}}} P_b(f_*|M_{\text{tot}}) \frac{n_h(M_{\text{nb}})}{n_b(M_{\text{nb}})} J(M_{\text{tot}}), \]

(3.3)

where \( \tilde{M}_{\text{nb}} \) is the solution to the equation \( M_{\text{tot}} = \langle M_{\text{tot}}|\tilde{M}_{\text{nb}} \rangle \) and

\[ J(M_{\text{tot}}) = \left| \frac{d(M_{\text{tot}}|M_{\text{nb}})}{dM_{\text{nb}}} \right|^{-1}_{M_{\text{nb}} = \tilde{M}_{\text{nb}}}. \]

(3.4)

The conditional probability density in Eq.(3.3) represents the necessary tool to associate the masses of dark matter haloes defined in N-body simulations with the total masses of galaxy clusters in the presence of a mass dependent fraction of baryons. In general, \( P(M_{\text{tot}}|M_{\text{nb}}) \) is very well approximated by a log-normal distribution with a mass-dependent log-scatter \( \sigma_{\ln M_{\text{tot}}|M_{\text{nb}}} \approx \ln [1 + \sigma_b/(1 - (f_b|M_{\text{tot}}))] \) (see panel (c) of Fig. 1 for an example).

Cluster masses are not observed directly and need to be inferred from observational proxies (e.g. optical richness, X-ray temperature or flux, Sunyaev-Zel’dovich signal, lensing shear). We assume that the observed mass, \( M_{\text{obs}} \), is an unbiased estimate of the total mass, with a log-normal probability density function of the residuals \( P(M_{\text{obs}}|M_{\text{tot}}) \) [36, 37], for which we use a constant log-scatter \( \sigma_{\ln M_{\text{obs}}|M_{\text{tot}}} = 0.1 \). In general the scaling relation \( P(M_{\text{obs}}|M_{\text{tot}}) \) is obtained observationally, and for this reason there is no need to explicitly emphasize its dependence on the baryonic fraction.
Figure 1. (a) Ratio between the cluster mass function $n$ and a model assuming that the baryonic mass fraction is always equal to the cosmic value (i.e. $M_{\text{tot}} = M_{\text{nb}}$), $n_0$. Line styles correspond to the observed baryon fraction reported by different authors. (b) Signal-to-noise ratio for the difference $n - n_0$ as a function of the observed cluster mass in 20 equispaced log-bins. A survey volume of $0.8 \ (h^{-1}\text{Gpc})^3$ centered at $z = 0.1$ is assumed. (c) The conditional probability density $P(\log_{10} M_{\text{tot}} | M_{\text{nb}})$ at a mass scale $M_{\text{nb}} = 1.3 \times 10^{14} h^{-1} M_\odot$ as given in Eq.(3.3).

(e.g., the X-ray luminosity depends on the total amount of gas) in our formalism. The mass function of galaxy clusters in terms of their observed mass is then

$$n(M_{\text{obs}}, z) = \int_0^\infty n_b(M_{\text{nb}}, z) P(M_{\text{obs}} | M_{\text{nb}}) \, dM_{\text{nb}},$$  

(3.5)

with

$$P(M_{\text{obs}} | M_{\text{nb}}) = \int_0^\infty P(M_{\text{obs}} | M_{\text{tot}}) P(M_{\text{tot}} | M_{\text{nb}}) \, dM_{\text{tot}}.$$  

(3.6)

Note that the conditional probability $P(M_{\text{obs}} | M_{\text{nb}})$ can be written as a log-normal distribution with scatter $[\sigma_{\ln M_{\text{obs}} | M_{\text{tot}}}^2 + \sigma_{\ln M_{\text{tot}} | M_{\text{nb}}}^2]^{1/2}$. The main effect of the variable baryon fraction is therefore to introduce a systematic mass-dependent offset between $M_{\text{obs}}$ and $M_{\text{nb}}$ while the correction to the intrinsic scatter plays a sub-dominant role, at least for mass proxies with broad distributions ($\sigma_{\ln M_{\text{tot}} | M_{\text{nb}}} \ll \sigma_{\ln M_{\text{obs}} | M_{\text{tot}}}$).
Table 1. Summary of the Bayesian-inference cases discussed in this paper. Note that \( n_0 \) is obtained assuming that all clusters have \( M_{\text{tot}} \equiv M_{\text{ab}} \) (i.e. \( f_b = f_c \)) in Eq. (3.6) while \( n \) takes into account that the baryonic mass fraction varies from object to object.

| Case | Model | Notes | Data | Description |
|------|-------|-------|------|-------------|
| I    | \( n_0 \) |       | \( n_{\text{obs}} \) | The halo mass function from N-body simulations is used to fit the observed mass function |
| II   | \( n_0 \) \( \uparrow \) cov | \( n_{\text{obs}} \) | \( n_{\text{obs}} \) | As in I but after artificially inflating the covariance matrix of the \( n_0 \) parameters |
| III  | \( n \) |       | \( n_{\text{obs}} \) | Accounting for a varying baryon fraction and simultaneously fitting \( P(f_b|M_{\text{obs}}) \) |
| IV   | \( n \) |       | \( n_{\text{obs}}, f_b \) | As in III combining the mass function data with 30 measurements of \( f_b \) |
| V    | \( n \) |       | \( n_{\text{obs}} \) | As in III but assuming perfect a priori knowledge of \( P(f_b|M_{\text{tot}}) \) |
| I_G  | \( n_0 \) |       | \( n_{\text{obs}} \) | As in I marginalizing over the mass-measurement error with a broad flat prior (see text) |
| I_F  | \( n_0 \) |       | \( n_{\text{obs}} \) | As in I marginalizing over the mass-measurement error with a broad flat prior (see text) |
| IV_G | \( n \) |       | \( n_{\text{obs}}, f_b \) | As in IV marginalizing over the mass-measurement error with a Gaussian prior (see text) |
| IV_F | \( n \) |       | \( n_{\text{obs}}, f_b \) | As in IV marginalizing over the mass-measurement error with a broad flat prior (see text) |

In panel (a) of Fig. 1 we compare the cluster mass function, \( n(M_{\text{obs}}) \), with its counterpart obtained by neglecting the effects of the varying baryon fraction (i.e. assuming that \( M_{\text{tot}} = M_{\text{ab}} \)) that we dub \( n_0(M_{\text{obs}}) \) - note that \( n_0 \) differs from \( n \) due to the mass-measurement errors. We use the observational scaling relations \( P_b(f_b|M_{\text{tot}}) \) by [22] and [23]. The latter corresponds to Eq. (3.1) with \( A \approx 0.8 \) and \( B = 0.136 \pm 0.028 \). The main effect is a reduction of the cluster counts by \( 5 \% - 15 \% \), depending on the cluster mass and the details of the model.

Will the discrepancy between the actual cluster mass function, \( n(M_{\text{obs}}) \), and the predictions of N-body simulations (convolved with the mass-measurement error), \( n_0(M_{\text{obs}}) \), be noticeable with future observational campaigns? As a prototype for the forthcoming cluster surveys, we consider a catalog spanning a comoving volume of \( 0.8 \; (h^{-1}\text{Gpc})^3 \) (corresponding to a survey covering the full sky down to \( z < 0.2 \) or half of the sky to \( z < 0.25 \)) and containing \( 2.79 \times 10^4 \) entries in the mass range \( 10^{13.5} < M_{\text{obs}}/(h^{-1}M_\odot) < 10^{15} \) (for our fiducial model) with mean redshift \( z = 0.1 \). We assume full completeness in the mass range we are probing and compute the cluster mass function in 20 mass bins of width \( \Delta \log_{10} M_{\text{obs}} = 0.075 \). Assuming Poisson errorbars of size \( \sigma \), in panel (b) of Fig. 1 we show the signal-to-noise ratio for the difference \( n(M_{\text{obs}}) - n_0(M_{\text{obs}}) \) as a function of the cluster mass. Highly statistically significant deviations are detectable for \( M_{\text{obs}} < 10^{14.5} h^{-1} M_\odot \).

4 Cosmological parameters

We want to quantify the bias and the uncertainty in the measurement of the cosmological parameters \( \Omega_m \) and \( \sigma_8 \) obtained by fitting the observed cluster mass function, \( n_{\text{obs}} \), with different models. In order to do this, we use the mock cluster catalog described in the previous section and we sample the posterior distribution of the model parameters with a Markov Chain Monte Carlo algorithm. We use different models and combinations of data which are briefly summarized in Table 1 and extensively described below. The (marginalized) posterior mean and rms values for \( \Omega_m \) and \( \sigma_8 \) are given in Table 2 together with the corresponding “figure of merit” (FoM, defined as the inverse of the area of the joint 68.3\% credibility region in the \( \{ \Omega_m, \sigma_8 \} \) plane). The joint 95.4\% credibility regions are instead shown in panel (a) of Fig. 2.

First, we consider the mass function extracted from N-body simulations with no corrections for a varying \( f_b \) (case I). Specifically, we use the function \( n_0(M_{\text{obs}}) \) to fit the observed mass distribution. The resulting estimate for \( \Omega_m \) is significantly biased low. To first order, this is because the normalization of the function \( n_0 \) scales proportionally to \( \Omega_m \) and, as shown in panel (a) of Fig. 1, the effect of the varying baryon fraction is to reduce the overall normalization of the observed cluster counts. On the other hand, \( \sigma_8 \) is slightly biased high (for our fiducial model) as expected from the location of the exponential cutoff in the mass function. However, the bias is not very significant given the corresponding statistical uncertainty. The size and sign of the bias on \( \sigma_8 \) markedly depends on the values adopted in Eq. (3.1) while this is not true for \( \Omega_m \), which is always biased low for \( A < 1 \).

To gain freedom in the shape of the theoretical mass function and reduce systematic effects when fitting current data, it is common practice to let the parameters that define \( n_b \) vary within some predefined range [3]. We have investigated what happens applying this technique to our mock sample.
Table 2. Mean and rms value of the marginalized posterior distribution for $\Omega_m$ and $\sigma_8$. The third column shows the figure-of-merit, defined as the inverse of the area of the joint 68.3\% credibility region.

| Case | $\Omega_m$       | $\sigma_8$     | FoM         |
|------|------------------|----------------|-------------|
| Fiducial | 0.258           | 0.773          | 2.4 \times 10^4 |
| I    | 0.245 $\pm$ 0.002 | 0.780 $\pm$ 0.007 | 6.7 $\times$ 10^3 |
| II   | 0.244 $\pm$ 0.003 | 0.780 $\pm$ 0.007 | 2.8 $\times$ 10^3 |
| III  | 0.259 $\pm$ 0.005 | 0.782 $\pm$ 0.012 | 2.1 $\times$ 10^4 |
| IV   | 0.257 $\pm$ 0.003 | 0.773 $\pm$ 0.007 | 2.5 $\times$ 10^4 |
| V    | 0.258 $\pm$ 0.003 | 0.772 $\pm$ 0.007 | 2.5 $\times$ 10^4 |

In this case we have allowed the parameters of the mass function to vary within a four-dimensional Gaussian prior. We built the covariance matrix of the prior by multiplying the original covariance matrix of the parameters (kindly made available by Jeremy Tinker) with a positive constant so that the halo mass function at $M = 10^{15.3} h^{-1} M_\odot$ is $\sim$ 10\% uncertain (case II). With respect to case I, this method does not improve the bias of $\Omega_m$ and $\sigma_8$ while the corresponding FoM decreases by a factor of $\sim 3.6$.

In order to eliminate the bias, we release the assumption that $M_{\text{tot}} = M_{\text{bh}}$ in Eq. (3.6) and replace $n_0$ with $n$ to fit $n_{\text{obs}}$ (case III). This way, we simultaneously constrain the cosmological parameters and the scaling relation $P(f_b|M_{\text{tot}})$. This procedure is analogous to the “self calibration” method proposed to extract cosmological information from cluster surveys [38]. We have verified that this approach provides unbiased estimates of $\Omega_m$ and $\sigma_8$. However, the statistical uncertainty on the values of the cosmological parameters constrained by the cluster counts are significantly larger with respect to case I and II.

The situation markedly improves by considering additional information on $f_b$ extracted from multi-wavelength studies of a small subset of galaxy clusters. To show this, we randomly select 30 clusters out of the full sample and imagine that the baryonic fraction of their mass content has been measured with 10\% precision in an unbiased way (case IV). We then build new Markov chains assuming that the information from the measurement of the baryon fraction is independent of the cluster counts. The resulting estimates for $\Omega_m$ and $\sigma_8$ are unbiased and errorbars small as expected from experiments for “precision cosmology”. In this case, the parameters of the scaling relation $P(f_b|M_{\text{tot}})$ are also recovered to good accuracy, namely (to 68.3\% credibility), $A = 0.74 \pm 0.01$, $B = 0.09 \pm 0.01$ and $\sigma_b = 4.5^{+3.1}_{-1.6} \times 10^{-3}$.

Finally, we consider the ideal case in which the scaling relation $P(f_b|M_{\text{tot}})$ is perfectly known from independent data (case V, not shown in Fig. 2). This allows us to conclude that case IV gives cosmological constraints which are nearly optimal.

To simplify the discussion, so far, we have assumed that the scatter of the observational mass estimates for the galaxy clusters, $\sigma_{\text{in},M_{\text{obs}}|M_{\text{bh}}}$, is perfectly known. This, however, does not accurately reflect reality where one has only limited knowledge on the actual size of the measurement error which should then be treated as a nuisance parameter in the model-fitting procedure [37]. This will generally broaden the credibility region of the cosmological parameters and, in principle, might reduce the statistical significance of the biased estimates obtained with model $n_0$. In order to evaluate the impact of this subtlety on our results, we have repeated the measurement of the cosmological parameters (for case I and IV) after marginalizing over $\sigma_{\text{in},M_{\text{obs}}|M_{\text{bh}}}$.

First, as a realistic option, we have adopted a Gaussian prior on $\sigma_{\text{in},M_{\text{obs}}|M_{\text{bh}}}$, with mean 0.1 and rms error 0.02 (cases I_G and IV_G) - note that a 20\% standard error of the standard deviation corresponds to a sample of 14 objects. As a (very) pessimistic option, instead, we have considered a flat prior between $0 < \sigma_{\text{in},M_{\text{obs}}|M_{\text{bh}}} < 1$ (cases I_F and IV_F). The corresponding joint posterior distributions for $\Omega_m$ and $\sigma_8$ are shown in panel (b) of Fig. 2. In all cases, the bias in the estimates based on the model $n_0$ is still evident.
5 Discussion and conclusions

The fractional baryon content of galaxy clusters within an overdensity of $\Delta = 500$ is observed to be a random variable which, on average, decreases with the cluster mass [22–25]. Therefore the cluster mass function must differ from the predictions of N-body simulations where the baryon fraction is implicitly held constant to the cosmic value.

The forthcoming generation of cluster surveys will provide number counts with an accuracy ranging between 1 and 10% depending on the cluster mass. Such an error is substantially smaller than the difference between models of the mass function with and without baryonic physics. Considering a prototypic catalog of galaxy clusters containing $2.79 \times 10^4$ entries, we have shown that constraints on the cosmological parameters $\Omega_m$ and $\sigma_8$ derived from the cluster mass function would be severely biased if this signal is modelled with fitting formulae based on N-body simulations of collisionless matter. In particular, $\Omega_m$ would be always biased low while the bias on $\sigma_8$ depends on the details of the scaling relations between $f_b$ and the cluster mass.

The widespread technique of artificially inflating the covariance matrix of the parameters that describe the halo mass function to gain freedom and minimize systematic effect will not be of much help. Our study shows that it would enlarge the statistical uncertainties on the cosmological parameters without eliminating (or reducing) the bias.

In order to obtain accurate estimates for the cosmological parameters, complementary information on the fraction of baryons as a function of the total mass is required. The optimal method to eliminate this systematic effect requires two ingredients: I) an accurate model for the conditional probability density of finding a particular value for $f_b$ given the cluster total mass, $P(f_b|M_{\text{tot}})$; II) A small, random subsample of clusters with follow-up data for which simultaneous measurements of $f_b$ and $M_{\text{tot}}$ can be made. We have shown that, if the scatter around the mean $f_b - M_{\text{tot}}$ relation, $\sigma_b$, is independent of mass, nearly 30 objects would be enough for precision cosmology. The required size of the subsample should grow bigger if $\sigma_b$ has a strong mass dependence.

It is interesting to compare the precision we can achieve with the different methods in our fiducial case (see Table 2). Using the N-body mass function (case I) returns estimates for $\sigma_8$ and $\Omega_m$ with statistical errors of 1% and 2%, respectively (but with a systematic shift which is approximately
1 and 6.5 times larger). On the other hand, accounting for the variable baryon fraction (case III) eliminates the biases but nearly doubles the rms of the marginal probabilities due to the inclusion of three additional parameters. However, combining this method with sufficient follow-up information (case IV), we can optimally recover the same uncertainties as in case I.

Let us now critically discuss the formalism we have outlined in this Paper. By performing an object by object comparison between collisionless and hydrodynamic simulations with the same initial conditions, recent studies have shown that the dark-matter mass of a cluster (within an overdensity of $\Delta = 500$) does not change when baryonic physics is included, provided a strong form of feedback is considered. Our analysis is based on this result. This is a conservative assumption which, if violated, would introduce additional deviations in the cluster mass function and thus make estimates of the cosmological parameters based on the standard approach even more biased (since it is very unlikely that this effect would exactly cancel the mass dependence of $f_b$). However, in Section 2, we have indicated how our calculations could be easily generalized to that case.

To facilitate understanding, our simple analysis considers a complete sample in a narrow redshift bin and only two cosmological parameters but our results are of general value. It is straightforward to generalize them including more parameters and accounting for possible evolutionary effects in the baryon fraction along the past light cone and for the radial selection function of a realistic survey. This, however, is beyond the scope of this Paper.

We conclude that considering the effect of baryons on the cluster mass function is central to extract unbiased estimates of the cosmological parameters from forthcoming large-volume surveys such as DES [39], eROSITA [40, 41], ASKAP-EMU [42], CCAT [43], LSST [44] and Euclid [45].

Acknowledgments

We acknowledge support through the SFB-Transregio 33 “The Dark Universe” by the Deutsche Forschungsgemeinschaft (DFG).

References

[1] A. Vikhlinin, A. V. Kravtsov, R. A. Burenin, H. Ebeling, W. R. Forman, A. Hornstrup, C. Jones, S. S. Murray, D. Nagai, H. Quintana, and A. Voevodkin, Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints, ApJ 692 (Feb., 2009) 1060–1074, [arXiv:0812.2720].

[2] C. Cunha, D. Huterer, and J. A. Frieman, Constraining dark energy with clusters: Complementarity with other probes, Phys. Rev. D. 80 (Sept., 2009) 063532, [arXiv:0904.1589].

[3] A. Mantz, S. W. Allen, H. Ebeling, D. Rapetti, and A. Drlica-Wagner, The observed growth of massive galaxy clusters - II. X-ray scaling relations, MNRAS 406 (Aug., 2010) 1773–1795, [arXiv:0909.3099].

[4] D. Rapetti, C. Blake, S. W. Allen, A. Mantz, D. Parkinson, and F. Beutler, A combined measurement of cosmic growth and expansion from clusters of galaxies, the CMB and galaxy clustering, ArXiv e-prints (May, 2012) [arXiv:1205.4679].

[5] S. Ettori, P. Tozzi, and P. Rosati, Constraining the cosmological parameters with the gas mass fraction in local and $z>0.7$ galaxy clusters, A&A 398 (Feb., 2003) 879–890, [astro-ph/].

[6] S. W. Allen, D. A. Rapetti, R. W. Schmidt, H. Ebeling, R. G. Morris, and A. C. Fabian, Improved constraints on dark energy from Chandra X-ray observations of the largest relaxed galaxy clusters, MNRAS 383 (Jan., 2008) 879–896, [arXiv:0706.0033].

[7] S. D. M. White, J. F. Navarro, A. E. Evrard, and C. S. Frenk, The baryon content of galaxy clusters: a challenge to cosmological orthodoxy, Nature 366 (Dec., 1993) 429–433.

[8] D. Eckert, S. Ettori, S. Molendi, F. Vazza, and S. Paltani, The X-ray/SZ view of the virial region. II. Gas mass fraction, ArXiv e-prints (Jan., 2013) [arXiv:1301.0624].

[9] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida, The mass function of dark matter haloes, MNRAS 321 (Feb., 2001) 372–384, [astro-ph/].
[10] M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro, Precision Determination of the Mass Function of Dark Matter Halos, *ApJ* 646 (Aug., 2006) 881–885, [astro-ph/].

[11] J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz, Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality, *ApJ* 688 (Dec., 2008) 709–728, [arXiv:0803.2706].

[12] A. Pillepich, C. Porciani, and O. Hahn, Halo mass function and scale-dependent bias from N-body simulations with non-Gaussian initial conditions, *MNRAS* 402 (Feb., 2010) 191–206, [arXiv:0811.4176].

[13] D. S. Reed, R. E. Smith, D. Potter, A. Schneider, J. Stadel, and B. Moore, Toward an accurate mass function for precision cosmology, *ArXiv e-prints* (June, 2012) [arXiv:1206.5302].

[14] R. Stanek, D. Rudd, and A. E. Evrard, The effect of gas physics on the halo mass function, *MNRAS* 394 (Mar., 2009) L1–L15, [arXiv:0809.2805].

[15] I. G. McCarthy, J. Schaye, T. J. Ponman, R. G. Bower, C. M. Booth, C. Dalla Vecchia, R. A. Crain, V. Springel, T. Theuns, and R. P. C. Wiersma, The case for AGN feedback in galaxy groups, *MNRAS* 406 (Aug., 2010) 822–839, [arXiv:0911.2641].

[16] W. Cui, S. Borgani, K. Dolag, G. Murante, and L. Tornatore, The effects of baryons on the halo mass function, *MNRAS* 423 (July, 2012) 2279–2287, [arXiv:1111.3066].

[17] M. P. van Daalen, J. Schaye, C. M. Booth, and C. Dalla Vecchia, The effects of galaxy formation on the matter power spectrum: a challenge for precision cosmology, *MNRAS* 415 (Aug., 2011) 3649–3665, [arXiv:1104.1174].

[18] E. Rasia, M. Meneghetti, R. Martino, S. Borgani, A. Bonafede, K. Dolag, S. Ettori, D. Fabjan, C. Giocoli, P. Mazzotta, J. Merten, M. Radovich, and L. Tornatore, Lensing and x-ray mass estimates of clusters (simulations), *New Journal of Physics* 14 (May, 2012) 055018, [arXiv:1201.1569].

[19] A. Balaguera-Antolínez, A. G. Sánchez, H. Böhringer, and C. Collins, Constructing mock catalogues for the REFLEX II galaxy cluster sample, *MNRAS* 425 (Sept., 2012) 2244–2254, [arXiv:1207.2138].

[20] V. R. Eke, S. Cole, and C. S. Frenk, Cluster evolution as a diagnostic for Omega, *MNRAS* 282 (Sept., 1996) 263–280, [astro-ph/].

[21] S. Andreon, The stellar mass fraction and baryon content of galaxy clusters and groups, *MNRAS* 407 (Sept., 2010) 263–276, [arXiv:1004.2785].

[22] S. DiGionidi, D. Pierini, A. Finoguenov, G. W. Pratt, H. Boehringer, A. Leauthaud, L. Guzzo, H. Aussel, M. Bolzonella, P. Capak, M. Elvis, G. Hasinger, O. Ilbert, J. S. Kartaltepe, A. M. Koekemoer, S. J. Lilly, R. Massey, H. J. McCracken, J. Rhodes, M. Salvato, D. B. Sanders, N. Z. Scoville, S. Sasaki, V. Smolcic, Y. Taniguchi, D. Thompson, and COSMOS Collaboration, Stellar and Total Baryon Mass Fractions in Groups and Clusters Since Redshift 1, *ApJ* 703 (Sept., 2009) 982–993, [arXiv:0904.0448].

[23] T. F. Lagana, Y.-Y. Zhang, T. H. Reiprich, and P. Schneider, XMM-Newton/Sloan Digital Sky Survey: Star Formation Efficiency in Galaxy Clusters and Constraints on the Matter-density Parameter, *ApJ* 743 (Dec., 2011) 13, [arXiv:1108.3678].

[24] Y.-T. Lin, S. A. Stanford, P. R. M. Eisenhardt, A. Vikhlinin, B. J. Maughan, and A. Kravtsov, Baryon Content of Massive Galaxy Clusters at z = 0–0.6, *ApJ Lett.* 745 (Jan., 2012) L3, [arXiv:1112.1705].

[25] M. Sun, Hot gas in galaxy groups: recent observations, *New Journal of Physics* 14 (Apr., 2012) 045004, [arXiv:1203.4228].

[26] D. J. Eisenstein and W. Hu, Baryonic Features in the Matter Transfer Function, *ApJ* 496 (Mar., 1998) 605, [astro-ph/].

[27] O. Y. Gnedin, A. V. Kravtsov, A. A. Klypin, and D. Nagai, Response of Dark Matter Halos to Condensation of Baryons: Cosmological Simulations and Improved Adiabatic Contraction Model, *ApJ* 616 (Nov., 2004) 16–26, [astro-ph/].

[28] D. H. Rudd, A. R. Zentner, and A. V. Kravtsov, Effects of Baryons and Dissipation on the Matter Power Spectrum, *ApJ* 672 (Jan., 2008) 19–32, [astro-ph/].
[29] A. R. Duffy, J. Schaye, S. T. Kay, C. Dalla Vecchia, R. A. Battye, and C. M. Booth, *Impact of baryon physics on dark matter structures: a detailed simulation study of halo density profiles*, MNRAS 405 (July, 2010) 2161–2178, [arXiv:1001.3447](http://arxiv.org/abs/1001.3447).

[30] V. R. Eke, J. F. Navarro, and C. S. Frenk, *The Evolution of X-Ray Clusters in a Low-Density Universe*, ApJ 503 (Aug., 1998) 569, [astro-ph/](http://arxiv.org/abs/astro-ph/).

[31] C. S. Frenk, S. D. M. White, P. Bode, J. R. Bond, G. L. Bryan, R. Cen, H. M. P. Couchman, A. E. Evrard, N. Gnedin, A. Jenkins, A. Khokhlov, A. Klypin, J. F. Navarro, M. L. Norman, J. P. Ostriker, J. M. Owen, F. R. Pearce, U.-L. Pen, M. Steinmetz, P. A. Thomas, J. V. Villumsen, J. W. Wadsley, M. S. Warren, G. Xu, and G. Yepes, *The Santa Barbara Cluster Comparison Project: A Comparison of Cosmological Hydrodynamics Solutions*, ApJ 525 (Nov., 1999) 554–582, [astro-ph/](http://arxiv.org/abs/astro-ph/).

[32] S. Ettori, K. Dolag, S. Borgani, and G. Murante, *The baryon fraction in hydrodynamical simulations of galaxy clusters*, MNRAS 365 (Jan., 2006) 1021–1030, [astro-ph/](http://arxiv.org/abs/astro-ph/).

[33] S. Gottlöber, G. Yepes, A. Khalatyan, R. Sevilla, and V. Turchaninov, *Dark and baryonic matter in the MareNostrum Universe*, in The Dark Side of the Universe (C. Manoz and G. Yepes, eds.), vol. 878 of American Institute of Physics Conference Series, pp. 3–9, Nov., 2006. [astro-ph/](http://arxiv.org/abs/astro-ph/).

[34] R. A. Crain, V. R. Eke, C. S. Frenk, A. Jenkins, I. G. McCarthy, J. F. Navarro, and F. R. Pearce, *The baryon fraction of ΛCDM haloes*, MNRAS 377 (May, 2007) 41–49, [astro-ph/](http://arxiv.org/abs/astro-ph/).

[35] B. Rasheed, N. Bahcall, and P. Bode, *Where are the missing baryons in clusters?*, ArXiv e-prints (July, 2010) [arXiv:1007.1980](http://arxiv.org/abs/1007.1980).

[36] M. Lima and W. Hu, *Self-calibration of cluster dark energy studies: Counts in cells*, Phys. Rev. D. 70 (Aug., 2004) 043504, [astro-ph/](http://arxiv.org/abs/astro-ph/).

[37] C. E. Cunha and A. E. Evrard, *Sensitivity of galaxy cluster dark energy constraints to halo modeling uncertainties*, Phys. Rev. D. 81 (Apr., 2010) 083509, [arXiv:0908.0526](http://arxiv.org/abs/0908.0526).

[38] W. Hu, *Self-consistency and calibration of cluster number count surveys for dark energy*, Phys. Rev. D. 67 (Apr., 2003) 081304, [astro-ph/](http://arxiv.org/abs/astro-ph/).

[39] The Dark Energy Survey Collaboration, *The Dark Energy Survey*, ArXiv Astrophysics e-prints (Oct., 2005) [astro-ph/](http://arxiv.org/abs/astro-ph/).

[40] P. Predehl, *eRosita - X-ray all sky survey and clusters*, in Galaxy Clusters: Observations, Physics and Cosmology, July, 2010.

[41] A. Pillepich, C. Porciani, and T. H. Reiprich, *The X-ray cluster survey with eRosita: forecasts for cosmology, cluster physics and primordial non-Gaussianity*, MNRAS 422 (May, 2012) 44–69, [arXiv:1111.6587](http://arxiv.org/abs/1111.6587).

[42] R. P. Norris, *Overcoming the challenges of wide deep continuum surveys*, in Panoramic Radio Astronomy: Wide-field 1-2 GHz Research on Galaxy Evolution, 2009. [arXiv:0909.3666](http://arxiv.org/abs/0909.3666).

[43] S. J. E. Radford, R. Giovanelli, J. Glenn, T. A. Sebring, D. Woody, and J. Zmuidzinasa, *The cornell caltech atacama telescope (ccat)*, in 2009 USNC/URSI Annual Meeting, Jan, 2009.

[44] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, and et al., *LSST Science Book, Version 2.0*, ArXiv e-prints (Dec., 2009) [arXiv:0912.0201](http://arxiv.org/abs/0912.0201).

[45] R. Laureijs, J. Amiaux, S. Arduini, J. Augusteins, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, and et al., *Euclid Definition Study Report*, ArXiv e-prints (Oct., 2011) [arXiv:1110.3193](http://arxiv.org/abs/1110.3193).