

Understanding entropy in massive halos: The role of baryon decoupling

S. Molendi^{1,*}, M. Balboni^{1,2,3}, I. Bartalucci¹, S. De Grandi⁴, M. Gaspari⁵, F. Gastaldello¹, S. Ghizzardi¹,
L. Lovisari¹, G. Riva^{1,6}, M. Rossetti¹, and P. Tozzi⁷

¹ INAF – IASF Milano, Via A. Corti 12, I-20133 Milano, Italy

² Dipartimento di Scienza e Alta Tecnologia, Università dell’Insubria, Via Valleggio 11, I-22100 Como, Italy

³ DIFA – Università di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy

⁴ INAF – Osservatorio Astronomico di Brera, Via E. Bianchi 46, I-23807 Merate (LC), Italy

⁵ Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, I-41125 Modena, Italy

⁶ Dipartimento di Fisica, Università degli Studi di Milano, Via G. Celoria 16, I-20133 Milano, Italy

⁷ INAF – Osservatorio Astrofisico di Arcetri, Largo E. Fermi, I-50122 Firenze, Italy

Received 4 December 2024 / Accepted 13 February 2025

ABSTRACT

Aims. The goal of the work presented in this paper is to use observed entropy profiles to infer constraints on the accretion process in massive halos.

Methods. We compare entropy profiles from various observational samples with those generated by an updated version of the semi-analytical models developed in the early 2000s, that have been modified to reflect recent advancements in our understanding of large-structure formation.

Results. Our model reproduces the growing departure from self-similarity observed in data when moving inward in individual profiles and down in mass across different profiles. These deviations stem from a phase of extremely low gas content centered around $10^{13} M_{\odot}$. According to our model, halos at this mass scale are missing between 50% and 90% of their baryons, corresponding to a gas fraction ranging between 2% and 8%.

Conclusions. Baryon decoupling, the mechanism at the heart of our model, proves effective in explaining much of the behavior we sought to understand.

Key words. galaxies: abundances – galaxies: clusters: general – galaxies: clusters: intracluster medium – intergalactic medium

1. Introduction

It has long been understood that the high entropy of gas in massive halos ($M_h \gtrsim 10^{13} M_{\odot}$) is a direct consequence of the virialization process. In simple terms, gas falls onto the halo, and around the virial radius, it converts its kinetic energy into thermal energy by impacting other gas, that is at rest in the potential well of the system. Since the sound speed in the infalling gas is much smaller than the infall velocity, the impact generates a shock that raises the entropy of the gas. In stark contrast, once inside the halo, the processes operating on the gas are largely adiabatic, and both conduction and mixing are highly suppressed (see Molendi et al. 2023, and references therein), leaving the entropy unaltered. Thus, by measuring the entropy of the gas in massive halos, one can acquire valuable information on shock heating at the virial radius and, ultimately, on the accretion process.

The goal of the work presented in this paper is to use observed entropy profiles to infer constraints on the accretion process. In pursuing this objective, we adopted a semi-analytical approach. As discussed by a few authors two decades ago (Tozzi & Norman 2001; Voit et al. 2003; Voit 2005), the process by which entropy is generated is a simple one: the bulk of the entropy is generated at the virial shock. Early semi-analytical work (Tozzi & Norman 2001; Voit et al. 2003; Voit 2005) met with some success, as the predicted slopes broadly

aligned with observations of massive clusters. However, with the advent of more precise measurements (see Pratt et al. 2006, 2010), it became clear that even beyond the core region, where processes such as radiative cooling and heating from the active galactic nucleus (AGN) heavily influence gas thermodynamics, these simple models fall short of explaining critical features. Specifically, lower mass systems exhibited entropy profiles that, when rescaled self-similarly, were offset relative to those of more massive systems. This discrepancy indicates that the self-similar scaling inherent in the models fails to capture some essential physical processes. To address these limitations, semi-analytical studies were quickly supplemented (Borgani et al. 2005; Voit et al. 2005) and later largely supplanted by simulation-based approaches (see Altamura et al. 2023, and references therein). While simulations have successfully overcome several obvious shortcomings of earlier models, they introduce their own set of challenges, such as computational complexity, resolution limitations, and uncertainties in modeling sub-grid physics such as feedback and cooling processes. Here, we aim to determine whether and to what extent an updated version of the semi-analytical models from the early 2000s, incorporating significant advancements in the understanding of large-structure formation, can account for the extensive observational evidence gathered over the past two decades.

As mentioned earlier, reproducing observed entropy profiles represents a step toward a broader goal: gaining deeper

* Corresponding author: silvano.molendi@inaf.it

insights into the accretion process. This focus distinguishes our approach from much of the simulation work in recent years (e.g., Barnes et al. 2017; Altamura et al. 2023; Braspenning et al. 2024). Simulations typically employ detailed descriptions of numerous processes governing the formation and evolution of large-scale structures in the Universe and are aiming at predicting a wide array of baryonic properties across different states and scales, with entropy profiles being just one among many. The two approaches are complementary: findings based on simulations offer a framework for conducting targeted studies of specific aspects of structure formation, while semi-analytical efforts, such as the one presented in this paper, provide insights into particular processes that can subsequently be incorporated into simulations.

The paper is organized as follows. Sect. 2 reviews the process of entropy generation as outlined in the semi-analytical studies of the early 2000s. In Sect. 3 we derive radial entropy profiles by combining the entropy-halo mass relation obtained in Sect. 2 with the radial mass distribution. We also show that these profiles align reasonably well with observations at large radii, though they fall significantly below observed values at smaller radii. In Sect. 4, we illustrate that incorporating pre-heating into the entropy generation process reduces the discrepancy between the model and the observed data, but it does not resolve it. In Sect. 5, we modify the self-similar accretion model by decoupling the accretion of baryons from that of dark matter. We demonstrate that the ensuing break in self-similarity can be recovered in part by rescaling entropy by the gas fraction. In Sect. 6, we compare entropy and density profiles derived assuming baryon decoupling with observed profiles from three different samples of galaxy clusters and one sample of X-ray bright galaxy groups, overall finding good agreement. We also argue that the difference between observed profiles of other, less X-ray bright groups and those predicted from our model can be resolved by invoking heating from the central AGN. In Sect. 7, we compare predictions from our model with those from other theoretical work. In Sects. 8 and 9 we respectively discuss and summarize our findings.

Throughout the paper we assume a Λ cold dark matter cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$. Furthermore, we define R_Δ as the radius inside which the cluster mass density is Δ times the critical density of the Universe, and M_Δ is the halo mass within R_Δ .

2. Entropy generation

We start by reviewing the spherical accretion model discussed by several authors (e.g., Tozzi & Norman 2001; Voit 2005) two decades ago. In simple terms, the accretion of gas in massive halos can be divided into three phases: a first phase where gas falls onto the halo, a second where it is shock heated and a third where it settles into near hydrostatic equilibrium within the halo. An important point is that the processes occurring in the first and third phase are assumed to be adiabatic and therefore leave the entropy of the gas unaltered. The variation in entropy occurs only in the second phase, this crucial aspect is nicely represented in Fig. 2 of Tozzi & Norman (2001). Furthermore, the whole process is assumed to be radially symmetric and can be described in terms of smooth accretion of spherical shells that are shocked at a single radius, namely the virial radius.

2.1. Baseline model

In the simplest case, the gas starts off cold and acquires all its entropy by going through a shock at the virial radius. Under these

conditions the equations describing the shock can be reduced to a set of four, (see Eqs. 67–70 in Voit 2005). We re-propose them here.

The mass flow through the accretion radius is

$$\dot{M}_g = 4\pi R_v^2 \rho_{g,u} v_v, \quad (1)$$

where \dot{M}_g is the gas accretion rate, R_v is the virial radius which coincides with the shock radius, $\rho_{g,u}$ is the upstream gas density and v_v is the gas velocity at the virial radius.

The velocity of the accreted gas is described with

$$v_v^2 = \frac{GM_h}{R_v}, \quad (2)$$

where M_h is the halo mass. Eq. (2) is derived by equating the kinetic energy at the virial radius to the variation in potential energy between the turn-around radius, where velocity is zero, and the accretion radius, under the assumption the turn around radius is, $R_{ta} = 2R_v$.

The equation for energy conservation is

$$k_b T = \frac{1}{3} \mu m_p v_v^2, \quad (3)$$

where k_b is the Boltzmann constant, μ is the mean molecular mass of the gas and m_p is the proton mass. Eq. (3) describes the conversion of the upstream kinetic energy (right hand side) into downstream thermal energy (left hand side).

For gas compression the equation is

$$\rho_{g,d} = 4\rho_{g,u}, \quad (4)$$

where $\rho_{g,d}$ and $\rho_{g,u}$ are respectively the gas density downstream and upstream of the shock. Eq. (4) relates the upstream to the downstream density of the gas under the assumption of high Mach number, in which case the downstream density is 4 times the upstream one. We note that this set of equations provides an approximate description of the process valid for strong shocks. For instance, it does not account for the compression and adiabatic heating undergone by the infalling gas before being shocked (see Tozzi & Norman 2001, for details). We nonetheless work within the framework provided by Eqs. (1)–(4) and evaluate the impact of the approximations at the appropriate time.

From Eqs. (1), (2) and (4) one can rewrite the downstream density as a function of halo mass, gas accretion rate and accretion radius:

$$\rho_{g,d} = \frac{1}{\pi G^{1/2}} R_v^{-3/2} M_h^{-1/2} \dot{M}_g. \quad (5)$$

Similarly, from Eqs. (2) and (3), the gas temperature may be expressed as a function of halo mass and accretion radius:

$$k_b T_v = \frac{1}{3} \mu m_p \frac{GM_h}{R_v}. \quad (6)$$

Making use of Eqs. (5) and (6), one can write the entropy¹, which following Balogh et al. (1999), we define as

$$K \equiv \frac{k_b T}{\mu m_p \rho_g^{2/3}}, \quad (7)$$

¹ It is important to note that in studies of the intra-cluster medium (ICM), the term “entropy” refers to a quantity related to the thermodynamic specific entropy, $s: s = 3/2 k_b \ln K + const$ (see Eqs. 9 and 10 in Donahue & Voit 2022). Additionally, specific entropy is an intensive variable.

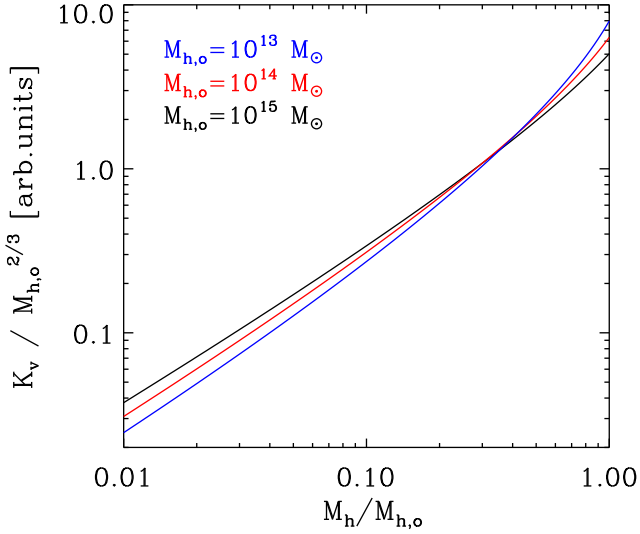


Fig. 1. Scaled entropy versus halo mass relation for three halos of $10^{13} M_\odot$, $10^{14} M_\odot$, and $10^{15} M_\odot$ at $z = 0$, respectively.

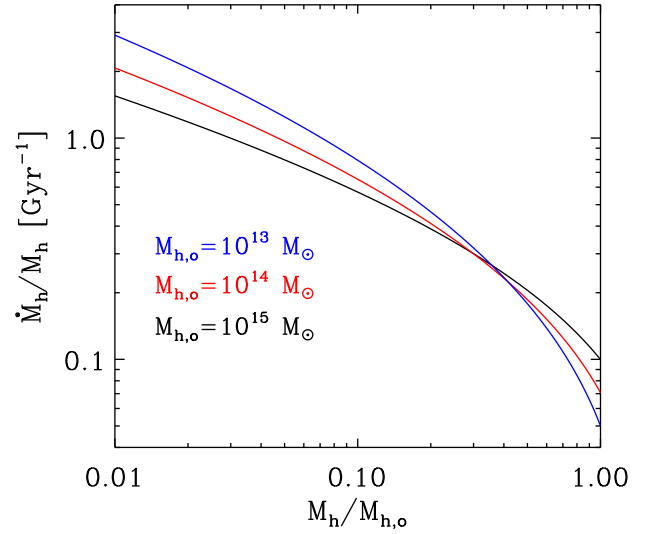


Fig. 2. Specific accretion rate versus halo mass relation for three halos of $10^{13} M_\odot$, $10^{14} M_\odot$, and $10^{15} M_\odot$ at $z = 0$, respectively.

in the form

$$K_v = \frac{1}{3} (\pi G^2)^{2/3} \frac{M_h^{4/3}}{\dot{M}_g^{2/3}}. \quad (8)$$

Assuming that the gas fraction of the shells undergoing accretion is constant and equal to the cosmic baryon value, we may write

$$\dot{M}_g = f_b \dot{M}_h, \quad (9)$$

where f_b is the cosmic baryon fraction and \dot{M}_h is the matter accretion rate on the halo. Thus, Eq. (8) can be rewritten in the form

$$K_v = \frac{1}{3} (\pi G^2)^{2/3} f_b^{-2/3} \left(\frac{M_h}{\dot{M}_h} \right)^{2/3} \dot{M}_h^{2/3}. \quad (10)$$

2.2. Entropy versus mass relation

Within the framework we have explored, the relationship between entropy and halo mass is described by Eq. (10). To estimate the halo mass, M_h , and its growth rate, \dot{M}_h , we make use of the analytical model described in [Salcido et al. \(2020\)](#) (see their Eq. 12 and Fig. 3). In Fig. 1 we show the case of three halos of $10^{13} M_\odot$, $10^{14} M_\odot$, and $10^{15} M_\odot$, respectively, at $z = 0$. The entropy has been rescaled by dividing by $M_{h,o}^{2/3}$, where $M_{h,o} \equiv M_h(z = 0)$.

There are two key points to highlight. First, the growth of entropy with halo mass for the three systems is similar. This arises from the weak dependency of the specific accretion rate, \dot{M}_h/M_h , on $M_{h,o}$, as illustrated in Fig. 2. Second, entropy increases more rapidly with decreasing $M_{h,o}$. This behavior can again be attributed to the relationship between entropy and the specific accretion rate. As shown in Fig. 2, the specific accretion rate decreases more significantly with decreasing $M_{h,o}$.

An interesting way of characterizing the entropy versus halo mass relationship is through the logarithmic slope $d \ln K_v / d \ln M_h$ (see [Voit et al. 2003](#)). It can be shown, see Appendix A for a derivation, that the following relation holds:

$$\frac{d \ln K_v}{d \ln M_h} = \frac{2}{3} \left(1 + \frac{1}{\alpha} - \frac{d \ln \alpha}{d \ln M_h} \right), \quad (11)$$

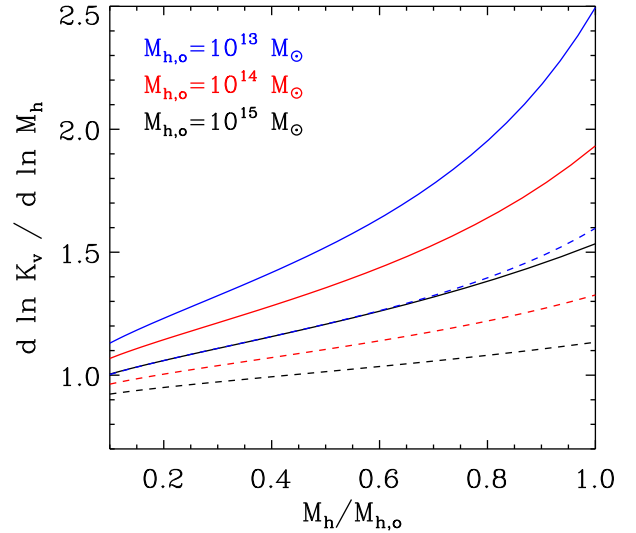


Fig. 3. Logarithmic slope, $d \ln K_v / d \ln M_h$, versus halo mass relation for three halos of $10^{13} M_\odot$, $10^{14} M_\odot$, and $10^{15} M_\odot$ at $z = 0$, respectively, as estimated from Eq. (11). Dashed lines were computed neglecting the $d \ln \alpha / d \ln M_h$ term.

where $\alpha \equiv d \ln M_h / d \ln t$.

The logarithmic slope $d \ln K_v / d \ln M_h$ is reported in Fig. 3 for three halos of $10^{13} M_\odot$, $10^{14} M_\odot$ and $10^{15} M_\odot$ at $z = 0$ respectively. As pointed out in [Voit et al. \(2003\)](#), the increase in slope at large values of $M_h/M_{h,o}$ is related to a decrease in accretion rate, see Fig. 2.

3. Radial entropy profiles

In this section we derive radial entropy profiles from the model presented in Sect. 2. We do this in three steps. First we make use of Eq. (10) that relates entropy to halo mass, then we assume a radial mass distribution, finally we combine the two to derive the radial entropy profile.

As we shall see later, in the last step we consider two distinct approaches to derive entropy profiles. The first is a straightforward combination of the functional forms: $K(R) =$

$K_v(M_h(R))^2$, the second imposes that gas within the halo follow hydrostatic equilibrium.

3.1. Mass versus radius relation

Several relations have been proposed in the literature to describe the distribution of mass within a halo (e.g., [Ettori et al. 2019](#), and references therein), we assume one of the mostly commonly used, the Navarro, Frank, and White (NFW) profile ([Navarro et al. 1997](#)). We write the mass within a radius R as:

$$M_h(<R) = 4\pi\rho_o R_s^3 \left[\ln(1+x) - \frac{x}{1+x} \right], \quad (12)$$

where ρ_o is a reference density, $x = R/R_s$, and R_s is the scale radius. Furthermore, we fix the shape of the profile by assuming that c_{200} , defined as: $c_{200} \equiv R_{200}/R_s$, assumes the value $c_{200} = 4$. We point out that variation of c_{200} within reasonable limits (e.g., [Ettori et al. 2019](#)) do not change results presented later. Furthermore, for a Λ CDM Universe, with $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$, the virial overdensity at $z = 0$ is $\Delta_v(z = 0) \sim 100$ (see Eq. 44 in [Voit 2005](#)) and the virial radius is: $R_v \sim 1.25 R_{200}$. Thus, the concentration computed using the virial radius, which is defined as: $c_v \equiv R_v/R_s$, turns out to be: $c_v \sim 5$.

3.2. Analytical derivation of the slope of the entropy profile

Interestingly, the logarithmic slope of the NFW mass profile can be expressed as

$$\frac{d \ln M_h}{d \ln R} = \frac{x^2}{(1+x)^2} \frac{1}{\ln(1+x) - \frac{x}{1+x}}. \quad (13)$$

By writing:

$$\frac{d \ln K}{d \ln R} = \frac{d \ln K_v}{d \ln M_h} \times \frac{d \ln M_h}{d \ln R} \quad (14)$$

and substituting Eqs. (11) and (13) into Eq. (14), we obtained

$$\frac{d \ln K}{d \ln R} = \frac{2}{3} \left(1 + \frac{1}{\alpha} - \frac{d \ln \alpha}{d \ln M_h} \right) \times \frac{x^2}{(1+x)^2} \frac{1}{\ln(1+x) - \frac{x}{1+x}}. \quad (15)$$

This expression allows us to perform a direct computation of the logarithmic slope of the radial entropy profile, which we report in Fig. 4, for three halos of $10^{13} M_\odot$, $10^{14} M_\odot$ and $10^{15} M_\odot$ at $z = 0$ respectively.

As can be appreciated, the slopes of the three entropy profiles differ considerably. Since the term in Eq. (15) coming from Eq. (13) does not depend on mass (the NFW profile is self-similar) the difference between the three curves is entirely due to the term coming from Eq. (11) and more specifically to α and $d \ln \alpha / d \ln M_h$.

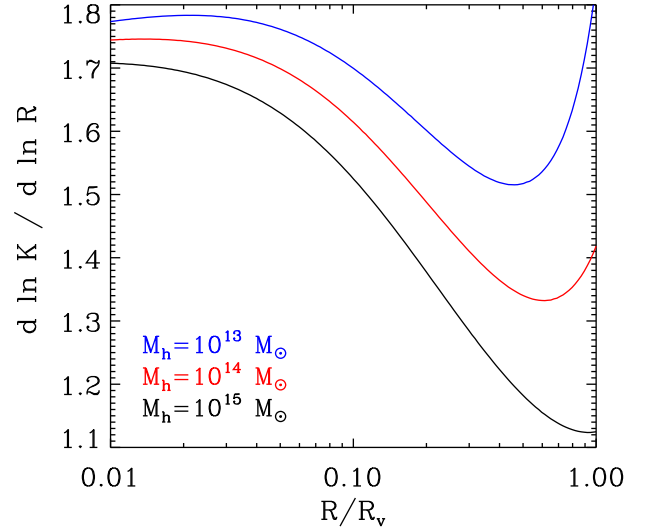


Fig. 4. Logarithmic slope of the radial entropy profile for three halos of $10^{13} M_\odot$, $10^{14} M_\odot$, and $10^{15} M_\odot$ at $z = 0$, respectively, as estimated from Eq. (15). The radial coordinate has been normalized to the virial radius.

3.3. Radial entropy profiles

As discussed at the beginning of this section, we compute entropy and other thermodynamic profiles in two different ways. In the first, we perform a straightforward combination of the functional forms: $K(R) = K_v(M_h(r))$, in the second we impose hydrostatic equilibrium on the shocked gas. We note that, in the first case, the gas fraction averaged within the shock radius remains unchanged with respect to the one we impose on the infalling shells. In the second, gas is allowed to reorganize itself. As in [Tozzi & Norman \(2001\)](#), we solve for hydrostatic equilibrium from the outside. We take the values of thermodynamic values at the shock radius as boundary conditions and numerically solve a set of three equations: the hydrostatic equilibrium equation, the gas mass conservation equation and the entropy conservation equation. This allows us to derive radial profiles for all thermodynamic variables. A detailed description of the procedure is provided in Appendix B.

In Fig. 5 we plot radial entropy profiles for three different halo masses: $10^{13} M_\odot$ (left panel), $10^{14} M_\odot$ (central panel) and $10^{15} M_\odot$ (right panel). We note that the radius has been normalized to the shock (virial) radius and the entropy has been divided by the entropy at the virial radius, K_v . As we may surmise from the figure, the two solutions, shown respectively in blue and black, do not differ much from one another, both become steeper as we move inward as highlighted by a comparison with a reference power-law profile (red line). Numerical computation of the $d \ln K / d \ln r$ curves, obtained from the solution derived through a straightforward combination of functional forms, shows them to be in excellent agreement with the ones computed analytically (see Eq. 15) and shown in Fig. 4.

It is worth noting that, for the cluster mass range, $10^{14} M_\odot - 10^{15} M_\odot$, these results, are consistent with those derived from non-radiative cosmological simulations performed using both Lagrangian smoothed particle hydrodynamics (SPH: see [Borgani et al. 2002](#); [Rasia et al. 2004](#)) and Eulerian adaptive mesh refinement codes (see [Voit et al. 2005](#)). This tells us something quite important: a simple one dimensional model, based on mass and energy conservation arguments, provides a description of the entropy generation process which is, at least to first

² The v subscript is present on the right hand side of the equation to highlight that this is the entropy generated at a time when the virial radius was going through the radius R , it is dropped on the left hand side of the equation because what we wish to emphasize in this case is that $K(R)$ is the entropy at a generic radius R .

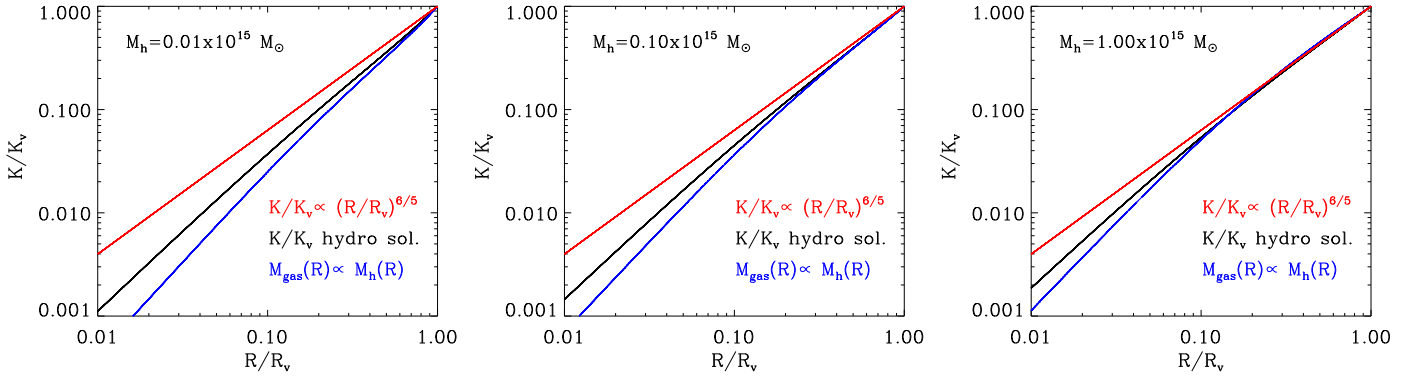


Fig. 5. Predicted entropy profile for a halo mass of $10^{13} M_{\odot}$ (left panel), $10^{14} M_{\odot}$ (central panel), and $10^{15} M_{\odot}$ (right panel). The radius has been normalized to the shock (virial) radius. All profiles are normalized at the shock radius. In red we plot a reference powerlaw with slope 1.2, in blue the profile derived from a combination of the functional forms: $K(R) = K_v(M_h(R))$ and in black the one derived by solving the hydrostatic equilibrium equation.

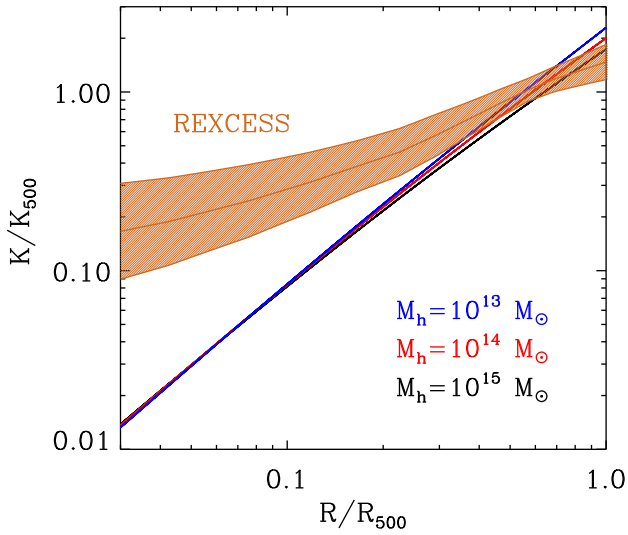


Fig. 6. Predicted entropy profiles for a halo mass of $10^{13} M_{\odot}$ (blue), $10^{14} M_{\odot}$ (red) and $10^{15} M_{\odot}$ (black) respectively, compared with the observed mean entropy profile and intrinsic scatter (orange shaded region) from the REXCESS sample of galaxy clusters, (Pratt et al. 2010), which features a median mass of $M_{500} \sim 2.5 \times 10^{14} M_{\odot}$. To ease comparison with observational data, the predicted entropy has been rescaled by K_{500} and radius by R_{500} .

order, comparable to that which can be achieved through a much more detailed and computationally intensive analysis.

3.4. A first comparison with observations

A detailed comparison with observations is presented in Sect. 6. Here, we wish to ascertain if the simple model outlined above provides estimates that are at least broadly consistent with available data. In Fig. 6 we compare the entropy profiles derived imposing hydrostatic equilibrium, and already presented in Fig. 5, with the median profile derived from the REXCESS sample of galaxy clusters (Pratt et al. 2010). To ease comparison with observational data, entropy has been rescaled by K_{500} , which is defined as in Pratt et al. (2010), see their Eq. (3), and radius by R_{500} . At large radii, $R \sim R_{500}$, the predicted entropy is broadly consistent with the measured one, however, as we move inwards, the behavior of measured and predicted profiles diverge, the latter featuring a much steeper decline than the for-

mer. In more physical terms, the entropy generated purely from gravitational shocks is consistent with the observed one in the outskirts but severely under-predicts it at smaller radii. Both agreement and disagreement are rich in significance. The agreement suggests that, to first order, a description of the virialization process that is as crude as the one we presented, can reproduce observed data. This implies that the entropy generation process is probably not too dissimilar from a total conversion of kinetic energy into thermal energy at a single shock radius³. It is worth pointing out that, although highly desirable, such an outcome is by no means guaranteed. Indeed, in light of the limited understanding of the physics of the accreted gas compounded by the dearth of observational data, the entropy generation process could just as easily have taken very different forms. Just as an example, Vazza et al. (2009) present simulations of the accretion process on massive halos in which thermalization occurs over an extremely broad radial range, with kinetic energy exceeding thermal energy for radii larger than $0.75 R_v$ and remaining as high as 30% of the thermal energy down to $\sim 0.3 R_v$.

Regarding the disagreement at smaller radii, it has been known since the time scaling relations were first established (Kaiser 1986, 1991), that non-gravitational forms of heating, including AGN feeding/feedback (e.g., Gaspari et al. 2020, for a review), may provide an important contribution. Some two decades ago, several authors (e.g., Tozzi & Norman 2001; McCarthy et al. 2004; Voit 2005) developed models in which the self-similar shape of entropy profiles was at least partially modified by assuming that the infalling gas was pre-heated to a fraction of the virial temperature. In the following section we revisit pre-heating.

4. Pre-heating

Pre-heating is the process by which gas is heated prior to being accreted. As pointed out in Tozzi & Norman (2001), the pre-heated gas contracts in an approximately adiabatic fashion as long as its temperature is larger than the virial temperature of the halo onto which it is falling. Under such conditions no virial shock forms and the entropy of the accreted gas is essentially that of the pre-heated gas. As soon as the virial temperature grows

³ As the attentive reader will have undoubtedly realized, the slopes of observed and predicted profiles differ significantly. This point is discussed at length in Sect. 6.1.

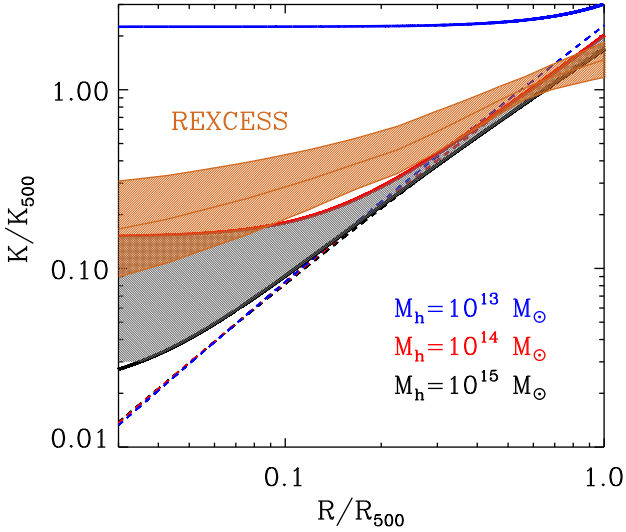


Fig. 7. Predicted entropy profiles for a halo mass of $10^{13} M_{\odot}$ (blue), $10^{14} M_{\odot}$ (red), and $10^{15} M_{\odot}$ (black). Dashed lines show entropy from gravitational processes only, solid lines include the effects of pre-heating. The mean entropy profile and intrinsic scatter from the REXCESS sample of galaxy clusters, (Pratt et al. 2010), which features a median mass of $M_{500} \sim 2.5 \times 10^{14} M_{\odot}$, is reported as an orange shaded region. The region between the $10^{14} M_{\odot}$ and $10^{15} M_{\odot}$ profiles is shown as a gray shaded region.

beyond the temperature of the infalling pre-heated gas, an accretion shock forms.

As discussed in a seminal paper (Kaiser 1991), pre-heating is likely associated to the galaxy formation process. Pre-heating introduces a flat region in the radial entropy profile. The gas residing in such an iso-entropic region has been accreted without being shocked and retains the entropy it had prior to its accretion. As shown in Voit et al. (2003), the entropy of the gas including the effects of preheating K_{WP} , can be expressed reasonably well as the sum of the entropy of the gas without pre-heating, K_v , and the entropy of the preheated gas, K_* multiplied by a correction factor:

$$K_{WP} \simeq K_v + 0.84K_*. \quad (16)$$

A crucial point is the estimate of the pre-heating entropy. Several authors (e.g., Tozzi & Norman 2001; McCarthy et al. 2002; Ponman et al. 2003; Voit et al. 2003) have proposed values in the range $1-5 \times 10^{33} \text{ erg cm}^2 \text{ g}^{-5/3}$. These estimates were arrived at by comparing observed scaling relations or entropy profiles with models including pre-heating. In Sect. 5 we discuss a different approach to estimating K_* .

In Fig. 7 we show entropy profiles for a halo mass of $10^{13} M_{\odot}$, $10^{14} M_{\odot}$ and $10^{15} M_{\odot}$ respectively. The entropy of the pre-heated gas is set to the value generated through gravitational processes at the virial radius of a $10^{13} M_{\odot}$ halo: $K_* = K_v(M_h = 10^{13} M_{\odot}) \sim 3.5 \times 10^{33} \text{ erg cm}^2 \text{ g}^{-5/3}$, for all three halo masses (for a justification of this choice see Tozzi & Norman 2001, and Sect. 5.6). As can be observed, the effect of pre-heating strongly depends on the halo mass. For the $10^{13} M_{\odot}$ system, it provides the dominant contribution at all radii. At the opposite mass end, $10^{15} M_{\odot}$, its effects can be appreciated only for $R < 0.1 R_{500}$. In somewhat different terms, we can say that pre-heating produces a departure from self-similarity that is strongly halo mass dependent.

As far as the comparison with observations is concerned, pre-heating does significantly reduce the differences between mod-

eled and observed profiles but it does not allow for agreement between the two. With the pre-heating modification to the model, predicted entropy profiles of clusters fall between the thick red ($10^{14} M_{\odot}$) and blue ($10^{15} M_{\odot}$) curves, observed ones do not. In particular, in observed clusters, the excess with respect to the gravity only model (dashed lines in Fig. 7) is present up to intermediate radii $R \sim 0.5 R_{500}$, while in modeled systems, with halo masses $M_h \gtrsim 10^{14} M_{\odot}$, it is found only at smaller radii.

It is worth pointing out that cooling has not been introduced in the gravity only or the pre-heating modified model, this limits the applicability of the models to regions beyond the core, $R \gtrsim 0.1 R_{500}$, where the cooling time of the gas is sufficiently long, we return to this point in Sect. 6.2, where we perform a more detailed comparison with observations.

5. Baryon decoupling

The pre-heating model described in Sect. 4 and in the literature (see Voit 2005, and references therein) manifestly fails to reproduce the shape of entropy profiles in clusters.

In this section, we investigate a different description of feedback processes in the hope that it will lead to a better agreement with observations. In doing so, we take stock of the considerable advances that have taken place over the past two decades.

5.1. Advancements

One of the most important things we have learned is that the halo mass range between a few $10^{12} M_{\odot}$ to $10^{13} M_{\odot}$ plays a critical role in the evolution of baryons. This is where star formation shuts down (e.g., Leauthaud et al. 2012; Coupon et al. 2012, 2015; Behroozi et al. 2013, 2019; Moster et al. 2013; Cowley et al. 2018; Legrand et al. 2019; Girelli et al. 2020; Shuntov et al. 2022). Equally important is the fact that the position of the star formation peak is largely insensitive to cosmic time, for $z \lesssim 4$ as shown in (Behroozi et al. 2013, 2019; Legrand et al. 2019). The cold baryon reservoir from which stars form is suddenly and dramatically emptied, quite possibly through feedback from the central AGN that heats up and disperses the cold gas. The same process likely leads to a substantial reduction in the accretion on the central black hole. However, the significantly different scales on which star formation and AGN feeding operate likely lead to a delay between the shutting down of the latter with respect to the former. It is worth pointing out that the observed correlation between stellar mass and Black Hole mass, (see Zhang et al. 2024a, and references therein) as well as the one between hot halo properties and super-massive black hole (SMBH) mass (Gaspari et al. 2019) are consistent with this scenario.

Relatively little is known of the heated gas, it is most likely in a hot phase either within the halo, in which case it is referred to as the hot phase of the circum galactic medium (see Tumlinson et al. 2017, and references therein) or outside it, in which case it takes the name of warm hot intergalactic medium (see Nicastro et al. 2022, and references therein). Another important piece of the puzzle that has emerged over the past decade is that the gas fraction, f_g , in massive halos, correlates with halo mass (see Eckert et al. 2021, and references therein), with lower mass systems featuring a significantly smaller baryon fraction than more massive ones. Putting the two pieces together, we may surmise that somewhere around $\sim 10^{13} M_{\odot}$, the feedback events responsible for shutting down the accretion also lead to a significant evacuation of baryons from the halo. The substantial scatter in gas fraction observed in

groups most likely reflects the stochastic nature of the process. Feedback processes operate on a wide range of halo masses, and to avoid confusion, we refer to feedback on $\sim 10^{13} M_\odot$ halos as early-time feedback, while those operating on more massive halos are described as late-time feedback. Within this framework, the increase in gas fraction observed for halo masses larger than $\sim 10^{13} M_\odot$, corresponds to a phase of slow re-accretion of gas previously ejected by less massive halos. It is worthwhile reminding our readers that we do have indirect evidence of the accretion of pre-enriched hot gas in massive halos through metals (see Molendi et al. 2024).

5.2. A simple prescription

Having taken stock of what has been learned over the past two decades, we modify the self-similar accretion model described in Sects. 2 and 3 by decoupling the accretion of baryons from that of dark matter. Baryon decoupling provides a very simple modification, one might argue the simplest possible modification, to satisfy the constraints described in Sect. 5.1. In this section we develop a model based on baryon decoupling and see how far it takes us in interpreting the thermodynamic properties of massive accretors. Before proceeding with the actual modification it is instructive to perform a little bit of deductive reasoning to see how differences between measured and predicted entropy might be explained by baryon decoupling.

Our approach is somewhat analogous to methods used by geologists when studying rock stratification. Similarly, in massive halos, gas stratifies by entropy, with gas accreted earlier positioned closer to the center, while gas accreted later resides in the outer regions. Comparison with observations, see Figs. 6 and 7, shows that the entropy generated when the halo was a fraction of its current mass is higher than expected. From Eq. (7) we see that a higher entropy might result from a lower density. A lower density likely follows from a lower gas accretion rate with respect to the one assumed in Eq. (9), namely:

$$\dot{M}_g < f_b \dot{M}_h, \quad (17)$$

or in other words, from baryon decoupling.

At least qualitatively, baryon decoupling goes in the direction required to reconcile the self-similar model with observations. A more quantitative assessment requires the definition of a gas fraction versus halo mass relation. To this end, we assume decoupling to take place at a specific halo mass, roughly $10^{13} M_\odot$. At much smaller and larger masses we assume baryon accretion to be coupled to dark matter accretion. A simple way to describe this process mathematically is through the formula:

$$\log(f_g) = \log(f_b) + l_n \exp \left[-\frac{1}{2} \left(\frac{\log M_h - l_m}{l_\sigma} \right)^2 \right], \quad (18)$$

where $\log(f_g)$ is the base ten logarithm of f_g , $\log(f_b)$ is the base ten logarithm of f_b , and $10^{l_m} M_\odot$ is the halo mass at which baryon decoupling is strongest (i.e., where the deviation of the gas fraction, f_g , from the cosmic baryon fraction reaches its maximum). The second parameter quantifies the decoupling: 10^{l_n} is the ratio of the gas fraction over the cosmic baryon fraction at $M_h = 10^{l_m} M_\odot$, and finally l_σ determines the halo mass range over which the gas fraction is decoupled from the cosmic baryon fraction.

In Fig. 8 we show, as a red solid line, a realization of the gas fraction versus halo mass relation for a representative choice of parameters. Additionally, the figure shows the gas mass versus

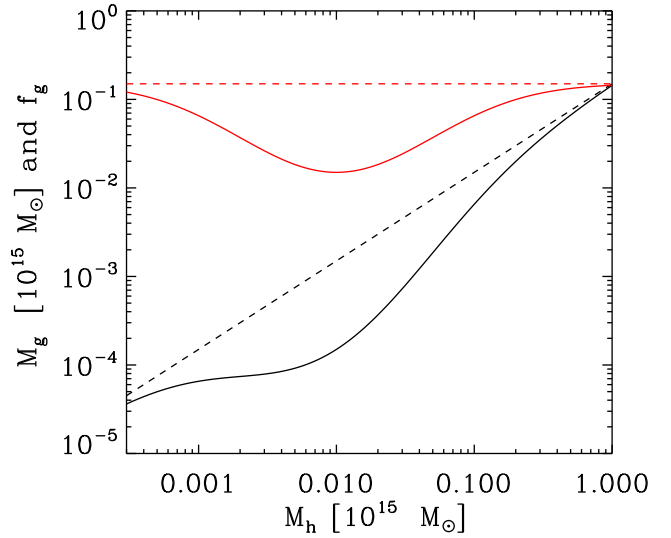


Fig. 8. Gas fraction, f_g , shown in red and gas mass, M_g , shown in black, as a function of halo mass. The dashed lines show the case where $f_g = f_b$ and $M_g = f_b M_h$ at all halo masses. The solid lines show the case modeled by Eq. (18) with parameters chosen as follows: $l_m = 13$, $l_n = -1$, and $l_\sigma = 0.6$.

halo mass relation as a black solid line, highlighting another key relation. We note how, despite the reduction in gas fraction with increasing halo mass for $M_h < 10^{l_m} M_\odot$, the gas mass always increases with halo mass.

Our implementation of baryon decoupling provides a very simple modification of the gravitational entropy generation process. It identifies a critical mass at which baryon decoupling operates, $10^{l_m} M_\odot$, and allows for variations in the strength and the range over which the modification operates, respectively through the l_n and l_σ parameters. In the next subsection we develop a model based on baryon decoupling and see how far it takes us in interpreting the thermodynamic properties of massive accretors.

5.3. Entropy versus mass relation with decoupling

In this subsection we investigate the effects of baryon decoupling on the entropy versus halo mass relation. The first thing to note is that, since the gas fraction of the accreted shells is no longer constant, Eq. (10) is not valid and we have to fall back on Eq. (8). Furthermore, the \dot{M}_g term, which in the case of $f_g = f_b$ is simply $\dot{M}_g = f_g \dot{M}$, for a non-constant f_g takes the form

$$\dot{M}_g = f_g \dot{M}_h \left(1 + \frac{d \ln f_g}{d \ln M_h} \right). \quad (19)$$

Since the logarithmic derivative $d \ln f_g / d \ln M_h$ is negative for $M_h < 10^{l_m} M_\odot$ and positive for $M_h > 10^{l_m} M_\odot$, the term in parenthesis on the right hand side of Eq. (19) reaches a minimum at $M_h = 10^{l_m} M_\odot$. Multiplication by the $f_g \dot{M}_h$ term shifts the minimum to somewhat lower halo masses. This is clearly seen in Fig. 9, where we plot \dot{M}_g as a function of halo mass for a halo of $10^{15} M_\odot$ at $z = 0$.

In more physical terms, baryon decoupling reduces the gas accretion rate with respect to the $f_g = f_b$ case. Most importantly, it introduces a critical mass $\sim 1/4 10^{l_m} M_\odot$, below which the gas accretion rate decreases and above which it increases, with increasing halo mass. For halos that grow well beyond the

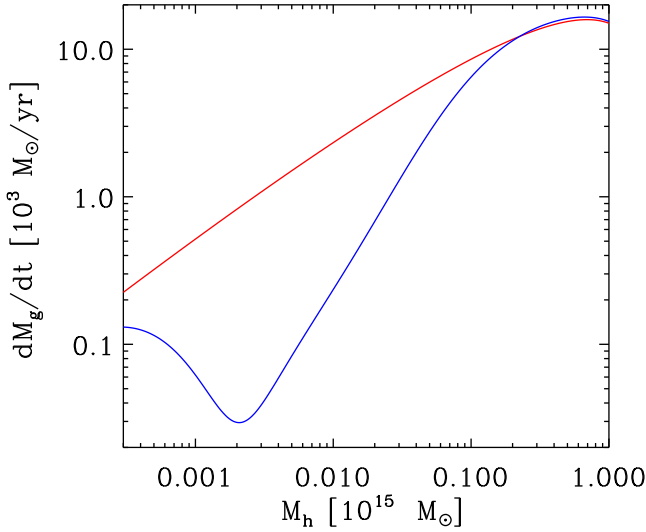


Fig. 9. Gas accretion rate versus halo mass relation for a halo of $10^{15} M_{\odot}$ at $z = 0$. The blue and red lines have been computed assuming f_g follows respectively Eq. (18) and $f_g = f_b$.

critical mass, the accretion rate slowly converges to $f_g \dot{M}_h$, see Fig. 8. In other words, baryon decoupling does not halt accretion, it delays it.

As shown in Fig. 10, baryon decoupling induced deviations propagate from the gas accretion rate to the entropy. Comparison of the solid curves derived assuming f_g follows Eq. (18) with dashed ones for which $f_g = f_b$, shows that baryon decoupling has a strong impact on the rescaled entropy versus halo mass relation. The reduction in accretion rate leads to an increase in entropy. As expected, the effect is largest for the least massive halos ($3 \cdot 10^{13} M_{\odot}$, $10^{14} M_{\odot}$ at $z = 0$), where the reduction in accretion rate is strongest. Another important point that is clearly observed in Fig. 10 is that baryon decoupling leads to a break in self-similarity; once again it is the less massive systems that deviate most. Finally, the change in slope observed for $M_h/M_{h,o} \sim 0.2$, for the cases $M_{h,o} = 3 \cdot 10^{13} M_{\odot}$ and $M_{h,o} = 10^{14} M_{\odot}$, is related to the minimum in \dot{M}_g observed in Fig. 9.

5.4. Gas fraction rescaling

With a little algebra Eq. 8 can be rewritten in the following fashion:

$$K_v = \frac{1}{3} (\pi G^2)^{2/3} \left(\frac{M_h}{f_g} \right)^{2/3} \left[\frac{\dot{f}_g}{f_g} + \frac{\dot{M}_h}{M_h} \right]^{-2/3}. \quad (20)$$

This form highlights the similarity between the role played by M_h and f_g . We recall that the self-similar scaling with respect to mass follows from the fact that \dot{M}_h/M_h depends only weakly on mass, see Fig. 2. Something similar may happen with f_g if \dot{f}_g/f_g is independent or at least weakly dependent on f_g . Having identified a specific prescription for f_g , see Eq. (18), we can verify to what extent f_g scaling might work. In Fig. 11 we plot \dot{M}_h/M_h , \dot{f}_g/f_g and $\dot{f}_g/f_g + \dot{M}_h/M_h$ as a function of halo mass for four halos of $3 \cdot 10^{13} M_{\odot}$, $10^{14} M_{\odot}$, $3 \cdot 10^{14} M_{\odot}$ and $10^{15} M_{\odot}$ at $z = 0$ respectively. The \dot{M}_h/M_h (dashed lines) are very similar to one another, this has already been shown in Fig. 2. It is this behavior that is responsible for self-similarity in the $f_g = f_b$ case. The \dot{f}_g/f_g curves (dotted lines) show somewhat larger vari-

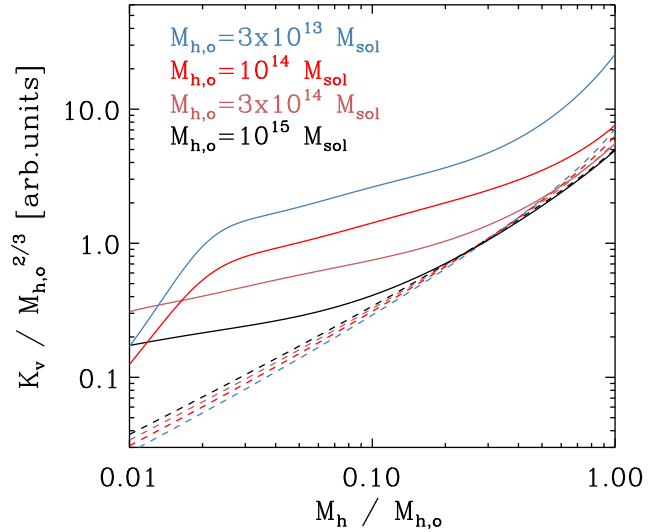


Fig. 10. Rescaled entropy versus halo mass relation for four halos of $3 \cdot 10^{13} M_{\odot}$, $10^{14} M_{\odot}$, $3 \cdot 10^{14} M_{\odot}$, and $10^{15} M_{\odot}$ at $z = 0$ respectively. Filled lines have been computed assuming f_g follows Eq. (18), dashed ones assuming $f_g = f_b$. We note how baryon decoupling produces a break in self-similarity. The change in slope observed at small $M_h/M_{h,o}$ values for the $3 \cdot 10^{13} M_{\odot}$ and $10^{14} M_{\odot}$ at $z = 0$ halos is associated to the minimum in \dot{M}_g , see Fig. 9. In the case of $3 \cdot 10^{14} M_{\odot}$ and $10^{15} M_{\odot}$ at $z = 0$ profiles, the minimum occurs at smaller $M_h/M_{h,o}$ values and falls outside the plotted range.

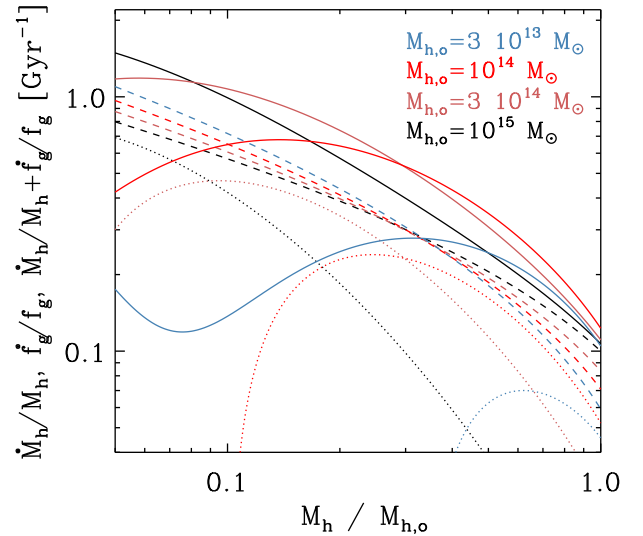


Fig. 11. \dot{M}_h/M_h (dashed lines), \dot{f}_g/f_g (dotted lines) and $\dot{f}_g/f_g + \dot{M}_h/M_h$ (filled lines) as a function of halo mass for four halos of $3 \cdot 10^{13} M_{\odot}$, $10^{14} M_{\odot}$, $3 \cdot 10^{14} M_{\odot}$, and $10^{15} M_{\odot}$ at $z = 0$, respectively.

ations. However, since \dot{M}_h/M_h and \dot{f}_g/f_g curves have opposite dependencies on halo mass, the $\dot{f}_g/f_g + \dot{M}_h/M_h$ (filled lines) show smaller variations on halo mass, with typical deviations from one another of less than $\sim 30\%$ for $M_h/M_{h,o} \gtrsim 0.3$, suggesting that self-similarity can be recovered, at least in part, by normalizing on the gas fraction.

A comparison of Fig. 12 with Fig. 10 shows that rescaling by gas fraction does reduce variability between the different halos in a substantial way. This implies that the break in self-similarity associated to the decoupling between gas and dark matter accretion can be in large part recovered by rescaling the entropy by the gas fraction. However, we note that while all profiles are broadly

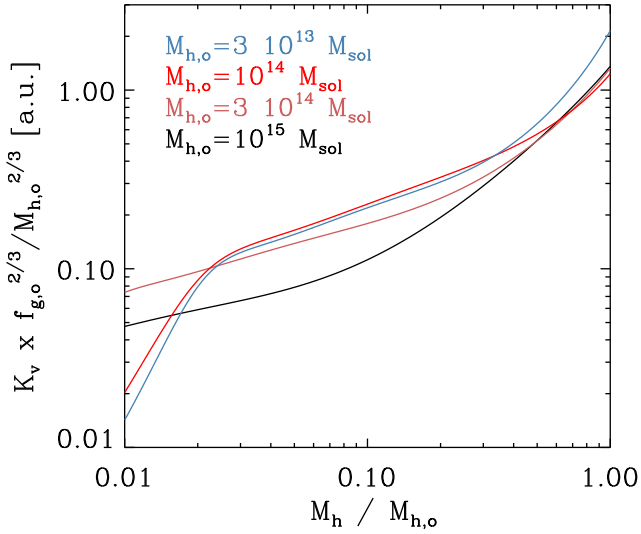


Fig. 12. Halo mass and gas fraction scaled entropy as a function of halo mass for four halos of $3 \cdot 10^{13} M_{\odot}$, $10^{14} M_{\odot}$, $3 \cdot 10^{14} M_{\odot}$, and $10^{15} M_{\odot}$ at $z = 0$, respectively.

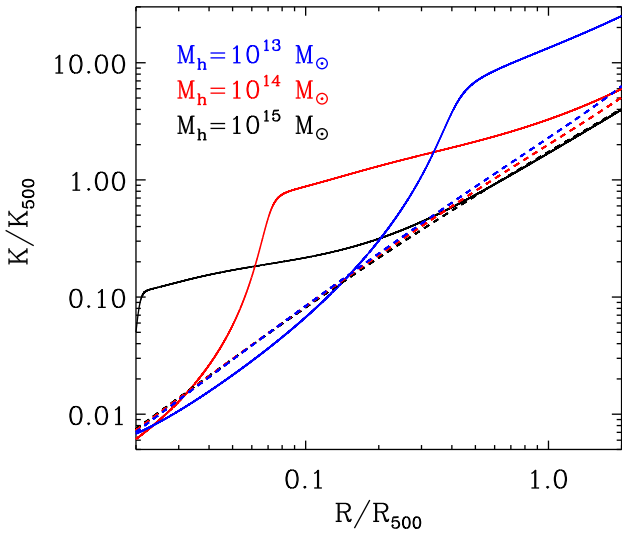


Fig. 13. Radial entropy profiles derived by imposing hydrostatic equilibrium for a halo mass of $10^{13} M_{\odot}$ (blue), $10^{14} M_{\odot}$ (red) and $10^{15} M_{\odot}$ (black) respectively. The radius has been normalized to R_{500} and the entropy to K_{500} . Filled lines have been computed with baryon decoupling, dashed ones without.

similar at large fractional masses, as we move to smaller fractional masses, the lower mass systems deviate before the higher mass ones.

5.5. Entropy profiles with baryon decoupling

Following the procedure described in Sect. 3.3 we derive radial entropy profiles with baryon decoupling. As expected, we observe substantial deviations with respect to the self-similar profiles, see Fig. 13. More specifically, baryon decoupling augments entropy for all halo masses but it does so in different radial ranges for different halo masses. In broad terms we may say that, for $M_h = 10^{13} M_{\odot}$ (blue line) the increase is concentrated at large radii, for $M_h = 10^{14} M_{\odot}$ (red line) at intermediate radii and at small radii for $M_h = 10^{15} M_{\odot}$ (black line). Intriguingly, the black solid line, which includes baryon decoupling, coincides with the

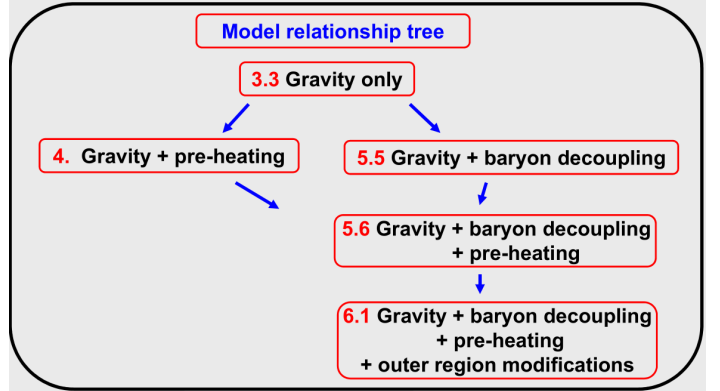


Fig. 14. Relationship tree connecting the models explored in this paper. The numbers in red at the left of the model name identify the section where the model is discussed.

dashed one which is derived assuming the baryon fraction of the accreting gas is constant, for radii larger than $\sim 0.4 R_{500}$. This is expected, as shown in Fig. 8, matter accreted on massive halos is close to having a cosmological baryon fraction. As pointed out in Sect. 5.3, baryon decoupling delays accretion, it does not stop it. An important consequence is that, in the outer regions of the most massive systems the entropy generation process is dominated by gravity, or in more explicit terms by the conversion of kinetic into thermal energy at the virial shock. We return to this important point in Sect. 6.1. The careful reader may note that also for the inner radii of low mass systems deviations from the gravitational heating model are modest. However, as discussed in Sect. 6.2, for $R < 0.1 R_{500}$ radiative cooling, which is not included in our model, significantly modifies entropy profiles.

To help readers keep track of the modifications we make to our model (see Sects. 5.6 and 6.1), in Fig. 14, we provide a model relationship tree.

5.6. Entropy profiles with decoupling and preheating

Decoupling and pre-heating are not mutually exclusive, quite the contrary, early-time feedback processes that remove gas from shallow potential wells may also heat it. In this subsection we modify entropy profiles derived in Sect. 5.5 to include the effects of preheating.

As already discussed in Sect. 4, to provide a realistic entropy model, it is critical to estimate K_* , the pre-heating contribution. To this end, we make use of what has been learned about feedback mechanisms over the past two decades. We proceed by estimating the temperature and the density of the preheated gas. Expulsion from low mass halos occurs when the temperature of the hotter phase of the gas exceeds the virial temperature. We set the temperature of the expelled gas to be: $T_{\text{exp}} = \Phi T_v(M_h = 0.3-1 \times 10^{13} M_{\odot})$, where Φ is some yet to be determined factor. Regarding density, we start by noting that the mass of gas that is accreted onto the halo is much larger than what is involved in the star formation or central black hole feeding process. For halos of $\sim 0.3-1.0 \times 10^{13} M_{\odot}$ roughly 1/7 of the baryon mass goes into stars (see Behroozi et al. 2019, and references therein). Assuming that roughly 1/3 of the gas that goes into stars is fed back (see discussion on the r_o factor in Ghizzardi et al. 2021, and references therein), and that the contribution from the AGN feedback is negligible, as suggested by the fact that the black hole mass is less than ~ 0.01 of the stellar mass (e.g., Reines & Volonteri 2015), we have that roughly

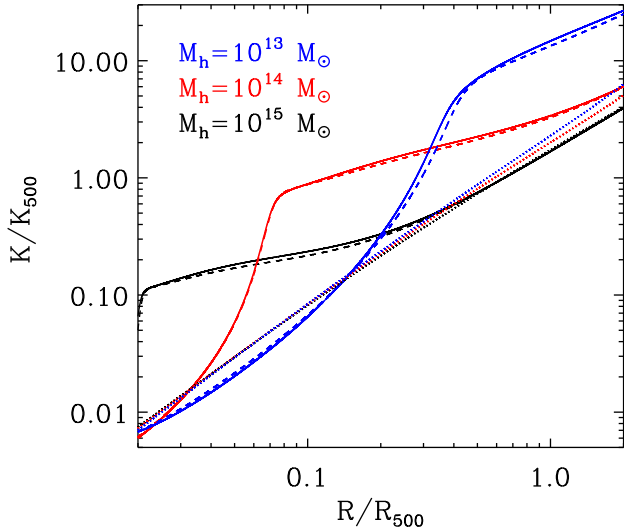


Fig. 15. Entropy profiles derived by imposing hydrostatic equilibrium for a halo mass of $10^{13} M_{\odot}$ (blue), $10^{14} M_{\odot}$ (red), and $10^{15} M_{\odot}$ (black), respectively. The radius has been normalized to R_{500} and the entropy to K_{500} . Filled lines have been computed with baryon decoupling and pre-heating, dashed lines with baryon decoupling and no preheating, dotted lines without either.

$1/7 \times 1/3 \sim 5\%$ of the baryons are involved in the feedback process. This implies that while the energy of the outflow will contribute significantly to that of the expelled gas, the mass contribution will be modest and for the present estimate negligible. Since the gas that requires the least energy to be expelled is the one that has just been accreted, we set the density of the expelled gas to that of the gas that has recently sustained a virial shock, $\rho_{\text{exp}} \sim \rho_v(M_h = 0.3-1 \times 10^{13} M_{\odot})$. Coupling this with $T_{\text{exp}} = \Phi T_v(M_h = 0.3-1 \times 10^{13} M_{\odot})$, which we derived earlier, we get $K_{\text{exp}} = \Phi K_v(M_h = 0.3-1 \times 10^{13} M_{\odot})$. Finally, we estimate that Φ cannot be much larger than a few, if it were $\geq 10^{2/3}$ then the thermal energy of the gas would have been sufficient to inhibit accretion on objects that are 10 times more massive (clusters), which would in turn require clusters to have gas fractions comparable to those of less massive systems (groups). On the basis of this discussion, we include a preheating entropy of $K_* = K_v(M_h = M_*) \sim 3.5 \times 10^{33} \text{ erg cm}^2 \text{ g}^{-5/3}$, where $M_* = 10^{13} M_{\odot}$. Since preheating only affects gas that is accreted on halos that are larger than the critical mass, M_* , we implemented the preheating correction on entropy by multiplying K_* by the factor $\exp(-M_*/M_h)$, which goes respectively to one for $M_h \gg M_*$ and to zero for $M_h \ll M_*$. This accounts for the fact that gas accreting on halos with mass comparable to M_* is likely a mixture of cold and preheated, with cold becoming dominant as we move downwards in halo mass and preheated if we move in the opposite direction.

In Fig. 15 we show entropy profiles for a halo mass of $10^{13} M_{\odot}$ (blue), $10^{14} M_{\odot}$ (red) and $10^{15} M_{\odot}$ (black) respectively with baryon decoupling and pre-heating (filled lines), with baryon decoupling and no preheating (dashed lines), without either (dotted lines). As already shown in Fig. 13, baryon decoupling introduces significant modification to the entropy profiles. Conversely, including pre-heating on top of baryon decoupling introduces negligible modifications. This is fairly easy to understand, for small halo masses, $M_h < 10^m M_{\odot}$, this is achieved by construction through the exponential cutoff term, $\exp(-M_*/M_h)$ described above. For large halo masses, the accreted gas is all

pre-heated, $\exp(-M_*/M_h) \sim 1$, but its contribution to entropy is negligible as in the preheating only model (see Fig. 6). We point out that changing the value of K_* by a factor of a few results in very minor modifications, this justifies, in retrospect, the order of magnitude estimate we have made of this quantity.

In summary, within massive halos, $M_h \sim 10^{15} M_{\odot}$, almost all the heat conferred to the gas is of gravitational nature. The role of feedback processes, while vital, requires relatively little energy, all that needs to be done is to take gas out of shallow potential wells and let gravity do the rest. The situation is radically different at the low mass end, $M_h \sim 10^{13} M_{\odot}$, here the energy of the expelled gas is sufficient to prevent, at least in part, its re-accretion.

6. Comparison with observations

This section is dedicated to a detailed comparison of the model we have developed with state of the art observations. We begin by focusing on the outer regions of massive clusters.

6.1. The outer regions of massive halos

On the base of the model described in the previous Section, it seems that the only place where entropy profiles depend only upon cosmological accretion is large radii in massive systems. Indeed, this is where both baryon decoupling and pre-heating play a negligible role, see Fig. 15. As shown in Figs. 6 and 7, the entropy predicted by the model is broadly consistent with the measured one, perhaps even a little on the high side: however, if one looks a little more in detail, one sees that the slopes are quite different, with observed profiles typically $d \ln K/d \ln R \sim 0.8$ (Pratt et al. 2006; Ghirardini et al. 2019) and our model $d \ln K/d \ln R \sim 1.2$.

As pointed out in Sect. 2, our modeling of the entropy generation at the accretion shock is relatively simple. Here we take a closer look at the process to gauge how a more realistic description might solve this issue. In very general terms there are two classes of corrections that may be applied. In the first, for some reason, the temperature and therefore the entropy of the shock heated gas does not increase with halo mass as $M_h^{2/3}$ but at a slightly milder rate. In the second, the kinetic to thermal conversion occurs over a radial range, rather than just at one radius. In this case the entropy, rather than jumping from zero to the value in Eq. (10) at a specific radius, grows over a yet to be determined radial range.

6.1.1. From smooth to mixed accretion

In this subsection we discuss a minor modification to the smooth accretion framework that can have an effect on the entropy profiles in the outer regions of massive halos. Part of the gas undergoing shock heating is unbound, part is bound to infalling sub-halos. For the former part, all the kinetic energy associated to the infall is transformed into thermal energy, for the latter a fraction is required to strip the gas from the sub-halo. Thus, the mean temperature of the shocked gas will be lower than that estimated in the model described in Sect. 2, because part of the kinetic energy has been used up to free gas from sub-halos. As can be seen in Fig. 7 of Eckert et al. (2021), the fraction of gas bound to the halo, goes up with increasing halo mass. This implies that the mean temperature of the shock heated gas is progressively reduced as we go up in mass, with respect to what is predicted on the basis of Eq. (3). Through some relatively straightforward

algebra we find that the ratio between the ‘‘mixed accretion’’ temperature, T_{ma} and the temperature computed from Eq. (3), can be expressed as

$$\frac{T_{\text{ma}}(M_{\text{h}})}{T(M_{\text{h}})} = 1 - \left(\frac{M_{\text{h}}}{M_{\text{h},\text{o}}}\right)^{-2/3} \frac{\int_{M_{\text{min}}/M_{\text{h},\text{o}}}^{M_{\text{h}}/M_{\text{h},\text{o}}} f_{\text{g}}(m_{\text{h}}) m_{\text{h}}^{5/3} dn/dm_{\text{h}} dm_{\text{h}}}{\int_{M_{\text{min}}/M_{\text{h},\text{o}}}^{M_{\text{h}}/M_{\text{h},\text{o}}} m_{\text{h}} dn/dm_{\text{h}} dm_{\text{h}}}, \quad (21)$$

where dn/dm_{h} is the halo mass function, m_{h} is a renormalized halo mass, $M_{\text{h},\text{o}}$ is the mass over which renormalization is carried out, M_{min} is the minimum mass over which integration is extended, for which we adopt a value $5 \times 10^{12} M_{\odot}$. We note that integration is fairly insensitive to the specific value of M_{min} as long as it is smaller than $\sim 10^{13} M_{\odot}$. For the gas fraction we adopt the expression presented in Eq. (18) with parameters used in Fig. 8. We note that the computation is insensitive to small variations on the parameters. A derivation of Eq. (21) is provided in Appendix C. Integration of Eq. (21) leads to variation of $T_{\text{ma}}(M_{\text{h}})/T(M_{\text{h}})$ of a few percent in the range $10^{13} M_{\odot}$ to $10^{15} M_{\odot}$. More specifically, when going from $10^{14} M_{\odot}$ to $10^{15} M_{\odot}$ we find a fractional temperature reduction of about 4%, see Fig. C.1, this is because more massive halos feature a larger fraction of bound gas. The net effect is a slightly lower growth of temperature and therefore entropy with increasing mass with respect to the case where evacuation of gas from sub-halos is not considered (see Sect. 5.6 and Fig. 14), leading in turn to a flattening of the entropy profile in the outer regions of massive clusters. The modification to the slope of the entropy profile, $\Delta(\ln K/\ln r) \sim 0.04$, is however very modest and insufficient to reach agreement with observations. It is worth pointing out that ours is not the only approach that can be adopted to estimate a correction for mixed accretion. Furthermore, given the very gradual increase in the fraction of bound (or unbound) gas with increasing halo mass, the correction will likely be a modest one, regardless of the approach used.

6.1.2. From shock radius to shock region

The model described in Sect. 2 is based on the assumption that all the kinetic energy generated by gas motions within the gravitation field of the forming halo is converted into thermal energy through one shock occurring at one specific radius. There are some good reasons to think this cannot be the case.

As discussed in Sect. 5, a sizable fraction of the gas falling onto a massive cluster has been previously accreted and ejected by smaller halos. It will be characterized by a distribution in entropy and other thermodynamic variables. Within the infall region, conversion of kinetic into thermal energy through adiabatic compression will operate more effectively in hotter subregions (see Eqs. 3 and 3.5 respectively in Tozzi & Norman 2001; Pringle & King 2007), this will lead to velocity gradient with cooler subregions falling in more rapidly than hotter ones. As soon as velocity differences become of the order of the sound speed, shocks (i.e. non adiabatic thermalization) will kick in and further amplify velocity gradients, likely leading to a runaway process. Furthermore, the process by which collisionless shocks convert kinetic into thermal energy is far from being well understood (see Goodrich et al. 2023, and references therein). For all of these reasons, the radial extent of the virialization region is not easy to estimate, so we took a heuristic approach. We substituted the step function operating on entropy in the model described in

Sect. 2 with a sigmoid that we implement through an hyperbolic tangent of the form

$$S(R) = \frac{1}{2} \tanh\left(-s \frac{R - hR_{\text{v}}}{R_{\text{v}}}\right) + \frac{1}{2}, \quad (22)$$

where R_{v} is the virial radius and s and h are two parameters that control the shape of the sigmoid. We note that $S(R)$ goes to 1 for $R \ll R_{\text{v}}$ and 0 for $R \gg R_{\text{v}}$. We also note that the modified entropy profile, $K'(R)$, is related to the non-modified one through the relation: $K'(R) = S(R) K(R)$.

We constrained the parameters characterizing the sigmoid through two measurements: (1) the entropy profiles in massive clusters, which, as pointed out above, have flatter slope than those predicted from our model; (2) the fraction of pressure in non-thermal form. The latter has been measured by several authors using different methods. Eckert et al. (2019), by comparing the hydrostatic mass with the total mass derived from the gas mass under the assumption that the gas fraction of massive clusters is close to the baryon fraction, find a non-thermal pressure support of $\sim 5\%$ at R_{500} and $\sim 10\%$ at R_{200} . Wicker et al. (2023), applying a similar approach, come up with an hydrostatic mass bias factor⁴ b of $\sim 20\%$. Dupourqué et al. (2024) analyzing surface brightness fluctuations on the CHEX-MATE sample, estimate a non-thermal pressure support of $9 \pm 6\%$ at R_{500} . Lovisari et al. (2024), applying a similar approach to temperature maps from the CHEX-MATE sample, estimate a non-thermal pressure support of $10^{+20}_{-10}\%$. Finally, Muñoz-Echeverría et al. (2024) derive an hydrostatic to lensing mass bias factor of $\sim 25 \pm 7\%$. Given the relatively large uncertainty in the non-thermal pressure contribution suggested by these measurements, we consider two scenarios: one assuming a larger contribution and the other a smaller one.

In Fig. 16 we plot the radial entropy profile for a halo of $10^{15} M_{\odot}$ at $z = 0$. The black filled line shows the profile in the case of a single shock at the virial radius, the blue filled line shows the case of a profile modified with a sigmoid to account for an extended shock region. Comparison with the dashed lines shows that the change introduced by the sigmoid flattens the slope to a value consistent with the one measured in observations.

The relationship between entropy modification and non-thermal pressure support depends on the details of the virialization process which, as pointed out above, are unknown. Assuming a simplified model in which the gas undergoes gradual heating within a given radial range accompanied by full instantaneous compression, $n'(R) = n(R)$, we have that the relationship connecting the modified temperature with the unmodified one is: $T'(R) = S(R) T(R)$. Similarly the modified pressure can be expressed as: $p'(R) = S(R) p(R)$ and the fraction of pressure support of non-thermal origin is given by $(p(R) - p'(R))/p(R) = 1 - S(R)$. Since $S(R_{500}) \sim 0.8$ and $S(R_{200}) \sim 0.7$, the modification shown in Fig. 16 implies a relatively high fractional non-thermal pressure support of 20% at R_{500} and 30% at R_{200} , comparable to measurements such as Muñoz-Echeverría et al. (2024) and Wicker et al. (2023). We note also that if we relax the condition of full instantaneous compression and write: $n'(R) = \mathcal{R}(R) n(R)$, with $0 < \mathcal{R}(R) < 1$, the relation between modified and unmodified pressure becomes: $p'(R) = S(R) \mathcal{R}(R)^{5/3} p(R)$

⁴ The factor b is defined through the relationship $b = 1 - M_{\text{HSE}}/M_{\text{T}}$, where M_{HSE} is the mass derived from the hydrostatic equilibrium equation and M_{T} is the total mass of the halo; to first approximation, the non-thermal pressure fraction, defined as the ratio of non-thermal pressure to total pressure is equal to the b parameter (Eckert et al. 2019).

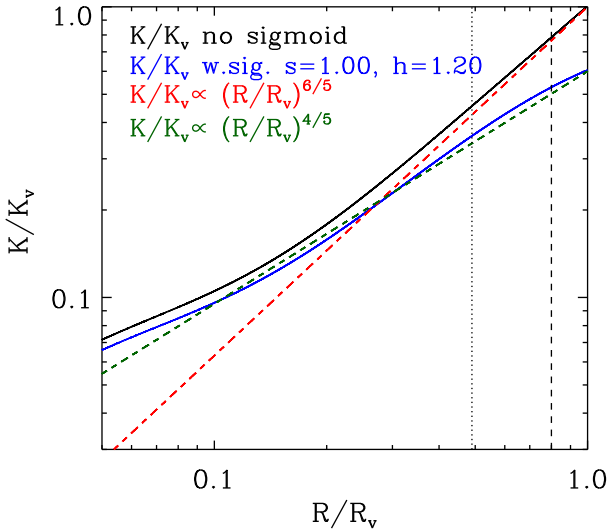


Fig. 16. Radial entropy profiles for a halo of $10^{15} M_{\odot}$ at $z = 0$. The black filled line shows the profile in the case of single shock at the virial radius. The blue filled line shows a profile modified with a sigmoid of the form reported in Eq. (22) and parameters shown in the figure. The vertical dotted and dashed lines mark the position of R_{500} and R_{200} respectively. The slope of the profiles can be gauged by comparison with the power laws plotted as dashed lines.

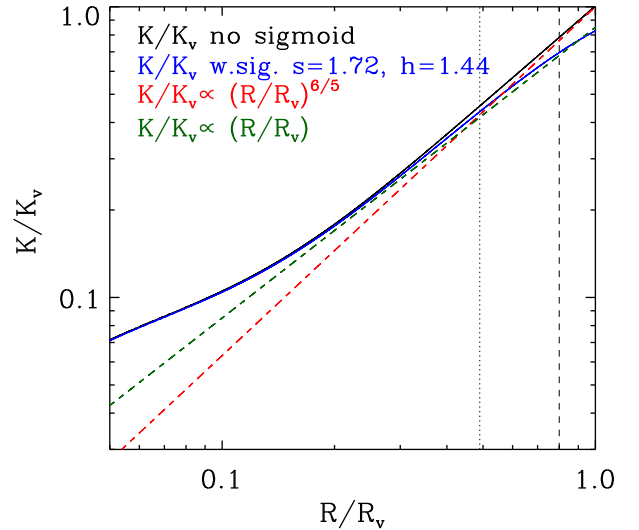


Fig. 17. Radial entropy profiles for a halo of $10^{15} M_{\odot}$ at $z = 0$. The black filled line shows the profile in the case of single shock at the virial radius. The blue filled line shows a profile modified with a sigmoid of the form reported in Eq. (22) and parameters shown in the figure. The vertical dotted and dashed lines mark the position of R_{500} and R_{200} respectively. The slope of the profiles can be gauged by comparison with the power laws plotted as dashed lines.

and the fraction of pressure support of non-thermal origin: $1 - S(R) \mathcal{R}(R)^{5/3}$, leading to an even larger non-thermal pressure support.

In Fig. 17 we show that by applying modifications that satisfy tighter constraints on the non-thermal pressure, roughly 5% at R_{500} and 10% at R_{200} , changes in the entropy profile are more modest, with the slope of the modified entropy profile in the outskirts reaching about 1.0.

We note that the flattening in the entropy profile could have a different physical origin from the one we have proposed above. In such a case our heuristic modeling would need to be reinterpreted within a different framework.

By combining modifications discussed in Sect. 6.1.1 with the ones that assume a larger contribution from non-thermal pressure, we come up with an entropy profile that is in good agreement with the observed one. Conversely, if we assume a smaller contribution from non-thermal pressure, we derive an entropy profile that is still steeper than the observed one but only slightly so. The difference is sufficiently small to allow us to question if the remaining deviations might be associated with issues on the observational side. We will return to this point in Sect. 6.2.

6.2. Comparison with clusters

We compare predictions based on the model presented in Sect. 5 with modifications described in Sects. 6.1.1 and 6.1.2 (see also Fig. 14) with 3 different cluster samples: HIGHMz (Riva et al. 2024), XCOP (Ghirardini et al. 2019) and REXCESS (Pratt et al. 2010) with median halo masses roughly scaling by a factor of 2, see Riva et al. (2024)⁵. Since the choice between larger and smaller non-thermal pressure support discussed in Sect. 6.1.2 has an impact only in the outermost regions we consider only one namely the latter.

⁵ All entropy profiles have been extracted from Riva et al. (2024), density profiles are extracted from the papers where they were first presented.

Before proceeding with the comparison, we corrected the observed profiles for a hydro-static scaling bias of 15%, $b = 0.15^6$. We do not limit our comparison to entropy, we also include another thermo-dynamic variable, we choose density, for reason that will become apparent later. Since the remaining two variables can be derived from the former, the comparison is informative of thermo-dynamic variables in general.

In Fig. 18 we show the comparison between entropy and density profiles predicted from our model and those measured from the HIGHMz sample (Riva et al. 2024), the density profiles are renormalized through the characteristic density $\rho_{g,500}$ defined by the relation: $\rho_{g,500} = 500 f_b \rho_c$, where f_b is the cosmic baryon fraction and ρ_c is the critical density of the Universe at redshift zero. The model was computed for a mass enclosed within an overdensity of 500 of $10^{15} M_{\odot}$. The agreement between observed data and model has been achieved by operating on the gas fraction parameters (see Eq. 18). This was not done through a formal fitting procedure but by adopting a trial and error approach. At the end of this procedure, two of the three parameters have been fixed to the following values: $l_m = 13.2$ and $l_{\sigma} = 0.75$. The shaded blue region is obtained by allowing the third parameter, l_n , to vary in the range $[-1.1, -0.3]$, with a central value of -0.7 . This is an attempt to explain the scatter within the observed sample as a variation in l_n . We note that shading stops at $R \sim 0.1 R_{500}$, for smaller radii the cooling time is shorter than the time since the gas was shock heated and entropy is reduced via radiation, something we do not account for in our model. For the radial region where the model can be meaningfully compared to observation, $0.1 R_{500} \leq R \leq R_{500}$, the agreement is quite good both for the entropy and the density (and therefore for all thermo-dynamic profiles).

It is worth pointing out that there is a certain degree of degeneracy between the 3 parameters describing the gas fraction versus halo mass relation. For example, a more gradual increase of

⁶ Moderate (± 0.1) variation on this quantity do not change our results in a substantial way.

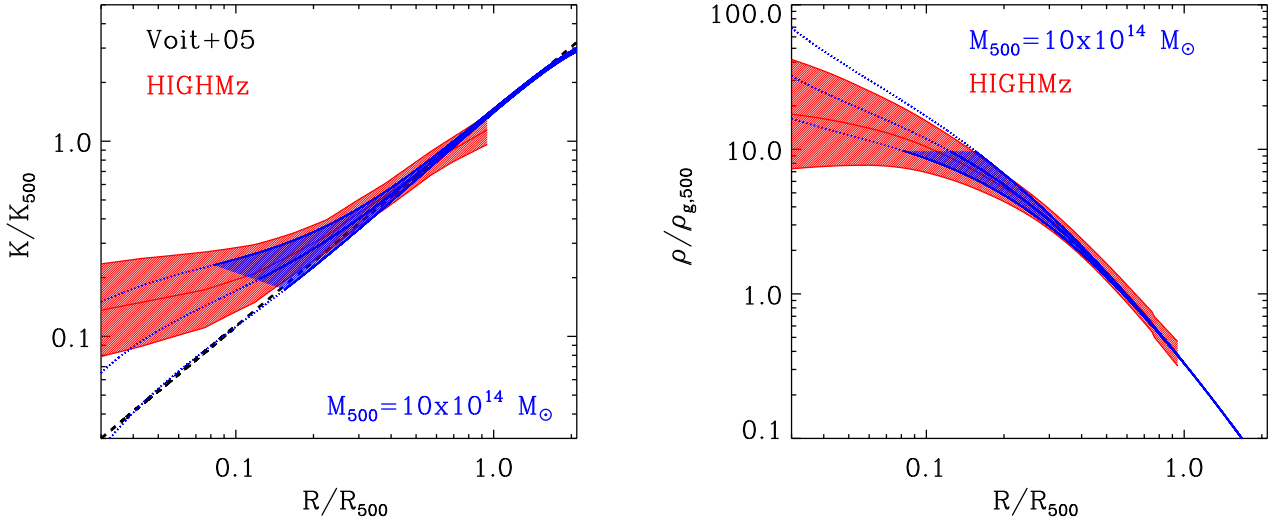


Fig. 18. *Left panel:* Entropy in units of K_{500} versus radius in units of R_{500} . The three blue lines are derived from the model presented in Sect. 5 with corrections described in Sects. 6.1.1 and 6.1.2. All 3 profiles are computed assuming $M_{500} = 10^{15} M_{\odot}$ and for the gas fraction parameters: $l_m = 13.2$ and $l_s = -0.75$. The l_n parameter is set to -1.1 , -0.7 and -0.3 respectively for the top, medium and bottom curves. The region between the top and bottom curves is shaded for radii where the cooling time is longer than the time since the gas was shock heated. The red shaded region is the median entropy and intrinsic scatter from the HIGHMz sample presented in Riva et al. (2024). The black dashed line shows the self-similar approximation introduced in Voit et al. (2005). *Right panel:* Density in units of $\rho_{g,500}$ versus radius in units of R_{500} . The three blue lines are derived from the model presented in Sect. 5 with corrections described in Sects. 6.1.1 and 6.1.2 using the same parameters for the gas fraction adopted for the entropy profile presented in the left panel. The region between the top and bottom curves is shaded for radii where the cooling time is longer than the time since the gas was shock heated. The red shaded region is the median and intrinsic scatter derived from the HIGHMz sample.

the gas fraction can be obtained by increasing the value of l_{σ} or by reducing that of l_m . As already alluded to, we have chosen to reproduce the observed scatter by varying just one parameter, l_n . This allows us to associate a physical interpretation with the intrinsic scatter measured in the observed sample. More specifically, objects that feature an entropy larger than average are interpreted as those for which deviation from the self-similar prediction is strongest, $l_n \sim -1.1$, conversely, systems that feature an entropy smaller than average, $l_n \sim -0.3$, are understood as having the smallest deviations from self-similarity.

Having derived a set of f_g parameters that reproduces reasonably well the observed profiles of the HIGHMz sample, we perform a comparison with the lower mass samples XCOP and REXCESS. We note that in performing these comparisons we change the halo mass, which we fix to a value close to that of the median of the sample, while the parameters describing the gas fraction are fixed to the values determined for the HIGHMz sample. In Fig. 19 we show a comparison of the model with density and entropy profiles for the three datasets. For the most part, we find good agreement; we will return in the next paragraphs on minor differences. However, the most important point is that the model reproduces two key features that are found in the data. First: for each entropy profile deviations from the self-similar prediction increase with decreasing radius. Second: when comparing entropy profiles, we see that, as we go down in halo mass, observations and model feature an increasing offset with respect to the self-similar prediction. Within the model that we propose, these two features are closely connected, indeed in both instances, the growing departure from self-similarity as we move down in mass or radius stems from the increasing decoupling of baryon accretion from dark matter accretion. It is this result, much more than the quality of the individual fits, that suggests to us that baryon decoupling captures a key feature of entropy generation.

The discrepancy between the modeled and observed entropy at large radii could be resolved by assuming a greater contri-

bution from non-thermal pressure support (see Sect. 6.1.2). This would suggest a substantial non-thermal pressure fraction, $\sim 20\%$ at R_{500} . However, we are cautious about adopting this interpretation because the deviation between the model and the data (see Fig. 19, bottom panel) is at the threshold of statistical significance and background related systematics may play a role (see Rossetti et al. 2024). Therefore, we find it more prudent to conclude that, within the scope of our model, current entropy measurements are consistent with a non-thermal pressure fraction within a range of 5% to 20% at R_{500} .

6.3. Comparison with groups

We move on to comparing our model with groups, $M_{500} \lesssim 10^{14} M_{\odot}$. Much less is known about these systems. While clusters thermodynamic profiles are reasonably well constrained, the same is not true for groups. More specifically there is no general consensus on the mean, median and scatter of thermodynamic profiles (see Popesso et al. 2024a, and references therein). One point recognized by all workers (see Eckert et al. 2021; Oppenheimer et al. 2021) is that group thermodynamic profiles feature a significantly larger scatter than those of clusters. We start off by making a comparison with the Sun et al. (2009) sample (see also Sun 2012), which was drawn from archival data and, although not X-ray selected in a strict sense, is X-ray bright.

In Fig. 20 we compare our model with observations from the Sun et al. (2009) sample, the only one for which density and entropy profiles are currently available. As for the comparison with the XCOP and REXCESS samples no tuning of the f_g parameters is made, they are fixed to the values determined by comparison with the HIGHMz sample. Since the sample comprises mostly objects in the $5.0 \times 10^{13} M_{\odot}$ to $10^{14} M_{\odot}$ range, we compare it with two different realizations of the model: one for $M_{500} = 5.0 \times 10^{13} M_{\odot}$ and the other for $M_{500} = 10^{14} M_{\odot}$. The model shows reasonable agreement with the data. As might be

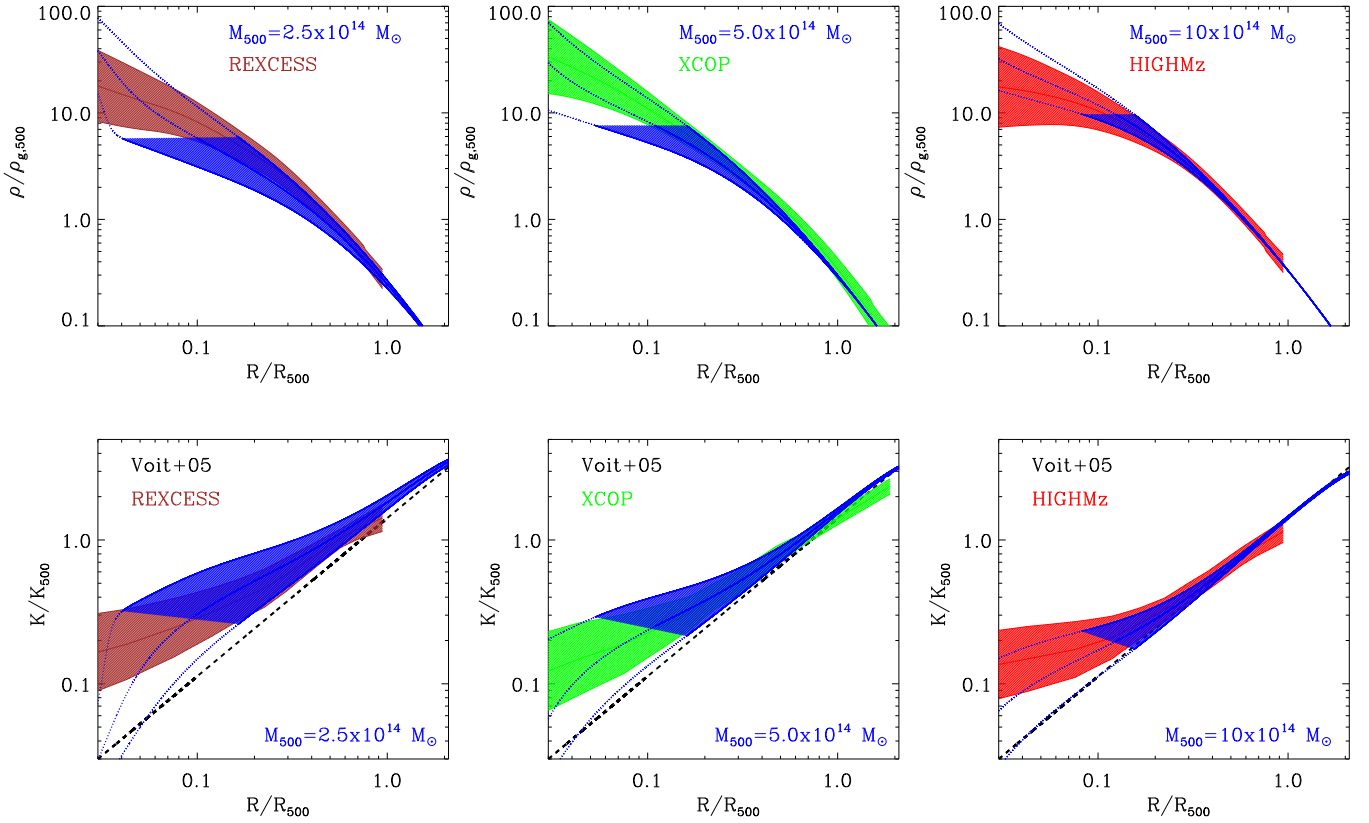


Fig. 19. *Top panels:* Gas density profiles from the model presented in Sect. 5 with corrections described in Sects. 6.1.1 and 6.1.2 for $M_{500} = 2.5 \times 10^{14} M_{\odot}$ (left), $M_{500} = 5.0 \times 10^{14} M_{\odot}$ (center), and $M_{500} = 10^{15} M_{\odot}$ (right) compared with observations from the REXCESS, XCOP, and HIGHMz samples respectively. For further details see the label of Fig. 18. *Bottom panels:* Same as top panels but for the entropy profiles.

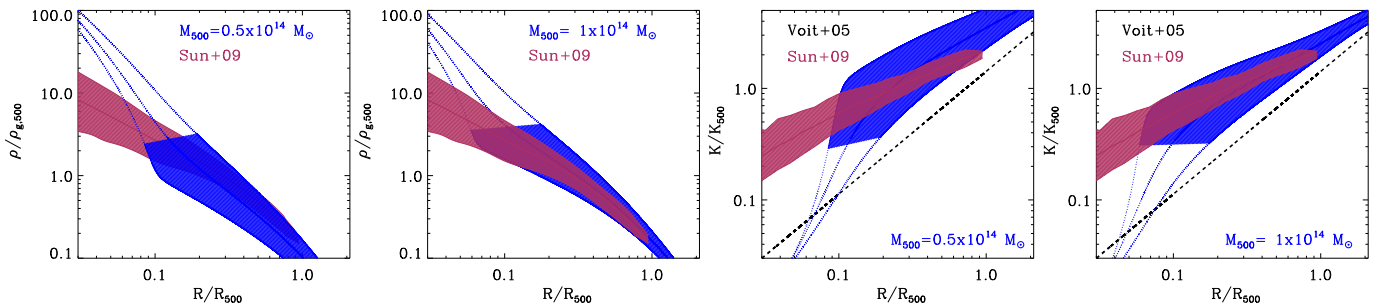


Fig. 20. *Left panels:* Density profiles predicted from our model for halo masses of $5.0 \times 10^{13} M_{\odot}$ (left) and $10^{14} M_{\odot}$ (right), shown as blue shaded regions, compared with observations from the Sun et al. (2009) sample, shown as a red shaded region (for further details, see the label of Fig. 18). *Right panels:* Same as left panels but for entropy profiles.

expected, since the median mass of the sample is $\sim 7 \times 10^{13} M_{\odot}$, the measured density profile is on the high side of the region predicted by the model for $M_{500} = 5.0 \times 10^{13} M_{\odot}$ and on the low side for a halo mass of $M_{500} = 10^{14} M_{\odot}$. A similar variation, albeit in the opposite direction, is observed for the entropy profiles. All in all, given that the parameters describing the model were tuned by comparison with data of systems that are 10 times more massive than those considered here, the agreement is remarkable. However, before drawing more general conclusions, we must point out that there are groups that feature very different radial profiles.

It has been known for some time that certain systems feature entropy profiles that display a rapid rise in the core followed by a quasi-iso-entropic range extending all the way to the outskirts.

Two such systems are reported in Baldi et al. (2009), a few more can be found in Humphrey et al. (2012). More recently, several more have been identified in the XGAP sample (Eckert et al. 2024).

In the left panel of Fig. 21, we can appreciate the difference between modeled and X-ray bright profiles on one side and those reported in Eckert et al. (2024). The latter are all characterized by a rapid rise followed by a flat (quasi-iso-entropic) region. The major difference between the four XGAP profiles is the radius at which the transition occurs and the value of the entropy in the flat region. For radii smaller than $0.1\text{--}0.2 R_{500}$, the entropy is in excess of both model predictions and observed data from the Sun et al. (2009) sample. The simplest possible explanation is that these systems have recently undergone significant heating

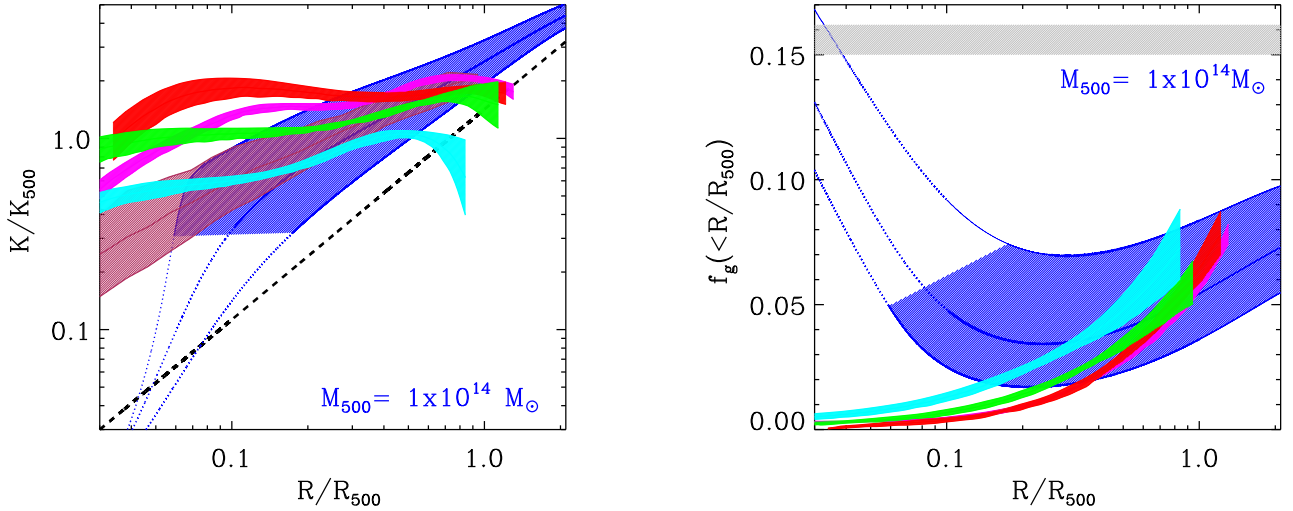


Fig. 21. *Left panel:* Radial entropy profiles. The blue and brown shaded regions respectively show the model for $M_{500} = 10^{14} M_{\odot}$ and the Sun et al. (2009) profiles, already reported in Fig. 20. The red, green, violet, and cyan shaded regions represent profiles reported in Eckert et al. (2024). *Right panel:* Same as the left panel but for the gas fraction profiles. The gray shaded region shows the cosmic baryon fraction.

events in their cores, the likeliest culprit being the central AGN. Given that our model does not include late-time AGN heating⁷, radiative cooling or self regulated Chaotic Cold Accretion (e.g., Gaspari et al. 2020), it is not surprising that it fails to reproduce the Eckert et al. (2024) profiles in the central regions. By looking at the right panel of Fig. 21 we see that for radii larger than $0.2-0.3 R_{500}$, the gas fraction of these systems is consistent with the one derived from our model, suggesting that the late-time AGN heating rearranges gas only within the inner regions of the halos (see Gaspari et al. 2012).

Although our model does not include a central heating source, it is relatively straightforward to estimate the amount of heat required to account for the entropy excess and gas defect characterizing the core of the Eckert et al. (2024) systems when compared to our model or the Sun et al. (2009) sample. To first approximation, the required heat has to be comparable to the thermal energy contained within $0.1-0.3 R_{500}$. A larger heat input would lead to an entropy excess extending to larger radii. The total thermal energy is the enthalpy integrated over the region of interest, and it can be expressed as

$$E_{\text{th}}(x) = \frac{15}{2} p_{500} V_{500} \int_0^x \rho(x) x^2 dx, \quad (23)$$

where p_{500} is the characteristic pressure (see Arnaud et al. 2010, for a definition), $V_{500} = 4/3\pi R_{500}^3$, $x = R/R_{500}$, and $\rho(x) = p(x)/p_{500}$. We also note that within the integral, we used the symbol x rather than x to avoid confusion with the upper integration limit. A derivation of the above formula is provided in Appendix D. By carrying out the calculation we find the thermal energy within $0.1-0.3 R_{500}$ to be somewhere between 10% and 30% of the thermal energy within R_{500} . Furthermore, since the energy required to go from a Voit (2005) profile to an XGAP one is of the order of the thermal energy content within R_{500} , we conclude that: if, prior to the heating event, the XGAP systems had thermodynamic profiles similar to Sun et al. (2009) ones, the required energy to make the transition is about 10% to 30%

⁷ As discussed in Sect. 5, our model assumes that early-time AGN feedback is responsible for expelling large amounts of gas in halos with masses around $10^{13} M_{\odot}$. It does not include a treatment for late-time AGN feedback at the group or cluster scales.

of what would be needed if the starting profiles featured a self-similar behavior down to $\sim 0.1 R_{500}$.

7. Comparison with theoretical work

Over the past two decades, there have been several attempts to reproduce observed entropy profiles, sometimes through models, more often via simulations.

7.1. Analytical modeling

Very recently Sullivan et al. (2024) presented an analytical model for cluster entropy incorporating non-thermal pressure contributions. They find, as we do (see our Fig. 15 and their Fig. 4, left panel) that, at large radii, the predicted entropy profile is steeper than the observed one. As we have done (see Sect. 6.1) they interpret this as evidence for non thermal pressure support. More specifically they show that the predicted entropy profile can be reconciled with the observed one by allowing roughly 20% of the energy content of the gas at R_{500} to be in a non-thermal form (see their Fig. 4, right panel). This is quite similar to what we achieve, through a somewhat more heuristic approach, see our Fig. 16.

7.2. Simulations

In Barnes et al. (2017), the authors produce thermodynamic profiles from a large suite of simulations. Their “hot” subsample, constituted by massive systems, features a median entropy profile which, beyond the core region, is in good agreement with the self-similar prediction (Voit et al. 2005), our own model for the $M_{500} = 10^{15} M_{\odot}$ case and the observed HIGHMz profile (Riva et al. 2024). Furthermore, the median entropy profile derived from their combined sample, with a median mass $M_{500} \sim 3 \times 10^{14} M_{\odot}$, is in agreement with our model for the $M_{500} = 2.5 \times 10^{14} M_{\odot}$ case and the observed REXCESS profile (see Fig. 22, left panel). According to the authors, the ability of their simulations to reproduce deviations from self-similarity observed in the data has to do with the inclusion of feedback from AGN. Unfortunately, from their paper it is not possible for us to determine how feedback actually achieves that objective.

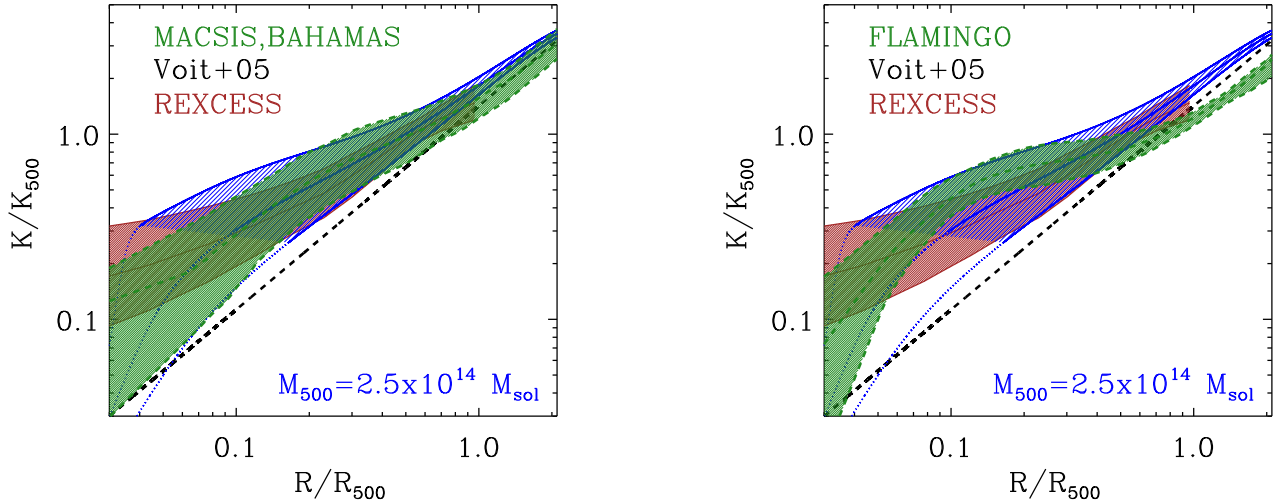


Fig. 22. *Left panel:* Radial entropy profiles for the combined sample extracted from the MACSIS and BAHAMAS simulations (Barnes et al. 2017) compared with the one derived from our model for $M_{500} = 2.5 \times 10^{14} M_{\odot}$ and with the observed REXCESS profile. *Right panel:* Radial entropy profiles from the FLAMINGO simulation for objects in the $M_{500} = 10^{14} - 10^{14.5} M_{\odot}$ mass bin compared with the one derived from our model for $M_{500} = 2.5 \times 10^{14} M_{\odot}$ and with the observed REXCESS profile.

Recently Braspenning et al. (2024) have published thermodynamic profiles from the FLAMINGO suite of simulations (Schaye et al. 2023). Their entropy profile for the $M_{500} = 10^{14} - 10^{14.5} M_{\odot}$ mass bin is compared to our prediction for $M_{500} = 2.5 \times 10^{14} M_{\odot}$ and the REXCESS observed profile in Fig. 22, right panel. As we can see, the FLAMINGO profile appears to be flatter than our own and, more importantly, of the observed profile. The tendency of simulated entropy profiles to be flatter than observed ones has been noted before (see Altamura et al. 2023). As discussed by these authors, there are several possible explanations. One possibility is that numerical viscosity and conduction, present in SPH simulations but absent from our model, favor mixing and heat transfer in the gas, which in turn lead to a reduction of the entropy gradient (see Fig. 6 of Altamura et al. 2023). As discussed in detail in Molendi et al. (2023) (see also Zhuravleva et al. 2019) there is significant evidence that both conduction and viscosity are heavily suppressed in the ICM.

7.3. Scaling relations

A different kind of connection can be made between our model and a phenomenological approach that has been pursued over the past two decades. Several authors have attempted to reduce the scatter in scaling relations and thermodynamic profiles that is left after applying self-similar scaling, by introducing modifications to the scaling laws (see Lovisari & Maughan 2022, for a review). Pratt et al. (2010) have shown that the scatter on the entropy profile of REXCESS clusters can be substantially reduced by multiplying it by the gas mass fraction. Recently, Ettori et al. (2023), have identified a correction scheme that modifies the self-similar predictions through two quantities: the gas mass fraction and the ratio between the global spectroscopic temperature and the temperature at R_{500} , with the former providing the most significant contribution. Interestingly, the authors derive a gas fraction versus halo mass relation that is broadly consistent with the one we derive by fitting the thermodynamic profiles of massive halos, see Sect. 6.2 and Fig. 23. It would therefore seem that a model based on baryon decoupling can provide a physical basis for the phenomenological findings reported in Pratt et al. (2010) and Ettori et al. (2023).

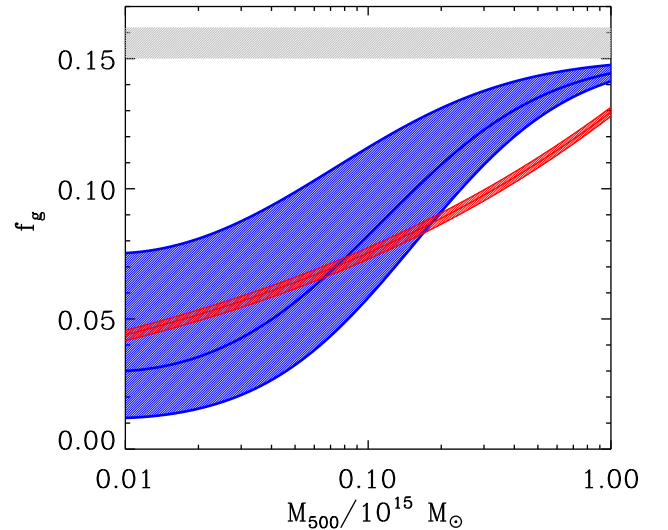


Fig. 23. Gas fraction as a function of halo mass. The solid blue lines show the profiles for the three choices of parameters adopted to fit cluster and group thermodynamic profiles. The blue shaded region spans the region between the different profiles. The red shaded region shows the gas fraction derived in Ettori et al. (2023) to minimize scatter in cluster scaling relations. The gray shaded region shows the cosmic baryon fraction.

Another connection can be made between our model and scaling relations. In Bower et al. (2008), the authors employ a semi-analytical model to investigate the luminosity temperature relation. To bring predictions into agreement with observations, they require the gas fraction in less massive halos to be smaller than in more massive ones. Intriguingly, they point out that the major problem for their model is the high level of heating required from the AGN, which follows from assuming that the heating occurs after the system has collapsed. As discussed in Sect. 6.3, baryon decoupling offers a simple and viable solution to this problem: the low gas fraction in less massive systems is achieved by delaying baryon infall rather than by ejecting the gas.

8. Discussion

We have shown that a simple model based on baryon decoupling is consistent with thermodynamic profiles at cluster scale and provides valuable insight on processes operating at the group scale. In this section we explore the implications of our results. Before doing so we take care of an important preliminary issue.

8.1. Evidence against a scale of massive gas ejection

Inspection of Fig. 23 reveals that the deviations from self-similarity in entropy at the cluster scale (see Figs. 18 and 19) stem from a phase of extremely low gas content centered around $10^{13} M_{\odot}$. According to our model, halos at this mass scale are missing between 50% and 90% of their baryons, corresponding to a gas fraction ranging between 2% and 8%. As discussed in Sect. 5.1, we interpret this severe depletion as the result of a phase of massive baryon evacuation. This finding is quite significant, and before proceeding, we must critically evaluate whether it aligns or conflicts with current observational data.

The gas content of halos around $10^{13} M_{\odot}$ remains largely unknown, primarily because the relevant measurements are extremely challenging to secure (see Tumlinson et al. 2017, and references therein). In a recent and brave attempt, Zhang et al. (2024b), stacking X-ray emission around $\sim 10^5$ SDD galaxies, found that the gas fraction within R_{500} of halos of mass $\sim 10^{13} M_{\odot}$ is in the order of 0.02–0.03, which is well within the range of our estimates. A detailed analysis of the systematics associated to the stacking procedure (Popesso et al. 2024b), while highlighting several criticalities inherent to the technique, confirm the Zhang et al. (2024b) result, albeit with significantly larger uncertainties: $f_g(R < R_{500}) \sim 0.00\text{--}0.07$ (Popesso et al. 2024c, see their Fig. 4). It is also worth pointing out that, in the few cases where measurements are available on individual systems, they are likely not representative of the general population (see Popesso et al. 2024a, and references therein). Furthermore, we do have indirect evidence that large amounts of enriched gas are expelled at these scales, (see Molendi et al. 2024). We conclude that, at least for the moment, the little observational evidence that is available does not challenge our inference of substantial gas mass ejection at the $10^{13} M_{\odot}$ scale.

8.2. Understanding entropy in massive halos

The success of our model in reproducing thermodynamic profiles of massive halos allows us make a few inferences about the underlying physics. The rough agreement between observed and modeled entropy and density in the outskirts of massive clusters (see Fig. 6), combined with evidence for moderate non-thermal pressure support in the $R_{500}\text{--}R_{200}$, see discussion at the end of Sect. 6.2, tells us that, to first order: (1) entropy is generated close to the virial radius, through an accretion shock; (2) the processes operating in such a shock, whatever they may be, must convert the bulk of the kinetic energy into thermal energy in a relatively narrow radial range.

The formation of an accretion shock at the virial radius can be understood in relatively simple terms. We know that within our Universe halos virialize at a specific overdensity (see Voit 2005, and references therein). In a collisionless fluid, such as dark matter, virialization is the process by which matter reorganizes itself by converting its velocity from radial to isotropic. This leads to a sharp increase in matter density around the virial radius. Since baryons are as affected by gravity as dark matter is, they also redistribute themselves in phase space like dark mat-

ter. However, unlike dark matter, baryons are collisional, which means that the interaction between low density baryons that are radially infalling and higher density baryons on isotropic orbits results in the formation of a shock front. It is for this reason that, as the halo mass grows by orders of magnitude, the shock tracks the virial radius as it moves outwards. Simulations (e.g., Vazza et al. 2009) suggest the gas undergoes multiple shocks spread out over a large radial range. However, from the point of view of entropy generation, the bulk is likely produced in just one shock, roughly located at the virial radius.

On the second point there is much less we can say. More than 60 years after their discovery, collisionless shocks in space are still very much the mystery they have always been (see Goodrich et al. 2023, and references therein). While we can infer that accretion shocks must be quite efficient in converting kinetic energy into thermal energy, we have little to say about how this is achieved.

The good agreement between observed thermodynamic profiles and predictions from our model, over a broad range of halo masses (see Figs. 18, 19 and 20), tells us that the modifications to the thermodynamics of massive halos, with respect to the self-similar model, can be explained through baryon decoupling. More specifically, for halo masses larger than $\sim 10^{13} M_{\odot}$, accreted matter contains a baryon fraction that is smaller than the cosmic one, this results in an increase in entropy with respect to the self-similar model. Since the baryon deficiency is larger for accretion on smaller halos: (1) for any given halo, the deviation in the radial profiles with respect to the self-similar prediction is larger at smaller radii and (2) the thermodynamic profiles of less massive halos feature a general offset, with respect to the self-similar model, that is larger than in more massive ones. It is this result, much more than the quality of the individual fits, that suggests to us that baryon decoupling captures a fundamental feature of entropy generation.

8.3. Non-thermal pressure support

As discussed in Sects. 6.1.2, 6.2 and 7.1, within our and similar models (e.g., Sullivan et al. 2024), there is a tight connection between non-thermal pressure support and the slope of the entropy profile in the outer regions of clusters, with steeper profiles requiring a smaller non-thermal pressure fraction. Current estimates place the non-thermal pressure fraction at R_{500} between 5% and 20%, while the entropy profile slopes in the $R_{500}\text{--}R_{200}$ range are estimated to be between 0.7 and 0.9. The limitations of current measurements for both parameters hinder our ability to use one to precisely constrain the other, leaving the radial extent of the shock-heated region uncertain.

Progress will require new and improved measurements of the entropy or the non-thermal pressure in cluster outskirts. Entropy profiles could be secured through a combination of high sensitivity surface brightness measurements in X-rays and Sunyaev-Zeldovich (SZ) observation in the sub-millimeter range. The former with missions such as AXIS (Reynolds et al. 2024), recently selected for a phase A study by NASA and through the WFI experiment on board the L-class mission NewAthena currently under study by ESA; the latter with future facilities such as AtLAST (Di Mascolo et al. 2024). The most effective way of estimating non-thermal pressure is through high resolution spectroscopic observations. The currently operational XRISM observatory (Tashiro et al. 2018) might provide some first results, much more will become possible with the XIFU experiment (Peille et al. 2025) on board NewAthena.

8.4. Baryon decoupling versus pre-heating

The baryon decoupling mechanism we propose exhibits substantial differences from the pre-heating models of the early 2000s, see Sect. 4. In our model, in massive halos, $M_h \sim 10^{14} - 10^{15} M_\odot$, almost all the heat conferred to the gas is of gravitational nature. Although feedback processes remain crucial, they require comparatively little energy; the key is to remove gas from shallow potential wells and allow gravity to do the rest. The situation is radically different at the low mass end, $M_h \sim 10^{13} M_\odot$, here the energy of the expelled gas is sufficient to prevent, at least in part, its re-accretion.

8.5. Gas fraction parametrization

The broad range of minimum baryon content suggested by our fits, $f_g/f_b = 0.1 - 0.5$ for $M_h \sim 10^{13} M_\odot$, see Fig. 23, likely follows from the stochastic nature of accretion onto the halo and of AGN feedback processes. As far as the mass range over which this minimum is reached, it cannot be estimated from our fits alone because they are essentially insensitive to accretion on halos less massive than about $M_h \sim 10^{13} M_\odot$. As discussed in Sect. 5.4, see also Figs. 9 and 10, the presence of a minimum in the gas fraction versus halo mass relation leads to a marked change in slope in the radial entropy profile. As it happens, for more massive halos (see Fig. 19), this variation occurs in the core of these systems where gas cooling becomes important and, as indicated by the comparison with data, washes it away. For less massive systems (see Fig. 19), although the variation occurs beyond the region where cooling becomes important, it is still within the reach of feedback from the SMBH (Gaspari et al. 2014) operating at late-times, that is, when the halo mass has grown well beyond $10^{13} M_\odot$.

The minimum on 10^m can be set at $\sim 5 \cdot 10^{12} M_\odot$ by recalling that the stellar mass fraction versus halo mass relation reaches a maximum around a few $10^{12} M_\odot$, for halos at $z \sim 2 - 3$, which are the progenitors of current groups and clusters, suggesting that up to this scale feedback effects do not prevent star formation by heating up the gas and therefore are unlikely to heat it up to virial temperatures. The maximum can be estimated at $\sim 3 \cdot 10^{13} M_\odot$ by noting that observed entropy profiles at the cluster mass scale show no evidence of the rapid change in slope discussed earlier in this subsection.

8.6. Scatter and bias in group samples

Our analysis indicates that reproducing the scatter in thermodynamic profiles at the cluster scale requires the gas fraction to exhibit an increasing scatter with decreasing halo mass, as illustrated in Fig. 23. We propose that this trend arises from a shift in accretion modes. At the “ejection” scale, around $10^{13} M_\odot$, outbursts from the central galaxy’s supermassive black hole (SMBH) expel gas from the halo. Due to the stochastic nature of these outbursts, different systems experience significantly varying levels of baryon loss. Over cosmic time, as the halo mass increases, the source of re-accreted baryons changes. Initially, the re-accreted material originates primarily from the halo’s own ejected baryons. However, as the halo mass continues to grow, re-accreted baryons increasingly come from a large number of smaller halos. Since individual halos expel varying amounts of gas, this process progressively reduces the scatter in the total amount of accreted gas, leading to the trend observed in Fig. 23.

Whatever the origin of the increase in scatter with decreasing halo mass, it implies that for $M_h \lesssim 5 \cdot 10^{13} M_\odot$, a substan-

tial portion of the population must have a very low baryon fraction. Consequently, any X-ray selected group sample is likely to be significantly biased. Notably, a similar conclusion has been reached by other studies using entirely independent methods (see Popesso et al. 2024a, and references therein). It is also important to note that this bias may be even more pronounced than our model predicts, as it does not account for late-time AGN heating effects, which are clearly observed (see Eckert et al. 2024).

8.7. Transitions

From an observational point of view, the “ejection” scale around $10^{13} M_\odot$ provides a physically meaningful divide between lower mass halos dominated by individual galaxies and higher mass multi galaxy systems. Below this threshold, stellar mass builds up rapidly from the accretion of cold gas, above it, the accreted gas is predominantly hot, leading to a sharp decline in the central galaxy’s star formation rate, often to the point of near cessation. As halo mass increases beyond the ejection scale, the stellar mass associated with sub-halos becomes more prominent, making the system more easily identifiable as a multi-galaxy system (see Fig. 9 of Coupon et al. 2015).

Baryons at the group and cluster scales clearly show significant differences. Groups feature a larger scatter in their properties, their baryon fraction is smaller and late-time heating from the central AGN contributes more significantly to their overall thermodynamic structure. That said, it is important to recognize that there is no clear separation between these two regimes. As far as we can tell, all these properties slowly vary with halo mass. In this sense the group/cluster scale separation appears to be less fundamental than the galaxy/group one, which is characterized by a change of the primary accretion mode from cold to hot.

8.8. Active galactic nucleus heating

In the framework we have sketched, the group and cluster scales ($M_h \gtrsim 10^{13} M_\odot$) can be viewed as the mass range over which halos reacquire expelled baryons. One may ask why the AGN feedback that is so effective in ejecting baryons in halos in the few $10^{12} M_\odot$ to $10^{13} M_\odot$ range seems to be incapable of doing so in more massive halos. The answer likely lies in the combination of two factors. One, as shown by the shut down of the star formation process (see Sect. 5 and references therein), the cold gas reservoir from which stars are formed and the black hole is fed, dries out at these scales never to be replenished. Two, as discussed in Gaspari et al. (2014), the maximum energy that can be extracted from the central AGN is comparable to the binding energy of poor groups. Thus, as halos grow in size, SMBH heating is restricted to progressively smaller self-similar radii. Although AGN feedback is incapable of baryon evacuation at the group and cluster scales, it does play a role in shaping the thermodynamics of these systems. At the group scale, where baryon decoupling is more substantial, it provides enough heat to reshape the thermodynamic profiles out to about a 1/5 of the virial radius (see Sect. 6.1.1). At the cluster scale its effects are contained within the core.

9. Summary

In this paper, we have presented a semi-analytical model that builds on work from the early 2000s. The model assumes that the bulk of the entropy of the hot gas in massive halos is generated near the virial shock. We investigated how differences

between entropy profiles predicted under gravitational collapse alone (known as self-similar profiles) and observed profiles can be reconciled by incorporating feedback processes. After demonstrating that adding pre-heating to the entropy generation process reduces, but does not fully resolve, the discrepancy between the model and observed data, we explored an alternative form of feedback. The core idea of our new approach is that by decoupling baryon accretion from dark matter accretion, we can reduce the density and increase the entropy of the accreted gas. We implemented this “baryon decoupling” concept through a parametric model for the gas fraction versus halo mass relation. Our main findings are summarized as follows:

- Our model provides a simple explanation for the effectiveness of gas fraction rescaling in reducing the scatter observed in entropy profiles of galaxy clusters.
- The difference in slope between predicted, $d \ln K / d \ln r \sim 1.2$, and observed, $d \ln K / d \ln r \sim 0.8$, entropy profiles of massive clusters at large radii can be reconciled, for the better part, through the adoption of a heuristic model and the assumption that entropy is generated over a radial range rather than just at the virial radius. Improved measurements of the entropy profile and the non-thermal pressure support in cluster outskirts are needed to further constrain the size of this region.
- Entropy and density profiles generated from a single set of parameters describing the gas fraction versus halo mass relation are in good agreement with observations from three different datasets that fully sample the cluster mass scale.
- Our model reproduces the growing departure from self-similarity observed in the data when moving inward in individual profiles and down in mass across different profiles. It is this result more than any other that suggests that baryon decoupling captures a key feature of entropy generation.
- In our model, deviations from self-similarity in entropy at the cluster scale stem from a phase of extremely low gas content centered around $10^{13} M_{\odot}$. We estimate that halos at this mass scale are missing between 50% and 90% of their baryons, corresponding to a gas fraction ranging between 2% and 8%.
- Our model reproduces entropy profiles of X-ray bright groups but not those of “quasi-isoentropic” groups. The latter clearly require late-time heating from the central AGN, which we do not include. Baryon decoupling does, however, significantly reduce the amount of AGN heating that is required.
- The increase of scatter in the gas fraction, when moving down in halo mass results from the reduction in the number of halos from which the re-accreted gas has previously been expelled.
- To reproduce the scatter in entropy profiles at the cluster scale, the gas fraction at the group scale must exhibit large variations. This implies that X-ray selected samples are likely to be significantly biased against gas-poor groups.
- The “ejection” scale around $10^{13} M_{\odot}$ that we have identified by modeling the gas fraction versus halo mass relation provides a physically meaningful divide between lower mass halos dominated by individual galaxies and higher mass multi-galaxy systems. Below this threshold, stellar mass builds up rapidly from the accretion of cold gas, and above it, the accreted gas is predominantly hot, and the central galaxy star formation rate is reduced to the point of almost being shut off.
- The AGN feedback that is so effective in ejecting baryons in halos in the few $10^{12} M_{\odot}$ to $10^{13} M_{\odot}$ range is incapable of doing so in more massive halos for two reasons: The cold

gas reservoir from which stars are formed and the black hole is fed dries out at these scales, and the maximum energy that can be extracted from the central AGN is comparable to the binding energy of poor groups.

What we have conducted in this paper might be described as a “thought experiment”. We set out to determine whether a single modification, the decoupling of baryon accretion from dark matter accretion, could account for a specific phenomenon, the observed deviation of entropy profiles from the self-similar prediction. Remarkably, baryon decoupling proves effective in explaining much of the behavior we sought to understand. While certain aspects inevitably require more complex modifications, the simplicity of our model offers valuable insights and a clearer understanding of the underlying processes.

Acknowledgements. We are grateful to the referee, Mark Voit, for helpful comments that improved the quality of our manuscript. S.M. thanks former Master student Giacomo Aprea for his contribution to this project. We acknowledge financial support from INAF mainstream project No.1.05.01.86.13. M.G. acknowledges support from the ERC Consolidator Grant *BlackHoleWeather* (101086804). L.L. acknowledges support from INAF grant 1.05.12.04.01. PT acknowledges support from the Next Generation European Union PRIN 2022 20225E4SY5 – “From ProtoClusters to Clusters in one Gyr”.

References

- Altamura, E., Kay, S. T., Bower, R. G., et al. 2023, *MNRAS*, **520**, 3164
- Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, *A&A*, **517**, A92
- Baldi, A., Forman, W., Jones, C., et al. 2009, *ApJ*, **694**, 479
- Balogh, M. L., Babul, A., & Patton, D. R. 1999, *MNRAS*, **307**, 463
- Barnes, D. J., Kay, S. T., Henson, M. A., et al. 2017, *MNRAS*, **465**, 213
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, **770**, 57
- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143
- Binney, J., & Tremaine, S. 1987, *Galactic Dynamics* (Princeton: Princeton University Press)
- Borgani, S., Governato, F., Wadsley, J., et al. 2002, *MNRAS*, **336**, 409
- Borgani, S., Finoguenov, A., Kay, S. T., et al. 2005, *MNRAS*, **361**, 233
- Bower, R. G., McCarthy, I. G., & Benson, A. J. 2008, *MNRAS*, **390**, 1399
- Braspenning, J., Schaye, J., Schaller, M., et al. 2024, *MNRAS*, **533**, 2656
- Coupon, J., Kilbinger, M., McCracken, H. J., et al. 2012, *A&A*, **542**, A5
- Coupon, J., Arnouts, S., van Waerbeke, L., et al. 2015, *MNRAS*, **449**, 1352
- Cowley, W. I., Caputi, K. I., Deshmukh, S., et al. 2018, *ApJ*, **853**, 69
- Di Mascolo, L., Perrott, Y., Mroczkowski, T., et al. 2024, *ArXiv e-prints* [arXiv:2403.00909]
- Donahue, M., & Voit, G. M. 2022, *Phys. Rep.*, **973**, 1
- Dupourqué, S., Clerc, N., Pointecouteau, E., et al. 2024, *A&A*, **687**, A58
- Eckert, D., Ghirardini, V., Ettori, S., et al. 2019, *A&A*, **621**, A40
- Eckert, D., Gaspari, M., Gastaldello, F., Le Brun, A. M. C., & O’Sullivan, E. 2021, *Universe*, **7**, 142
- Eckert, D., Gastaldello, F., O’Sullivan, E., et al. 2024, *Galaxies*, **12**, 24
- Ettori, S., Ghirardini, V., Eckert, D., et al. 2019, *A&A*, **621**, A39
- Ettori, S., Lovisari, L., & Eckert, D. 2023, *A&A*, **669**, A133
- Gaspari, M., Brighenti, F., & Temi, P. 2012, *MNRAS*, **424**, 190
- Gaspari, M., Brighenti, F., Temi, P., & Ettori, S. 2014, *ApJ*, **783**, L10
- Gaspari, M., Eckert, D., Ettori, S., et al. 2019, *ApJ*, **884**, 169
- Gaspari, M., Tombesi, F., & Cappi, M. 2020, *Nat. Astron.*, **4**, 10
- Ghirardini, V., Ettori, S., Eckert, D., & Molendi, S. 2019, *A&A*, **627**, A19
- Ghizzardi, S., Molendi, S., van der Burg, R., et al. 2021, *A&A*, **646**, A92
- Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, *A&A*, **634**, A135
- Goodrich, K., Schwartz, S., Wilson, L., et al. 2023, *BAAS*, **55**, 134
- Humphrey, P. J., Buote, D. A., Brighenti, F., et al. 2012, *ApJ*, **748**, 11
- Kaiser, N. 1986, *MNRAS*, **222**, 323
- Kaiser, N. 1991, *ApJ*, **383**, 104
- Leauthaud, A., Tinker, J., Bundy, K., et al. 2012, *ApJ*, **744**, 159
- Legrand, L., McCracken, H. J., Davidzon, I., et al. 2019, *MNRAS*, **486**, 5468
- Lovisari, L., & Maughan, B. J. 2022, in *Handbook of X-ray and Gamma-ray Astrophysics*, eds. C. Bambi, & A. Sanganello, 65
- Lovisari, L., Ettori, S., Rasia, E., et al. 2024, *A&A*, **682**, A45
- McCarthy, I. G., Babul, A., & Balogh, M. L. 2002, *ApJ*, **573**, 515
- McCarthy, I. G., Balogh, M. L., Babul, A., Poole, G. B., & Horner, D. J. 2004, *ApJ*, **613**, 811
- Molendi, S., De Grandi, S., Rossetti, M., et al. 2023, *A&A*, **670**, A104

- Molendi, S., Ghizzardi, S., De Grandi, S., et al. 2024, *A&A*, **685**, A88
- Moster, B. P., Naab, T., & White, S. D. M. 2013, *MNRAS*, **428**, 3121
- Muñoz-Echeverría, M., Macías-Pérez, J. F., Pratt, G. W., et al. 2024, *A&A*, **682**, A147
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, **490**, 493
- Nicastro, F., Fang, T., & Mathur, S. 2022, ArXiv e-prints [arXiv:2203.15666]
- Oppenheimer, B. D., Babul, A., Bahé, Y., Butsky, I. S., & McCarthy, I. G. 2021, *Universe*, **7**, 209
- Peille, P., Barret, D., Cucchetti, E., et al. 2025, ArXiv e-prints [arXiv:2502.10866]
- Ponman, T. J., Sanderson, A. J. R., & Finoguenov, A. 2003, *MNRAS*, **343**, 331
- Popesso, P., Biviano, A., Bulbul, E., et al. 2024a, *MNRAS*, **527**, 895
- Popesso, P., Marini, I., Dolag, K., et al. 2024b, *A&A*, submitted [arXiv:2411.16546]
- Popesso, P., Biviano, A., Marini, I., et al. 2024c, *A&A*, submitted [arXiv:2411.16555]
- Pratt, G. W., Arnaud, M., & Pointecouteau, E. 2006, *A&A*, **446**, 429
- Pratt, G. W., Arnaud, M., Piffaretti, R., et al. 2010, *A&A*, **511**, A85
- Pringle, J. E., & King, A. 2007, *Astrophysical Flows* (Cambridge: Cambridge University Press)
- Rasia, E., Tormen, G., & Moscardini, L. 2004, *MNRAS*, **351**, 237
- Reines, A. E., & Volonteri, M. 2015, *ApJ*, **813**, 82
- Reynolds, C. S., Miller, E. D., Hodges-Kluck, E., et al. 2024, *SPIE Conf. Ser.*, **13093**, 1309328
- Riva, G., Pratt, G. W., Rossetti, M., et al. 2024, *A&A*, **691**, A340
- Rossetti, M., Eckert, D., Gastaldello, F., et al. 2024, *A&A*, **686**, A68
- Salcido, J., Bower, R. G., & Theuns, T. 2020, *MNRAS*, **491**, 5083
- Schaye, J., Kugel, R., Schaller, M., et al. 2023, *MNRAS*, **526**, 4978
- Shuntov, M., McCracken, H. J., Gavazzi, R., et al. 2022, *A&A*, **664**, A61
- Sullivan, A., Shabala, S., Power, C., Bottrell, C., & Robotham, A. 2024, *PASA*, **41**, e064
- Sun, M. 2012, *New J. Phys.*, **14**, 045004
- Sun, M., Voit, G. M., Donahue, M., et al. 2009, *ApJ*, **693**, 1142
- Tashiro, M., Maejima, H., Toda, K., et al. 2018, *Proc. SPIE*, **10699**, 1069922
- Tozzi, P., & Norman, C. 2001, *ApJ*, **546**, 63
- Tumlinson, J., Peebles, M. S., & Werk, J. K. 2017, *ARA&A*, **55**, 389
- Vazza, F., Brunetti, G., Kritsuk, A., et al. 2009, *A&A*, **504**, 33
- Voit, G. M. 2005, *Rev. Mod. Phys.*, **77**, 207
- Voit, G. M., Balogh, M. L., Bower, R. G., Lacey, C. G., & Bryan, G. L. 2003, *ApJ*, **593**, 272
- Voit, G. M., Kay, S. T., & Bryan, G. L. 2005, *MNRAS*, **364**, 909
- Wicker, R., Douspis, M., Salvati, L., & Aghanim, N. 2023, *A&A*, **674**, A48
- Zhang, H., Behroozi, P., Volonteri, M., et al. 2024a, ArXiv e-prints [arXiv:2409.16347]
- Zhang, Y., Comparat, J., Ponti, G., et al. 2024b, *A&A*, **690**, A267
- Zhuravleva, I., Churazov, E., Schekochihin, A. A., et al. 2019, *Nat. Astron.*, **3**, 82

Appendix A: Deriving $d \ln K_v / d \ln M_h$

In this appendix, we derive the relation presented in Eq. 11. We start off with the expression for K_v in Eq. 10. As a first step, we compute the natural logarithm of K_v :

$$\ln K_v = \ln \left[\frac{1}{3} (\pi G^2)^{2/3} f_b^{-2/3} \right] + \ln \left(\frac{M_h}{\dot{M}_h} \right)^{2/3} + \ln M_h^{2/3}. \quad (\text{A.1})$$

Next, by noting that M_h/\dot{M}_h can be rewritten as $t \, d \ln t / d \ln M_h$ and defining α as $\alpha \equiv d \ln M_h / d \ln t$, we get

$$\ln K_v = \ln \left[\frac{1}{3} (\pi G^2)^{2/3} f_b^{-2/3} \right] + \frac{2}{3} (\ln t - \ln \alpha + \ln M_h). \quad (\text{A.2})$$

Next, we compute the derivative of $\ln K_v$ with respect to $\ln M_h$:

$$\frac{d \ln K_v}{d \ln M_h} = \frac{2}{3} \left(\frac{d \ln t}{d \ln M_h} - \frac{d \ln \alpha}{d \ln M_h} + \frac{d \ln M_h}{d \ln M_h} \right), \quad (\text{A.3})$$

which simplifies to

$$\frac{d \ln K_v}{d \ln M_h} = \frac{2}{3} \left(1 + \frac{1}{\alpha} - \frac{d \ln \alpha}{d \ln M_h} \right). \quad (\text{A.4})$$

Appendix B: Applying hydrostatic equilibrium

In this appendix we present the numerical method we have adopted to derive thermodynamic profiles under the condition of hydrostatic equilibrium.

B.1. Thermodynamic variables at the virial radius

We start by deriving expressions for thermodynamic variables at the shock radius, which is assumed to coincide with the virial radius, R_v . Assuming a strong shock, the following relation holds:

$$\rho_{g,d} = 4\rho_{g,u}, \quad (\text{B.1})$$

where $\rho_{g,u}$ and $\rho_{g,d}$ are respectively the gas density upstream and downstream of the shock. By combining Eq. B.1 with equations describing the mass flow (Eq. 1) and the velocity of the accreted gas (Eq. 2), we derive the following expression:

$$\rho_{g,d} = \frac{1}{\pi G^{1/2}} R_v^{-3/2} M_h^{-1/2} \dot{M}_g, \quad (\text{B.2})$$

where M_h is the halo mass and \dot{M}_g the gas accretion rate. Combining energy conservation (Eq. 3) with the equation describing the velocity of the accreted gas (Eq. 2) we derive the following expression for the gas temperature at the virial radius:

$$k_b T_v = \frac{1}{3} \mu m_p \frac{G M_h}{R_v}, \quad (\text{B.3})$$

where k_b is the Boltzmann constant, μ the mean molecular weight of the gas and m_p the proton mass. Assuming our gas obeys the equation of state for a monoatomic gas,

$$P_g = \frac{\rho_g}{\mu m_p} k_b T, \quad (\text{B.4})$$

and making use of Eqs. B.2 and B.3, we derive an expression for the gas pressure at the virial radius:

$$P_{g,v} = \frac{G^{1/2}}{3\pi} R_v^{-5/2} M_h^{1/2} \dot{M}_g. \quad (\text{B.5})$$

Similarly, starting from the definition of entropy:

$$K \equiv \frac{k_b T}{\mu m_p \rho^{2/3}}, \quad (\text{B.6})$$

and making use of Eqs. B.2 and B.3 we get the following expression for the entropy at the virial radius:

$$K_v = \frac{1}{3} (\pi G^2)^{2/3} \frac{M_h^{4/3}}{\dot{M}_g^{2/3}}. \quad (\text{B.7})$$

B.2. Hydrostatic equilibrium equation

The next step is to take the commonly adopted form of the equation of hydrostatic equilibrium equation,

$$\frac{1}{\rho_g} \frac{d P_g}{d R} = - \frac{G M_h(< R)}{R^2}, \quad (\text{B.8})$$

and, in light of the power-law like nature of thermodynamic profiles, rewrite it as

$$\frac{d \ln P_g}{d \ln R} = - \frac{\mu m_p}{k_b T} \frac{G M_h(< R)}{R}. \quad (\text{B.9})$$

Next, following (Tozzi & Norman 2001), we rewrite Eq. B.9 by dividing each variable by its value at the virial radius:

$$\frac{d \ln p}{d \ln x} = -C \tau^{-1} \frac{m(< x)}{x}. \quad (\text{B.10})$$

where

$$C = \frac{G M_h}{R_v} \frac{\mu m_p}{k_b T_v}, \quad \tau = T/T_v, \quad p = P_g/P_{g,v}, \quad m = M_h(R)/M_h(R_v), \quad x = R/R_v. \quad (\text{B.11})$$

By plugging Eq. B.3 into Eq. B.11, we get

$$C = 3. \quad (\text{B.12})$$

An advantage of Eq. B.10 is that it allows us to assign at a glance a physical interpretation to the numerical term, $C = 3$. Since, for $x = 1$, all terms on the right side of the equation are 1, -3 is the slope of the $\ln p$ versus $\ln x$ relation at the virial radius.

Equations B.1, B.3, B.4, B.6 completely describe physical conditions at the virial radius. By imposing hydrostatic equilibrium, we relate variations in thermodynamic quantities with the quantities themselves and, since these quantities are known at the virial radius, so are the variations which relate them. Furthermore, by rewriting the hydrostatic equilibrium equation in terms of $d \ln P_g / d \ln R$ the two terms on the right hand side of Eq. B.9 become proportional to one another when we evaluate Eq. B.9 at the virial radius (see Eq. B.3). This is what allows us to establish that $d \ln P_g / d \ln R$ computed at R_v is a constant.

Finally, following (Tozzi & Norman 2001), we rewrite the normalized temperature in Eq. B.10 in terms of normalized pressure and entropy:

$$\frac{d \ln p}{d \ln x} = -3 p^{-2/5} k^{-3/5} \frac{m(< x)}{x}, \quad (\text{B.13})$$

where $k = K/K_v$.

B.3. Gas mass equation

The next step is to take the differential relation between gas mass within a given radius and the gas density at that radius,

$$\frac{dM_g}{dR} = 4\pi\rho_g R^2, \quad (\text{B.14})$$

and subject it to the same sort of massaging we applied to the hydrostatic equilibrium equation. We start by rewriting Eq. B.14 through variables that are one at the virial radius

$$\frac{dm_g}{dx} = C_g \chi x^2, \quad (\text{B.15})$$

where

$$C_g = \frac{4\pi R_v^3 \rho_{g,v}}{M_{g,v}}, \quad m_g = M_g(R)/M_{g,v}(R_v) \text{ and} \quad \chi = \rho_g/\rho_{g,v}. \quad (\text{B.16})$$

We rewrite C_g in a more convenient form:

$$C_g = 3 \frac{\rho_{g,v}}{\langle \rho_{g,v} \rangle} \quad \text{where} \quad \langle \rho_{g,v} \rangle = \frac{M_{g,v}}{4/3\pi R_v^3}. \quad (\text{B.17})$$

Finally, as we have done for the pressure gradient, we rewrite Eq. B.15 in terms of natural logarithms and express density in terms of pressure and entropy:

$$\frac{d \ln m_g}{d \ln x} = C_g \left(\frac{p}{k} \right)^{3/5} \frac{x^3}{m_g}. \quad (\text{B.18})$$

As for the hydrostatic equilibrium equation, see Eq. B.10, this form allows us to understand that C_g is the slope of the $\ln M_g$ versus $\ln R$ relation at the virial radius:

$$C_g = \left(\frac{d \ln M_g}{d \ln R} \right)_{R=R_v}. \quad (\text{B.19})$$

B.4. Entropy equation

In our model, entropy generation occurs only at the virial (shock) radius, this implies that any further evolution of the gas will be adiabatic. The gas may compress or expand to obey the hydrostatic equation, but it will not change its entropy. This implies that the entropy of a gas element within the halo can be expressed as a function of the halo mass and gas accretion rate at the time the element was shock heated, see Eq. 8. In differential terms, we may say that the infinitesimal increase in entropy is determined by the infinitesimal increase in gas mass. Thus, the variable $d \ln K/d \ln M_g$ will not depend on the radius within the halo where the infinitesimal gas element eventually resides but only on the halo mass at the time the gas was accreted. Furthermore, once a description of how the gas mass, M_g , varies with the halo mass is provided, $M_g = M_g(M_h)$ (see Eq. 9 for the self-similar case and Eq. 19 for the baryon decoupling case), $d \ln K/d \ln M_g$ is completely specified and can be readily computed from M_h , which in mathematical terms is

$$\frac{d \ln K}{d \ln M_g} = f(M_h). \quad (\text{B.20})$$

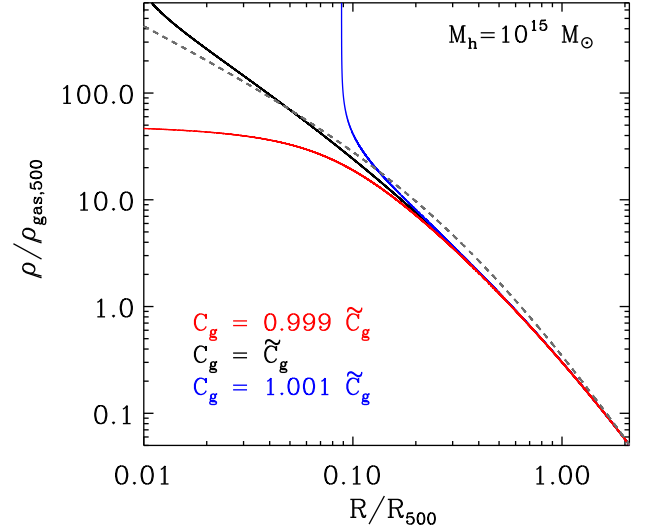


Fig. B.1. Gas density profiles derived by solving for hydrostatic equilibrium for 3 different values of C_g as detailed in the label. The dashed gray line shows the distribution of the gas as it is deposited in the gravitational well, without any redistribution.

B.5. Numerical solution

The differential equations, Eqs. B.13, B.18 and B.20, together with the boundary conditions provided in Eqs. B.2, B.5 and B.7, make up a system that can be solved numerically. To this end, we make use of the finite differences method and rewrite Eqs. B.13, B.18 and B.20 as

$$\Delta \ln p = -3 p^{-2/5} k^{-3/5} \frac{m(<x)}{x} \Delta \ln x, \quad (\text{B.21})$$

$$\Delta \ln m_g = C_g \left(\frac{p}{k} \right)^{3/5} \frac{x^3}{m_g} \Delta \ln x, \quad (\text{B.22})$$

$$\Delta \ln k = \frac{d \ln k}{d \ln m_g} (m(<x)) \Delta \ln m_g. \quad (\text{B.23})$$

As previously pointed out, we start from the virial radius where all normalized variables are 1 and C_g (see Eq. B.17) and $d \ln k/d \ln m_g$ (see discussion in Sect. B.4 and Eq. B.20) are known. We move inwards by selecting a value for $\Delta \ln x$ that is sufficiently small to adequately sample gradients and, through Eqs. B.21 and B.22, compute respectively values for $\ln p$ and $\ln m_g$ at $\ln x = 1 + \Delta \ln x$. From $\Delta \ln m_g$, through Eq. B.23, we compute $\ln k$ at $\ln x = 1 + \Delta \ln x$. Next we insert the values of p , $m(<x)$, m_g and k computed at $\ln x = 1 + \Delta \ln x$ in Eqs. B.21, B.22 and B.23 and proceed iteratively.

As previously pointed out, the system is solved going inwards. Since gas at smaller radii is shock heated before gas at larger radii, in a sense, we solve the system by going back in time. In other words, our scheme allows us to derive numerical solutions, which are not necessarily physical ones. This point is better understood by looking at Fig. B.1, where we show 3 different density profiles derived by solving for 3 distinct values of C_g that differ from one another by $\sim 0.1\%$. The gas distribution derived from the $C_g = \tilde{C}_g$ case (black solid line) differs by less than 10% over most of the radial range with respect to the distribution obtained by assuming the gas does not move from the position it is deposited in from the accretion process (gray dashed line). From Eq. B.22 we see that, for $C_g > \tilde{C}_g$ gas density will increase more rapidly as we go inwards with respect

to the $C_g = \widetilde{C}_g$. This effectively leads to gas being relocated to slightly larger radii, such a process cannot continue indefinitely because, sooner or later, we run out of gas. In the case reported here, where the variation in C_g is modest, the run away occurs at relatively small radii, for larger values of C_g it is located at larger radii. Clearly this is an unphysical solution, and we do not consider it further. The $C_g < \widetilde{C}_g$ is more interesting, here Eq. B.22 leads to a redistribution of the gas from larger to smaller radii. This is a mathematically viable solution, however from a physical point of view things are different.

The broad agreement observed at large radii between entropy predicted under the assumption of full conversion of kinetic into thermal energy and the observed one (see Fig. 6) tells us that the redistribution, that follows the initial deposition of the gas and which is required to reach hydrostatic equilibrium, must be quite modest. This is consistent with findings based on our analysis, in Fig. B.1 we see that, as we move away from the virial/shock radius, $R_v \sim 2.1R_{500}$, the density derived by imposing hydrostatic equilibrium remains close to the one derived by simply requiring that the gas remain where it is deposited by the accretion process over a large radial range. In summary we may say that, as the halo grows in size, a redistribution process sweeps outwards in the wake of the expanding accretion shock, leaving behind it a gas distribution which is in hydrostatic equilibrium and close to the one derived by assuming the gas remains where it is deposited by the accretion process. This line of reasoning strongly suggests that the favored solution is the one with $C_g = \widetilde{C}_g$.

We implement this solution in our numerical code through a recursive process that solves Eqs. B.21, B.22 and B.23 starting from the value of C_g derived from Eq. B.19. We then compare the density at $0.03 R_{500}$ with the one derived assuming gas does not move from where it is deposited by the accretion process and modify C_g to reduce the difference. We iterate until the difference in density at $0.03 R_{500}$ is smaller than 10%. The search for \widetilde{C}_g is carried out through a standard bisection method.

Appendix C: A derivation of the temperature correction

In this appendix we derive Eq. 21 presented in Sect. 6.1.1. We start by writing the gas mass falling onto the halo in a time interval Δt , $\dot{M}_g \Delta t$, as the sum of the masses of the infalling sub-halos:

$$\dot{M}_g(M_h^*) \Delta t = \int_{M_h^{\min}}^{M_h^*} M_g(M_h) \frac{dN}{dM_h} dM_h, \quad (\text{C.1})$$

where dN/dM_h is the halo mass function, M_h^{\min} is the minimum mass over which the sum is extended and M_h^* is the mass of the halo on which the gas is accreting. We note that this description of the process holds as long the timescale over which the halo mass function varies is much longer than Δt , the timescale over which thermalization occurs. The gravitational energy converted into thermal energy through shock heating in the same time interval may be written as

$$E_{\text{th}}(M_h^*) = \frac{\dot{M}_g(M_h^*)}{\mu m_p} k_b T(M_h^*) \Delta t. \quad (\text{C.2})$$

By substituting Eq. C.1 into Eq. C.2, we get

$$E_{\text{th}}(M_h^*) = \frac{k_b T(M_h^*)}{\mu m_p} \int_{M_h^{\min}}^{M_h^*} M_g(M_h) \frac{dN}{dM_h} dM_h. \quad (\text{C.3})$$

The binding energy of the gas in the potential wells of the sub-halos falling onto the main halo in the time interval Δt , can be expressed as

$$\widetilde{E}_b(M_h^*) = \int_{M_h^{\min}}^{M_h^*} E_b(M_h) \frac{dN}{dM_h} dM_h, \quad (\text{C.4})$$

where $E_b(M_h)$ is the binding energy of the gas within sub-halos of mass M_h . We recall that accreting gas with no energy at infinity, once settled into hydrostatic equilibrium, has a binding energy that equals its thermal energy (Binney & Tremaine 1987).

Thus the binding energy of the gas in a sub-halo of mass M_h can be expressed as the product of the number of gas particles multiplied by the temperature:

$$E_b(M_h) = \frac{M_g(M_h)}{\mu m_p} k_b T(M_h). \quad (\text{C.5})$$

Plugging Eq. C.5 into Eq. C.4, we get

$$\widetilde{E}_b(M_h^*) = \int_{M_h^{\min}}^{M_h^*} \frac{M_g(M_h)}{\mu m_p} k_b T(M_h) \frac{dN}{dM_h} dM_h. \quad (\text{C.6})$$

The next step is to compute the fraction of the energy made available by gravitational collapse that goes into unbinding the gas from sub-halos. This is

$$\frac{\widetilde{E}_b(M_h^*)}{E_{\text{th}}(M_h^*)} = \frac{\int_{M_h^{\min}}^{M_h^*} f_g(M_h) M_h T(M_h) \frac{dN}{dM_h} dM_h}{T(M_h^*) \int_{M_h^{\min}}^{M_h^*} f_g(M_h) M_h \frac{dN}{dM_h} dM_h}, \quad (\text{C.7})$$

where we have made use of expressions for $\widetilde{E}_b(M_h^*)$ and $E_{\text{th}}(M_h^*)$ provided respectively in Eqs. C.6 and C.3 and have substituted $M_g(M_h)$ with $f_g(M_h) M_h$. We rewrite $T(M_h)$ as

$$T(M_h) = T(M_{h,o}) \left(\frac{M_h}{M_{h,o}} \right)^{2/3}, \quad (\text{C.8})$$

where $M_{h,o}$ is a reference halo mass and $T(M_{h,o})$ the value temperature takes on at $M_{h,o}$. We change the integration variable from M_h to $m_h = M_h/M_{h,o}$, and with some algebra get

$$\frac{\widetilde{E}_b(M_h^*)}{E_{\text{th}}(M_h^*)} = m_h^{*-2/3} \frac{\int_{m_h^{\min}}^{m_h^*} f_g(m_h) m_h^{5/3} \frac{dN}{dm_h} dm_h}{\int_{m_h^{\min}}^{m_h^*} f_g(m_h) m_h \frac{dN}{dm_h} dm_h}. \quad (\text{C.9})$$

The energy left for the thermalization of the gas is: $1 - \widetilde{E}_b(M_h^*)/E_{\text{th}}(M_h^*)$. Thus, if we define as mixed accretion temperature, T_{ma} , the gas temperature computed accounting for the energy invested in unbinding gas from infalling sub-halos, we find

$$\frac{T_{\text{ma}}(M_h^*)}{T(M_h^*)} = 1 - m_h^{*-2/3} \frac{\int_{m_h^{\min}}^{m_h^*} f_g(m_h) m_h^{5/3} \frac{dN}{dm_h} dm_h}{\int_{m_h^{\min}}^{m_h^*} f_g(m_h) m_h \frac{dN}{dm_h} dm_h}, \quad (\text{C.10})$$

where $T(M_h^*)$ is the temperature computed assuming all gas is outside infalling sub-halos. In Fig. C.1 we show T_{ma}/T as a function of halo mass.

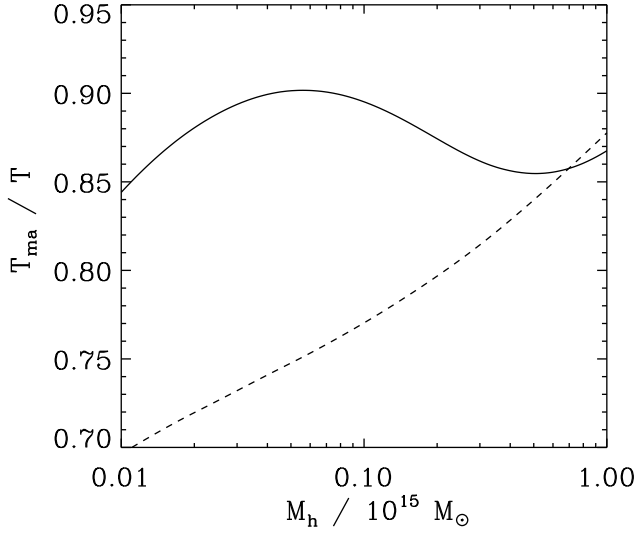


Fig. C.1. Mixed accretion temperature, see text for definition, divided by temperature computed assuming all gas is outside infalling subhalos. The thick line shows the case of a gas fraction that varies with halo mass as in Fig. 8. The dashed line shows the case of a gas fraction that does not depend on halo mass.

Appendix D: A derivation of the total gas energy

In this appendix we derive the expression for the thermal energy of the gas presented in Eq. 23. In general terms, the thermal energy within a given volume, V , can be written as

$$E_{\text{th}} = \int_V h \rho_g dV, \quad (\text{D.1})$$

where h , is the specific enthalpy (enthalpy per unit mass) and ρ_g is the gas mass density. Under the standard assumption that gas in the halo can be described as an ideal monoatomic gas:

$$h = \frac{5}{2} \frac{k_b T}{\mu m_p}. \quad (\text{D.2})$$

By assuming spherical symmetry and expressing the product of temperature with density as pressure we rewrite D.1 as

$$E_{\text{th}}(R) = 10 \pi \int_0^R p_g r^2 dr. \quad (\text{D.3})$$

Next, we rescale pressure in units of p_{500} and radius in units of R_{500} :

$$E_{\text{th}}(x) = \frac{15}{2} p_{500} V_{500} \int_0^x \mathcal{P}(x) x^2 dx, \quad (\text{D.4})$$

where: p_{500} is the characteristic pressure (see Arnaud et al. 2010, for a definition), $V_{500} = 4/3\pi R_{500}^3$, $x = R/R_{500}$ and $\mathcal{P}(x) = p(x)/p_{500}$. We note also that within the integral we used the symbol x rather than r to avoid confusion with the upper integration limit.