A&A 646, A92 (2021)
https://doi.org/10.1051/0004-6361/202038501

Astronomy & Astrophysics

Iron in X-COP: Tracing enrichment in cluster outskirts with high accuracy abundance profiles

Simona Ghizzardi[1], Silvano Molendi[1], Remco van der Burg[2], Sabrina De Grandi[3], Iacopo Bartalucci[1], Fabio Gastaldello[1], Mariachiara Rossetti[1], Veronica Biffi[4,5,6], Stefano Borgani[7,8,4,9], Dominique Eckert[10], Stefano Ettori[11,12], Massimo Gaspari[11,13,⋆], Vittorio Ghirardini[14], and Elena Rasia[7,4]

[1] INAF – Istituto di Astrofisica Spaziale e Fisica Cosmica di Milano, Via A. Corti 12, 20133 Milano, Italy
e-mail: simona.ghizzardi@inaf.it
[2] European Southern Observatory, Karl-Schwarzschild Str. 2, 85748 Garching bei Muenchen, Germany
[3] INAF – Osservatorio Astronomico di Brera, Via E. Bianchi 46, 23807 Merate, LC, Italy
[4] Institute of Fundamental Physics of the Universe, Via Beirut 2, 34151 Grignano, Trieste, Italy
[5] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA
[6] Universitäts-Sternwarte München, Scheinerstr. 1, 81679 München, Germany
[7] Dipartimento di Fisica, Sezione di Astronomia, Università di Trieste, Via Tiepolo 11, 34143 Trieste, Italy
[8] INAF-Osservatorio Astronomico Trieste, Via Tiepolo 11, 34123 Trieste, Italy
[9] INFN, Instituto Nazionale di Fisica Nucleare, Trieste, Italy
[10] Department of Astronomy, University of Geneva, Ch. d'Ecogia 16, 1290 Versoix, Switzerland
[11] INAF–Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, 40129 Bologna, Italy
[12] INFN, Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[13] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
[14] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Received 26 May 2020 / Accepted 12 October 2020

ABSTRACT

We present the first metal abundance profiles for a representative sample of massive clusters. Our measurements extend to R_{500} and are corrected for a systematic error plaguing previous outskirt estimates. Our profiles flatten out at large radii, admittedly not a new result, however the radial range and representative nature of our sample extends its import well beyond previous findings. We find no evidence of segregation between cool-core and non-cool-core systems beyond $\sim 0.3\,R_{500}$, implying that, as was found for thermodynamic properties, the physical state of the core does not affect global cluster properties. Our mean abundance within R_{500} shows a very modest scatter, <15%, suggesting the enrichment process must be quite similar in all these massive systems. This is a new finding and has significant implications for feedback processes. Together with results from the thermodynamic properties presented in a previous X-COP paper, it affords a coherent picture in which feedback effects do not vary significantly from one system to another. By combining intra-cluster medium with stellar measurements we have found the amount of Fe diffused in the intra-cluster medium to be about ten times higher than that locked in stars. Although our estimates suggest, with some strength, that the measured iron mass in clusters is well in excess of the predicted one, systematic errors prevent us from making a definitive statement. Further advancements will only be possible when systematic uncertainties, principally those associated with stellar masses, both within and beyond R_{500}, can be reduced.

Key words. X-rays: galaxies: clusters – galaxies: clusters: general – galaxies: clusters: intracluster medium – large-scale structure of Universe

1. Introduction

Fall into clusters of galaxies, which are the largest gravitational wells in the Universe, must lead to the heating of gas to very high temperatures; what is by no means a forgone conclusion is that the same gas is enriched in heavy elements. Indeed, while the former property of the intra-cluster medium (ICM) results from simple gravitational collapse, the latter requires that gas be processed in stars and re-ejected into the ICM in the form of heavy elements. Radiative cooling and feedback processes can of course leave their imprint on the thermodynamic properties of the ICM, however, beyond core regions, they do so in the form of modest modifications over substantial gravitational heating (Pratt et al. 2010); conversely, metals in the hot gas phase trace exclusively feedback mechanisms[1].

Metals can in principle be a powerful probe. Their quantity and distribution in the ICM can be used to provide important clues on the nature of feedback processes associated to stars and active galactic nuclei (AGN; see Fabjan et al. 2010; Biffi et al. 2018; Truong et al. 2019). An adequate radial range is needed both to measure the distribution and assess the total amount of Fe in the ICM and, ultimately, to constrain feedback processes occurring at high redshift during and perhaps

⋆ Lyman Spitzer Jr. fellow.

[1] Throughout this paper we will generically refer to "feedback" processes that impact on cluster global properties. These should not be confused with feedback processes in cluster cores (e.g., Gaspari et al. 2011).

even before the proto-cluster formation phase. Unfortunately, extending X-ray measurements out to large radii is quite difficult, particularly so for abundances. As discussed in Molendi et al. (2016), reliable measurements have been carried out to $\sim 0.6\,R_{500}$, sampling about one third of the total gas mass in clusters and only on archival samples with no guarantee of representativity (Leccardi & Molendi 2008a; Matsushita 2011; Mernier et al. 2017; Lovisari & Reiprich 2019). Moreover measurements are known to be plagued by a variety of systematic issues (Buote 2000; Leccardi & Molendi 2007, 2008a,b). In the last decade, some measurements have been performed up to $\lesssim R_{200}$ (Simionescu et al. 2011, 2017; Werner et al. 2013; Urban et al. 2014, 2017) on a few objects with Suzaku. For several of these systems the azimuthal coverage is limited, progressively decreasing when moving outward, thereby mapping a small portion of the cluster volume in the outskirts. Metal profiles for the Perseus cluster were derived along six azimuthal directions and show little scatter, suggesting a uniform enrichment.

A full census of metals in clusters requires that the stellar component also be estimated. This is not a trivial measurement in itself, moreover it requires that assumptions be made, for example, about the initial mass function (IMF), the stellar population synthesis models, and stellar formation histories (see Behroozi et al. 2010, for a detailed discussion). Thus, claims that the amount of Fe in the ICM is in excess of what can be produced by the stars in the cluster (Renzini & Andreon 2014; Loewenstein 2013) should be regarded with some caution.

In this paper we derive abundance profiles for the X-COP[2] sample (Eckert et al. 2017). Our measurements are unprecedented for four reasons: (1) they are carried out on a representative (Eckert et al. 2017), albeit small, sample; (2) they are made on massive systems ($M_{500} > 3.5 \times 10^{14}\,M_{\odot}$), which are systems that better approximate the ideal "closed box"; (3) they extend, for virtually all systems, to R_{500}, and (4) they have been corrected for a systematic error that has affected most past Fe abundance measurements.

The paper is organized as follows. In Sect. 2 we describe our *XMM-Newton* data set, data reduction, and analysis, including modifications to the spectral analysis, specifically developed for Fe abundance measurements. The most important of these is the exclusion of the L-shell region from spectral fitting. In Sect. 3 we present results from the X-ray analysis, we show the X-COP Fe abundance profile, and we compare it with those measured from other samples such as the ones reported in Leccardi & Molendi (2008a), Lovisari & Reiprich (2019) and Mernier et al. (2017). We also compute deprojected Fe profiles and Fe mass profiles, and we investigate scaling relations between the Fe mass and other cluster observables. In Sect. 4 we present stellar mass profiles for 7 of our 12 systems and investigate scaling relations between stellar mass and X-ray observables. In Sect. 5, we combine X-ray and optical measurements to take a census of iron in clusters and compare the total estimated Fe mass with the one expected from supernovae (SN). In Sect. 6 we interpret and discuss our results.

Throughout the paper, we assume a Λ cold dark matter cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.3$, $\Omega_\Lambda = 0.7$, and $E(z) = \sqrt{\Omega_\mathrm{m}(1+z)^3 + \Omega_\Lambda}$ for the evolution of the Hubble parameter. For each cluster we use as reference radii R_{500} and R_{200}, defined at the over-densities of $\Delta = 500$ and 200, respectively, with respect to the critical value $\rho_\mathrm{c} = 3H_0^2 \frac{E(z)^2}{8\pi G}$, and computed using the X-COP hydrostatic mass profiles in Ettori et al. (2019). All the quoted errors hereafter are at the 1σ confidence level.

[2] The XMM Cluster Outskirts Project (X-COP) is an *XMM-Newton* Very Large Program dedicated to Cluster outskirts. It has been extracted from the *Planck* PSZ1 catalog by making a high cut in signal-to-noise ratio and placing further constraints on: minimum apparent size, redshift range, and maximum galactic N_H.

2. X-ray data analysis

XMM-Newton observations, the X-ray data analysis pipeline, and the sample are described in detail in Ghirardini et al. (2018, 2019). In the following we summarize the main steps of this analysis.

2.1. Data reduction

All data were reduced using XMMSAS v13.5 and the Extended Source Analysis Software (ESAS; Snowden et al. 2008). The first basic steps produce calibrated events files for each observation (with *emchain* and *epchain*; the chain for the pn detector is run twice to create events files also for pn out-of-time events). We filtered out time periods affected by soft proton flares with *mos-filter* and *pn-filter*. We then estimated the contamination of residual soft protons to the spectrum by comparing IN and OUT count rates, where IN are the count rates of the MOS camera measured in the 7.5−11.8 keV energy band from regions inside the field of view (FoV), and OUT outside the FoV, in the unexposed corners of the MOS detectors (De Luca & Molendi 2004; Leccardi & Molendi 2008b). We ran the XMMSAS tool *ewavelet* to detect point sources within the FoV and corrected the resulting point source list for the spatial dependence of the fraction of the cosmic X-ray background (CXB) that is resolved by the instrument. This correction consists in excising only the sources with a measured count rate greater than a certain count rate threshold, determined by comparing our count rate distribution with the Log N–Log S distribution of CXB sources, down to which our source detection is complete (details in Ghirardini et al. 2018). We left the fainter sources to enforce a constant flux threshold across the FoV and avoid biasing local measurements of the CXB intensity.

2.2. Spectral analysis in X-COP

We extracted spectra in concentric annuli around the X-ray peak up to $\sim R_{500}$, estimated from the hydrostatic mass (see Ettori et al. 2019), with the ESAS routines *mos-spectra* and *pn-spectra*. Subsequently we used filter-wheel-closed data to estimate the high-energy particle background contribution with *mos-back* and *pn-back*. Products of these tools are also appropriate response matrices and effective areas files for extended sources. The output spectra were re-binned with a minimum of 5 counts per bin to ensure stable fitting results, and the data below 0.5 keV were discarded to avoid EPIC (European Photon Imaging Camera) calibration uncertainties in this energy range.

We used `XSPEC` v12 (Arnaud 1996) and `ATOMDB` v3.0.7 to fit the spectra and determine the plasma best-fit parameters according to the Cash statistics. Our strategy to estimate the physical quantities from the observed spectra consists in modeling all the individual background components and the source spectra. The background components, described in detail in Ghirardini et al. (2018), are the high-energy particle background, the sky background, and residual soft protons. We modeled the source emission of each region with an absorbed single temperature APEC

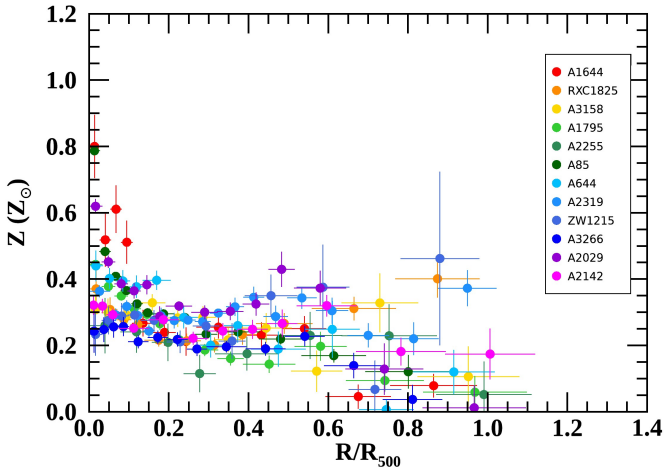

Fig. 1. Iron abundance profiles as a function of R/R_{500}, derived with standard X-COP analysis (see Sect. 2.2). Clusters in the inset are ordered by their total mass starting from the least massive, which is A1644. Abundances reported here and throughout the X-ray sections of this paper are in solar units as defined in Anders & Grevesse (1989).

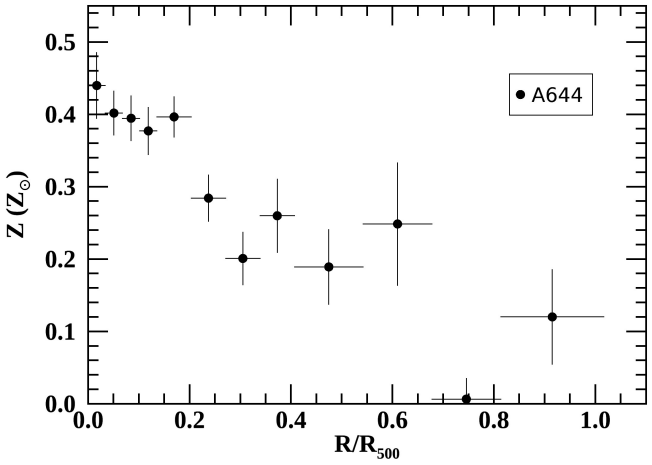

Fig. 2. Iron abundance profile for A644 as a function of R/R_{500}, derived with standard X-COP analysis (see Sect. 2.2).

model (Smith et al. 2001) with temperature, emission measure, and metal abundance free to vary. We note that, in the current version of XSPEC, the metal abundance parameter in APEC has a lower hard boundary at 0; this has significant implications that will be discussed in Sect. 2.3.3. The solar abundance table was set to Anders & Grevesse (1989). MOS and pn spectra of each region were fitted jointly, as are spectra of different observations of the same regions. The energy range considered in the fitting is 0.5−12 keV. We ignored energy ranges where bright and time variable fluorescence lines are present; these are 1.2−1.9 keV for the two MOS and 1.2−1.7 keV and 7.0−9.2 keV for the pn. Gas temperature, density, entropy, and pressure profiles for the whole X-COP sample are presented in Ghirardini et al. (2019), and Fe abundance profiles are plotted in Fig. 1. Hence forward we will refer to this analysis as the "standard" analysis.

2.3. Modifications to X-COP spectral analysis

In this subsection we shall look into modifications of the X-COP spectral analysis specifically designed to improve abundance measurements.

2.3.1. Vanishing metals

Inspection of Fig. 1 shows that, in several instances, the measured abundance drops to a value close to zero, in a specific bin, only to increase again in the following. What is even more puzzling is that the computed error on the near zero measure is very small; an example of this is provided by A644 (see Fig. 2). Similar jumps in abundance profiles have been found by other authors. In Mernier et al. (2017) the abundance profiles of A2029, A133, and A2597 (see their Appendix A) feature a prominent drop in metallicity; incidentally, for A2029, which is also in our sample, we measure a very similar jump. In Lovisari & Reiprich (2019), no individual profiles are reported, however, their Fig. 4 shows that for $R > 0.7\,R_{500}$ profiles tend to separate out with a lower branch located close to 0. We investigated the nature of these jumps through dedicated simulations.

In Fig. 3 we show a simulation comprising three main components: a source (red line), modeled as a thermal spectrum (APEC) with a temperature of 5 keV, an abundance of $0.3\,Z_\odot$, and a surface brightness of 2.5×10^{-14} erg cm^{-2} s^{-1} in the 0.5−2.0 keV band, typical of intermediate radii; a sky component (green line), comprising both X-ray foregrounds and backgrounds and last, but by no means least, an instrumental background component (blue line). In the bottom panel of Fig. 3 we show the simulated data over model ratio, where the abundance parameter of the APEC component has been set to 0. We clearly see that the abundance measure is driven by the Kα line at 6.7 keV. In Fig. 4 we show the same simulation with one difference: the source surface brightness has been dialed down by a factor of 10 to 2.5×10^{-15} erg cm^{-2} s^{-1} in the 0.5−2.0 keV band. In this new simulation the instrumental background dominates above ∼2 keV and the sky foregrounds below ∼0.7 keV. An intriguing consequence, highlighted in the bottom panel of Fig. 4, is that the intensity of the Fe L-shell emission with respect to the total continuum is comparable to that of the Fe Kα line and that both lines contribute to the abundance measurement.

From these exercises, we infer that as we move out to larger radii and the surface brightness of the ICM becomes progressively smaller, the L-shell emission, or more precisely what the fitting algorithm attributes to L-shell emission (see Sect. 2.3.2 for details), plays an increasing role in the measurement of the abundance.

In light of these findings, we decided to test whether the L-shell emission is responsible for the sudden jumps and drops in abundance seen in Figs. 1 and 2. We refitted our spectra excluding the 0.9−1.3 keV energy range, where the emission is observed. Our revised profiles are plotted in Fig. 5. As we can see: (1) most measurements close to 0 are shifted to higher values; (2) the errors associated to the new measurements are larger, and (3) several measurements located at relatively high abundances have shifted down. The overall result is that, in the outskirts, the abundance profiles of the different clusters are now less scattered and are indicative of a flat metal abundance distribution.

To quantify the changes obtained by fitting only the Fe Kα line, we compare mean abundances derived using the standard fitting method (that we label as "standard" procedure) with those derived by excluding the 0.9−1.3 keV energy range (that we refer to as "modified" procedure). In the comparison we exclude the central regions where the source surface brightness is high and spectral results are insensitive to systematics associated to

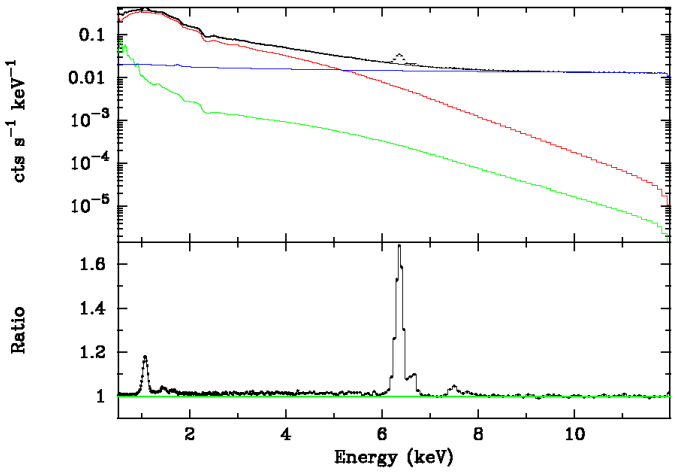

Fig. 3. Simulated ICM spectrum from an intermediate radius. *Top panel*: breakdown of the total spectrum (black) in source (red), sky foreground, and background (green), and instrumental background (blue) components. *Bottom panel*: to highlight the role of line emission, we show the ratio of simulated data to model with the abundance parameter in the source model set to 0 (see Sect. 2.3.1 for details).

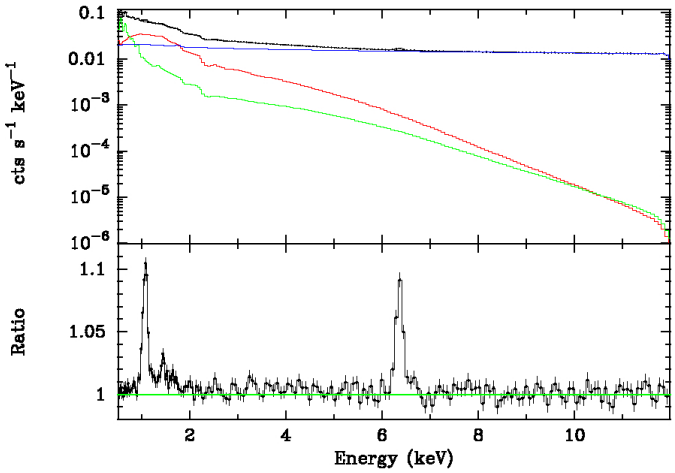

Fig. 4. As in Fig. 3 with the one difference that the surface brightness of the ICM emission has been reduced by a factor of 10 to be representative of cluster outskirts.

the background. Namely, we restrict the test to the radial range $R/R_{500} > 0.3$, which we further split into two sub-intervals: $0.3 < R/R_{500} < 0.6$, and $R/R_{500} > 0.6$. Results are reported in Table 1, where mean abundances and total scatters[3], with their statistical errors, are shown, both for the whole $R/R_{500} > 0.3$ radial range and for the two sub-intervals. The average metal abundance measured through the standard analysis drops significantly from $Z = 0.272 \pm 0.003\,Z_\odot$ for $0.3 < R/R_{500} < 0.6$ to $Z = 0.154 \pm 0.008\,Z_\odot$ for $R > 0.6\,R_{500}$. By applying the modified method, the average abundance remains unchanged from $Z = 0.244 \pm 0.006\,Z_\odot$ for $0.3 < R/R_{500} < 0.6$ to $Z = 0.242 \pm 0.016\,Z_\odot$ for $R > 0.6\,R_{500}$. In Table 1 we report the scatter of the data about the average value in the same radial ranges. When fitting with the standard method, the scatter in the outskirts, $R > 0.6\,R_{500}$, is high, $\sigma = 0.103 \pm 0.014$ (i.e., ∼67%), with a notable increase with respect to the $0.3 < R/R_{500} < 0.6$ radial range. The adoption of the modified recipe reduces the scatter to ∼20% over the whole range $R/R_{500} > 0.3$, with little variations from $0.3 < R/R_{500} < 0.6$ to $R > 0.6\,R_{500}$.

The global change in the profiles induced by this new method can be appreciated in Fig. 6, where we plot the average metallicity profile (derived following the approach detailed in Sect. 3.1) for our sample using the standard spectral analysis (blue area) and the modified spectral analysis (gray area). The new fitting procedure provides a flatter and more uniform profile up to $\sim R_{500}$, with a substantially reduced scatter. As a side note, we point out that the two profiles, which are indistinguishable at small radii, slowly separate as we move to larger radii, which is precisely what we expect if the difference is due to the increasing importance of the background in the spectral analysis.

The modified procedure does not affect temperature estimates. In Fig. 7 we compare the temperatures obtained with the modified recipe with those derived through the standard spectral analysis. Only points beyond $0.3\,R_{500}$ are plotted. The agreement is good, the small differences observed in the high temperature range are associated to known calibration issues (e.g., Nevalainen et al. 2010) between the EPIC soft and hard energy bands. These are, quite likely, the same calibration issues responsible for the underestimation of metal abundances from the L-shell (see Sect. 2.3.2 for details).

2.3.2. L-shell issue

We could simply adopt our new abundance measurements by pointing out that they are self-consistent Fe Kα line measurements and that the effects introduced by Fe L-shell emission contributions do not concern us. However, we prefer to take a different point of view, namely that Fe L-shell emission could, at least in principle, provide useful measurements and if they do not we can, at the very least, speculate as to what the problem might be. An important clue comes from the equivalent width of the L-shell emission, which, for a high temperature plasma ($kT > 4$ keV), is very modest, on the order of a few tens of electronvolts. The reason it becomes important and in some instances dominant, as shown by our analysis, is that the equivalent width of the Fe Kα line with respect to the dominant continuum, that is, the instrumental background, rapidly drops as a function of radius. So the question of why the Fe L measurements are unreliable and Fe Kα are not can be recast into a more meaningful form: why can the Fe Kα measurements be extended to a very low equivalent width while the Fe L ones cannot?

Let us start by examining the Fe Kα measurements. These are made with respect to a well-determined continuum dominated by one highly reproducible component, namely the instrumental background. Moreover, at the relevant energies, the spectral resolution is high and the measurement of the intensity of the line is limited to a narrow energy range. Under these circumstances, the recovered abundance will be characterized by large statistical and small systematic errors or, as some might say, it will be an accurate but not a precise measurement. Let us now consider the Fe L measurements. These are made in a region of the spectrum where more than one continuum component contributes; it is also where effective areas peak and calibration issues will impact most significantly on maximum-likelihood estimations. Under these circumstances, there is a strong

[3] In this paper we report total scatters even if formally intrinsic scatters may be more appropriate. The reason for this is that we suspect that, here and elsewhere, our statistical errors might be somewhat overestimated; if this is indeed the case, estimates of the intrinsic scatter will be biased low and incorrect conclusions about properties of the samples we are investigating might be drawn. By making use of the total scatter as an upper limit to the intrinsic scatter, we adopt a conservative approach that prevents us from making excessive claims.

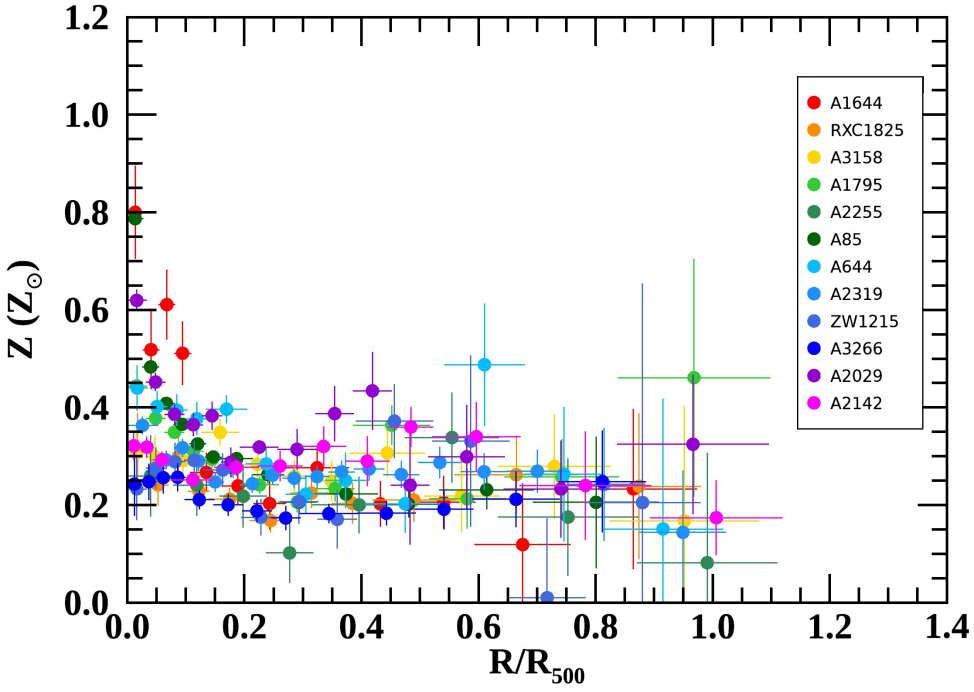

Fig. 5. Iron abundance profiles as a function of R/R_{500}. Abundances are measured solely through the Fe Kα line (see Sect. 2.3.1 for details). Clusters in the inset are ordered as in Fig. 1.

Table 1. Average abundance $\bar{Z}$ (and total scatter σ) in regions: $R/R_{500} > 0.3$, $0.3 < R/R_{500} < 0.6$, and $R/R_{500} > 0.6$ obtained through standard and modified spectral analysis.

	Standard analysis		Modified analysis	
	$\bar{Z}$	σ	$\bar{Z}$	σ
	($/Z_\odot$)			
$R/R_{500} > 0.3$	0.259 ± 0.003	0.067 ± 0.006	0.244 ± 0.005	0.052 ± 0.005
$0.3 < R/R_{500} < 0.6$	0.272 ± 0.003	0.046 ± 0.005	0.244 ± 0.006	0.052 ± 0.006
$R/R_{500} > 0.6$	0.154 ± 0.008	0.103 ± 0.014	0.242 ± 0.016	0.056 ± 0.008

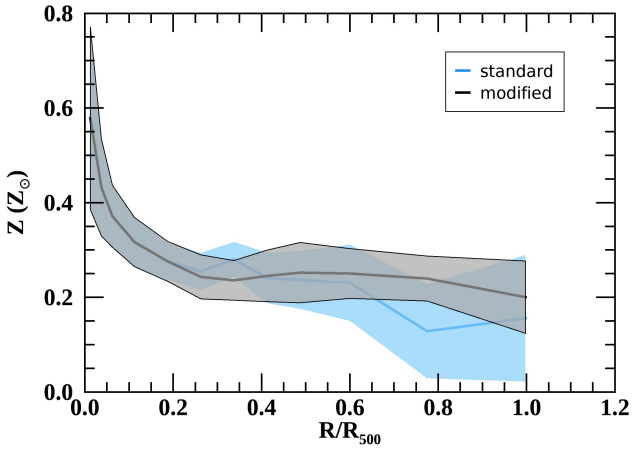

Fig. 6. Average metallicity profile (thick line) obtained through the standard (light blue) and the modified (gray) spectral analysis; the 1σ-scatter is shown as a shaded region.

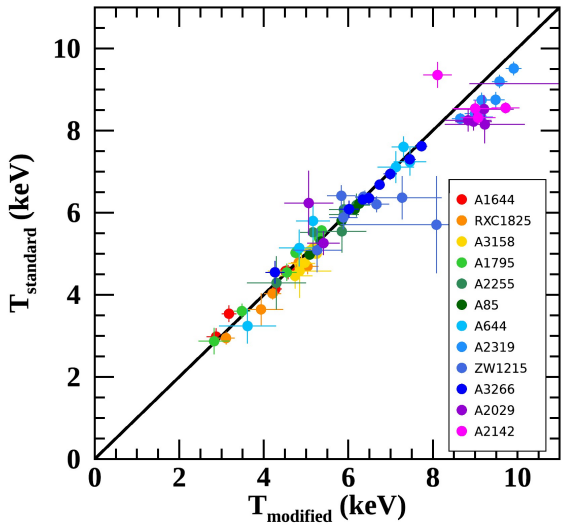

Fig. 7. Comparison between the temperatures for the sample derived through the standard method and the modified method. Only points beyond $0.3\,R_{500}$ are shown. Clusters in the inset are ordered as in Fig. 1.

possibility that the fitting algorithm will use the Fe L-shell emission to "fix" local residua associated either to an insufficient modeling of the continuum components or to the limited calibration of the instrument (see Appendix B for a detailed description of how this might happen). In other words, L-shell measures will have small statistical errors and, often, large undetected systematic errors leading to precise but inaccurate measurements.

Finally, before closing our discussion, we will consider the use of Fe L-shell measurements of hot (e.g., >4 keV) plasmas from an ideal instrument characterized by negligible calibration errors on effective areas and a highly reproducible background. This is not entirely idle speculation as ESA's next large X-ray mission ATHENA (Advanced Telescope for High Energy Astrophysics Nandra et al. 2013) is being designed with very stringent calibration requirements in mind. Under the standard assumption of collisional ionization equilibrium (CIE), for temperatures larger than ∼4 keV, the fraction of Fe in the form of FeXXIV is less then 10% of the total, the bulk being in FeXXV. This immediately implies that, to achieve systematic uncertainties below 10% on any Fe L-shell-based abundance measurement, the CIE will have to be understood to better than the 1% level. We also note that this is a necessary condition but by no means a sufficient one as several other issues enter into the computation of emission lines. Putting to the side atomic transitions, we note that the distribution between different excitation states of FeXXIV ions, which, like CIE, is critically dependent on collisions, needs to be know to a high precision. All these considerations suggest that a robust measure of Fe abundance should be based on transitions occurring in the dominant Fe ion, which, for temperatures in the 2−10 keV range, are Fe K-shell transitions in FeXXV. Use of less abundant ions, such as FeXXIV, requires a detailed understanding of the equilibrium properties of plasmas, which may become available with the advent of high spectral resolution measurements such as those afforded by XRISM (X-ray Imaging and Spectroscopy Mission Tashiro et al. 2018; Guainazzi & Tashiro 2018) and the ATHENA XIFU (X-ray Integral Field Unit Barret et al. 2013; Cucchetti et al. 2019).

2.3.3. Going negative

As pointed out in Sect. 2.2, the metal abundance parameter in APEC has a hard lower limit at 0. Such an approach may seem reasonable at a first glance, however it can lead to biased measurements. In the case at hand, for large radii where abundances are small and statistical errors large, enforcing a hard limit at 0 for the metallicity can lead to biased results. This is discussed in Appendix A of Leccardi & Molendi (2008a) and is treated in greater detail in Appendix C of this paper. The obvious solution to this potential bias is to allow abundances to assume negative values. Unfortunately the current implementation of the APEC model within XSPEC does not allow this and we had to resort to a more indirect approach. We refitted our spectra substituting the APEC component with MEKAL, which does allow for negative values. We compared results between runs where the metallicity was forced to be positive and runs where it was allowed to assume negative values. In the latter case we found no instance where the best fitting value was smaller than zero; we did however find a few measurements where the 1σ confidence interval extended to negative abundances. For these measurements, in the case of forced positive abundance fits, the confidence regions were of course cut off at zero, however, the errors from the show par command, which are the square roots of the diagonal elements of the covariance matrix, were the same as those measured when abundances were allowed to become negative and very close to the error determined by subtracting the best fit value from the upper bound. Thus, at least for the present sample, limiting abundance fits to positive values does not appear to introduce significant biases. We caution our readers that this is by no means a general result and that, for another sample, the result could be quite different. We refer to Appendix C for a more detailed discussion of this issue. Of course, the definitive solution will be to change the XSPEC code to allow negative values for APEC abundances.

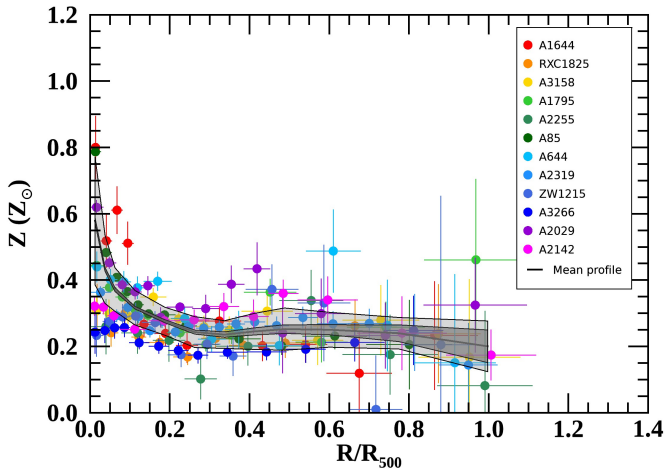

Fig. 8. Same as Fig. 5 with the average profile (thick line) overlaid. The dark and light shaded areas indicate respectively the statistical error and the total scatter. Clusters in the inset are ordered as in Fig. 1.

3. Iron in the intra-cluster medium

3.1. Abundance profiles

As discussed in detail in Sect. 2.3.1, we adopt measurements of the Fe abundance entirely based on the Fe Kα line. Ours are the first profiles extending out to $\sim R_{500}$ for a representative sample of massive systems (for more details see Sect. 1 and Eckert et al. 2017). As expected, they feature a significant spread, spanning from 0.23 to 0.80 $Z_\odot$, in the core and in the circum-core regions, where low core entropy systems have more prominent abundance peaks than high core entropy systems (Leccardi et al. 2010). Profiles flatten out beyond $R > 0.3\,R_{500}$ to values between 0.15 and 0.35 $Z_\odot$, with an averaged value of $Z = 0.244 \pm 0.005\,Z_\odot$, and they remain remarkably flat out to R_{500}. In addition, the distribution of metal abundance appears uniform across the sample, with little variation from cluster to cluster at radii larger than $0.3\,R_{500}$ (total scatter ∼20%, see Table 1).

In Fig. 8 we overlay the average profile on the individual ones. The bins adopted for the average profile are listed in the first column of Table 2. To derive the average profile, we followed the approach described in Leccardi & Molendi (2008b): in each bin, the average metallicity has been computed by performing a weighted average on values of the different clusters whose original bins are (even partially) included in the bin of the average profile. More precisely, we assign a weight $w_{i,j,k}$ to the jth bin of the ith cluster, which measures the fraction of the jth bin included in kth bin of the average profile; j bins that are totally included in a k bin are assigned a weight of 1. These weights are combined with the usual statistical weights $1/\sigma^2_{i,j}$ associated to the abundance measurements.

Thus, the average metallicity in the kth bin is given by

$$\bar{Z}_k = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i} w_{i,j,k}\frac{Z_{i,j}}{\sigma^2_{i,j}}}{\sum_{i=1}^{n}\sum_{j=1}^{m_i} w_{i,j,k}\frac{1}{\sigma^2_{i,j}}}, \quad (1)$$

where $Z_{i,j}$ and $\sigma_{i,j}$ are, respectively, the metal abundance and the error measured in the jth bin of the ith cluster. Index i runs from

Table 2. Mean abundance ($\bar{Z}$) profile for the full sample, the cool-core sub-sample, and the non-cool-core sub-sample.

	Full sample				Cool-core subsample			Non-cool-core subsample		
Radius	$\bar{Z}$	ϵ	σ	$Z_{\rm median}$	$\bar{Z}$	ϵ	σ	$\bar{Z}$	ϵ	σ
($/R_{500}$)						($/Z_\odot$)				
0.000–0.025	0.578	0.008	0.193	0.440 ± 0.024	0.648	0.009	0.158	0.330	0.017	0.048
0.025–0.050	0.432	0.006	0.103	0.363 ± 0.013	0.475	0.007	0.086	0.325	0.011	0.048
0.050–0.075	0.371	0.006	0.066	0.293 ± 0.014	0.408	0.007	0.040	0.297	0.010	0.038
0.075–0.150	0.317	0.004	0.052	0.299 ± 0.008	0.342	0.005	0.037	0.276	0.006	0.046
0.150–0.225	0.276	0.005	0.042	0.275 ± 0.010	0.291	0.006	0.024	0.258	0.007	0.051
0.225–0.300	0.243	0.006	0.046	0.241 ± 0.011	0.259	0.009	0.044	0.231	0.008	0.045
0.300–0.375	0.236	0.008	0.042	0.249 ± 0.013	0.236	0.014	0.044	0.236	0.010	0.041
0.375–0.450	0.245	0.011	0.054	0.274 ± 0.022	0.252	0.019	0.069	0.242	0.013	0.046
0.450–0.525	0.252	0.013	0.064	0.263 ± 0.034	0.237	0.023	0.065	0.259	0.016	0.062
0.525–0.675	0.250	0.015	0.053	0.268 ± 0.023	0.222	0.027	0.028	0.263	0.018	0.056
0.675–0.875	0.240	0.025	0.047	0.240 ± 0.038	0.231	0.048	0.041	0.243	0.030	0.049
0.875–1.120	0.200	0.049	0.076	0.174 ± 0.083	0.316	0.108	0.081	0.170	0.055	0.035

Notes. ϵ and σ are respectively the statistical error and the total scatter (see text). For the full sample we also report the median abundance profile $Z_{\rm median}$, along with the 1σ statistical error, which will be discussed in Appendix B.

1 to the number of clusters in the sample, n, while j runs from 1 to the number of bins of the profile of the ith cluster, m_i.

Similarly, statistical errors ϵ_k are given by

$$\epsilon_k = \frac{1}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m_i} w_{i,j,k}\frac{1}{\sigma_{i,j}^2}}}, \quad (2)$$

and the total scatter σ_k is obtained as

$$\sigma_k = \frac{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m_i} w_{i,j,k}\frac{(Z_{i,j}-\bar{Z}_k)^2}{\sigma_{i,j}^2}}}{\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m_i} w_{i,j,k}\frac{1}{\sigma_{i,j}^2}}}. \quad (3)$$

The average profile $\bar{Z}$ is plotted in Fig. 8 and values are reported in Table 2 along with the statistical errors ϵ and the total scatters σ (Cols. 2–4).

The mean abundances beyond $0.3\,R_{500}$ range from 0.2 to $0.25\,Z_\odot$, remaining extremely flat all the way out to R_{500}. Measurements have very small statistical uncertainties, smaller than 6% within $0.675\,R_{500}$ and still smaller than 25% in the two last bins covering the $[0.675-1]\,R_{500}$ range. Excluding the central bins where the scatter of the metallicity is dominated by the diversity between cool-core and non-cool-core systems, the total scatter stays below $0.065\,Z_\odot$ ($\lesssim$25%) on almost all the cluster volume, except for the last bin, where, due to the large statistical errors, the total scatter reaches $0.076\,Z_\odot$ (∼38%).

To investigate any dependence of the metal abundance on the dynamical state of the clusters, we divided our sample into two subsamples, namely cool-core and non-cool-core systems. To separate clusters into these two classes, we adopt as an indicator the central entropy K_0 provided by Cavagnolo et al. (2009). We consider as cool cores all the clusters with central entropy $K_0 < 30\,{\rm keV\,cm^2}$. Our sample includes four cool-core systems and eight non-cool-core systems. In Fig. 9 we report mean profiles separately for the two subsamples. Values for the mean profiles are reported in Table 2 along with the statistical errors ϵ and the total scatter σ (Cols. 6–11). As expected,

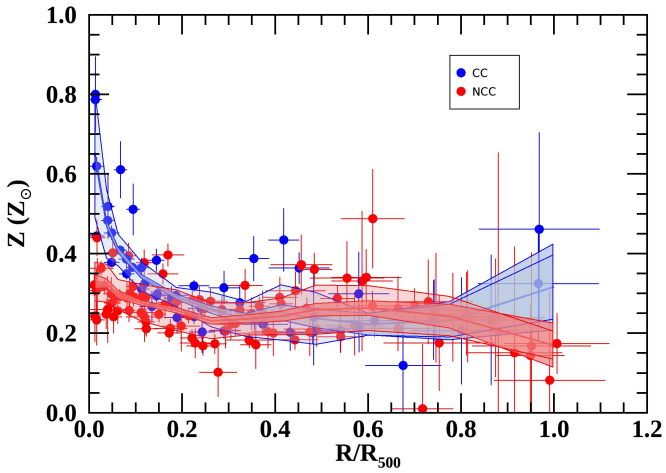

Fig. 9. Abundance profiles as a function of R/R_{500} for cool-core (blue) and non-cool-core (red) subsamples. Average profiles for both subsamples are overlaid. The dark and light shaded areas indicate respectively the statistical error and the total scatter.

the two average profiles differ in the innermost bins. On the contrary, they do not reveal any significant discrepancy beyond $0.3\,R_{500}$. The overall mean metal abundance at $R > 0.3\,R_{500}$ is $Z = 0.254^{+0.018}_{-0.013}\,Z_\odot$ with a total scatter $\sigma = 0.058 \pm 0.009\,Z_\odot$ (∼23%) for the cool-core subsample and $Z = 0.249^{+0.008}_{-0.008}\,Z_\odot$ with a total scatter $\sigma = 0.051 \pm 0.005\,Z_\odot$ (∼20%) for the non-cool-core subsample, showing an excellent agreement between the two subsamples.

From Figs. 5–9 we derive three key results: (1) the metal abundance profile is remarkably flat for $R > 0.3\,R_{500}$; (2) the profile beyond $0.3\,R_{500}$ shows a modest scatter implying that, whatever the enrichment mechanism may be, it must provide a uniform level of metal abundance for all the clusters over most of their volume; (3) cool-core and non-cool-core clusters have similar metal distributions beyond $0.3\,R_{500}$. If, on the one hand, differences in the metal content at the center of clusters can be ascribed to the early contribution of stars currently residing in the brightest cluster galaxies (BCGs) invariably found in

low-entropy systems (see De Grandi et al. 2004, 2014 for a discussion on this point), on the other, the lack of differences in the rest of the cluster volume implies the enrichment mechanisms at work in those regions are not related to phenomena that are typical of only one of the two classes, such as relatively recent merging events, which are signatures of disturbed systems (see also Urdampilleta et al. 2019; Mernier et al. 2018).

3.2. Systematic uncertainties

In Sect. 2.3.1 we have identified and corrected a major source of systematic errors on measurements in low surface brightness regions located at large radii. Here we investigate further sources of systematic errors through a complementary approach: we turn to hydro-dynamic simulations.

An important assumption, made in all estimates of the Fe abundance, is that spectra can be fit with single temperature and abundance models, although it is well known that some degree of multi-temperature and multi-abundance has to be present (e.g., Molendi et al. 2016) and that this can have an impact on measurements (e.g., Buote 2000). Simulations have been employed to address this point by several authors. Rasia et al. (2008) produced spectra starting from the distribution of temperature and abundance in their simulations, convolved them with instrumental response functions, and then fitted them with single temperature and abundance models. They found that, for temperatures above ∼3 keV where the abundance measure is dominated by the Kα line, the measured metallicity is within 5% of the emission weighted (EW) abundance derived directly from the simulation. Since all our spectra feature temperatures in excess of 3 keV (see Fig. 7) and the vast majority in excess of 4 keV, we can conclude that any bias in the measurements will be very small and confined to those spectra with temperatures smaller than 4 keV. An important aside, which we will pick up again in Sect. 6, is that only for the external regions of massive clusters, where Kα emission is sufficiently strong, can we proceed with a measure of the metallicity that is not plagued by the many biases (e.g., Rasia et al. 2008), and plasma code uncertainties (Mernier et al. 2018) associated to L-shell emission.

Having argued that Kα based measurements of the abundance are affected by very modest biases with respect to the EW abundance, we now quantify a possible systematic error due to the assumption that the spectrally measured abundances, which are emission-weighted, are equivalent to mass-weighted (MW) ones. Differences between EW and MW estimates arise when multi-temperature plasma is present, as is the case in outer regions where, on top of declining temperature profiles, clumps of denser and colder material might be present. Given the correlation between entropy and metallicity found both in observational data (e.g., Leccardi et al. 2010) and in cosmological simulations (Biffi et al. 2017), these clumps of high-density, low-entropy gas could bias high the EW abundance estimate.

The complex conditions described above cannot be captured by simple simulations such as those presented in Sect. 2.3. We therefore investigate this aspect through a set of clusters extracted from hydrodynamical cosmological simulations. The general properties of this suite, derived from a modified version of GADGET-3, are described in Rasia et al. (2015). The history of the ICM enrichment and the origin of the metallicity in the outskirts in present-day clusters are discussed respectively in Biffi et al. (2017, 2018), where details of the sub-grid models linked to stellar formation, stellar evolution, and chemical production are also provided. We consider 29 massive clusters (25 with temperatures within R_{500} in the range 4−10 keV and 4 with temperatures around 2−3 keV) and produce both EW and MW maps of Fe with the Smac code (Dolag et al. 2005). The maps are generated in pixels of about 4 kpc on the side and span 10 Mpc in projection along the line of sight, thereby including the contribution of gas particles in the cluster surroundings. We divided each map into concentric annuli centered on the emission peak with bounding radii 0−0.2 R_{500}, 0.2−0.4 R_{500}, 0.4−0.6 R_{500}, 0.6−0.8 R_{500}, 0.8−1.0 R_{500}, and 1.0−1.2 R_{500} and averaged the EW and MW abundances over these regions. Next, we computed the EW over MW ratio radial profile for each system and finally averaged it over all systems. As expected (see Fig. 10) the bias increases, but only moderately, as we move from center to periphery; it lies roughly between 5% and 10% over the radial range of interest ∼0.4−1.0 R_{500}.

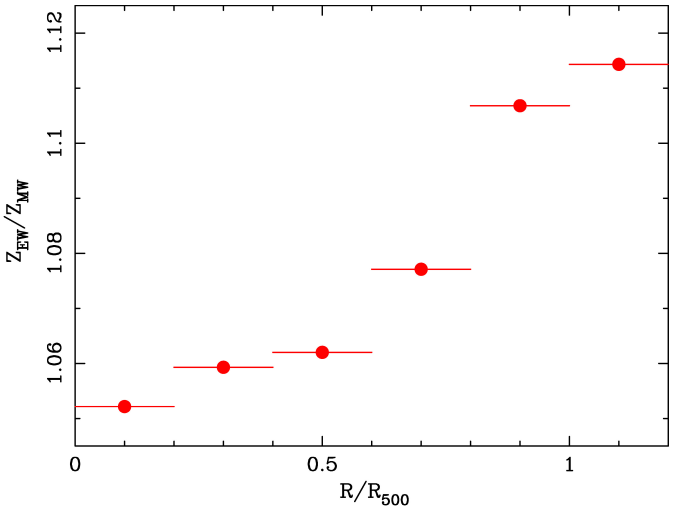

Fig. 10. Ratio of the EW to MW abundance profiles as a function of R/R_{500} averaged over 29 simulated clusters (see Sect. 3.2 for details).

Finally we combine the two effects: for the bias on the estimate of the EW abundance from the single phase spectral modeling we take the 5% value estimated by Rasia et al. (2008), and for the EW to MW bias, using results presented in Fig. 10, we take 10%. We combine these two measurements and assume an overall 15% systematic error on abundances. We note that this is a conservative estimate as the two effects go in opposite directions and should tend to cancel each other out, rather than add up.

Our systematic error is smaller than typical measurement errors for individual systems, implying that, on single clusters, we are dominated by statistical uncertainties. Conversely, as shown in Fig. 8 and Tables 1 and 2, our systematic error is larger than the statistical error on the average profile, suggesting that, despite the relatively small size of our sample, further improvements in the measurement of the mean cluster abundance in the outskirts will come from advancements in analysis methods rather than from an increase in sample size.

3.3. Comparison with previous measurements

In Fig. 11 we compare our abundance profile with those obtained by other authors. The profile reported in Leccardi & Molendi (2008a) extends to about 0.6 R_{500} and in its outermost bins is characterized by a very large scatter. It is worth pointing out that part of this scatter reflects background systematics, which at the time were identified, but only partially understood.

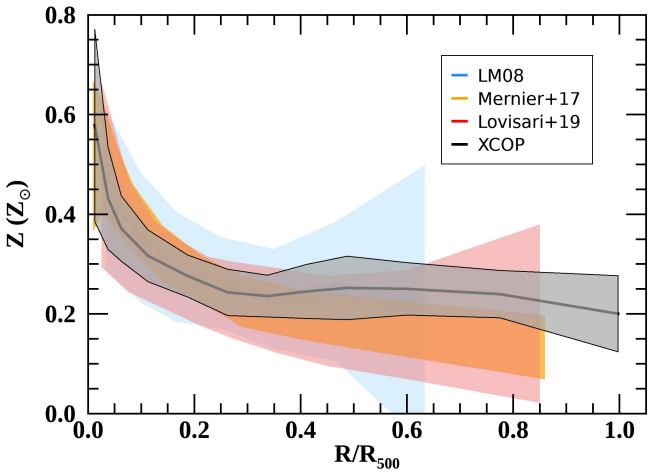

Fig. 11. Abundance profiles as a function of R/R_{500}: comparison between different measurements.

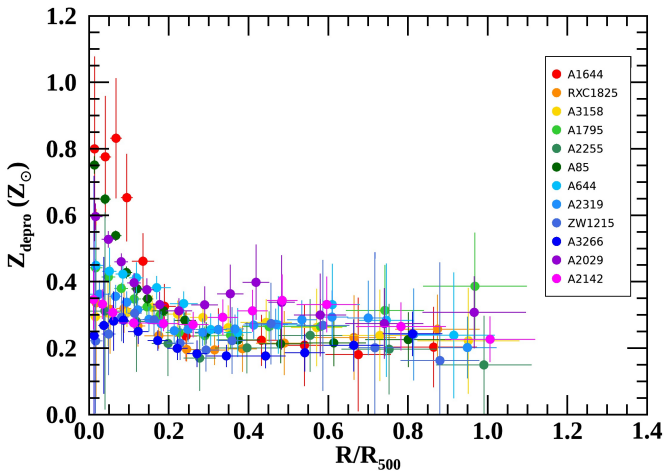

Fig. 12. Deprojected iron abundance profile as a function of R/R_{500}. Clusters in the inset are ordered as in Fig. 1.

Our profile fits comfortably within the rather weak bound posed by the Leccardi & Molendi (2008a) measurements.

The CHEmical Enrichment RGS Sample (CHEERS) sample (Mernier et al. 2017) was constructed to investigate cluster cores rather than outskirts; it comprises mostly nearby systems and only for a handful of systems can measurements be extended beyond intermediate radii. Starting from $\sim 0.3\,R_{500}$ their mean profile features a slow decline. While some overlap between the shaded regions indicating the scatter of the CHEERS and X-COP profiles is present all the way out to the largest radii, mean values start to differ significantly beyond $0.45\,R_{500}$. The decline seen in this profile might be due to the same artifact we have found in our own sample. Indeed, as can be seen in the appendix of Mernier et al. (2017), three of their systems, namely A133, A2029, and A2597, show evidence of a sudden drop in metal abundance at large radii. In the case of A2029, which is also part of our sample, the drop is essentially the same we find from our "standard" analysis. However, it must be recognized that other factors could contribute to shaping the mean abundance profile reported in Mernier et al. (2017), more specifically: (1) the CHEERS sample includes a lot of cooler systems (galaxy groups and elliptical galaxies); (2) it was fitted with SPEX rather than XSPEC.

Finally, we compare our measurements with those from Lovisari & Reiprich (2019). As for the CHEERS sample, the focus of the analysis is on the central regions, namely on the impact of the AGN on the central metal abundance, rather than on the outskirts, and only for a limited number of systems can abundances be measured out to large radii. Unfortunately, the authors do not provide profiles for individual systems, however, from their Fig. 4, we see that beyond $\sim 0.6\,R_{500}$ measurements tend to split up with some around zero and others at larger values. This could be yet another manifestation of the artifact described in Sect. 2.3.1.

In summary, what emerges from this comparison is that our abundance profile is unique in several ways: (1) it is constructed from a representative sample of massive systems; (2) it is based on individual profiles that, thanks to the offset pointings afforded by the X-COP Very Large Program, extend to R_{500}; (3) it features a small statistical scatter all the way to large radii and last but not least, (4) it has been corrected for a major systematic error that has impacted on previous measurements. As we shall see in the next sections, these properties allow us to perform an estimate of the metal content in massive clusters of unprecedented quality.

3.4. Deprojection and cumulative iron masses.

We define the deprojected iron abundance as: $Z_{\rm depro} = n_{\rm Fe}/(Z_{\rm n,\odot} n_{\rm H})$, where $n_{\rm Fe}$ and $n_{\rm H}$ are the iron and hydrogen densities, by number, respectively, and $Z_{\rm n,\odot}$ is the solar iron abundance by number. Assuming solar abundances reported in Anders & Grevesse (1989), $Z_{\rm n,\odot} = 4.68 \times 10^{-5}$. We adopted the standard onion-skin technique to deproject abundances (Kriss et al. 1983; Ettori et al. 2002), including a correction factor to account for the emission of the cluster beyond the outermost bin (see Ghizzardi et al. 2004; McLaughlin 1999, for details). Data have been slightly smoothed, with a boxcar average of three-points width, before deprojection, to reduce nonphysical fluctuations that would be enhanced by the deprojection process. In Fig. 12 we show the deprojected abundance profiles for the whole sample. Projected and deprojected Z profiles for individual cluster are reported in Appendix A. The average deprojected metallicity at $R > 0.3R_{500}$ is $Z_{\rm depro} = 0.242 \pm 0.008$.

The iron mass enclosed within a sphere of radius R, $M_{\rm Fe}(<R)$ (see De Grandi et al. 2004), can be expressed as

$$M_{\rm Fe}(<R) = 4\pi A_{\rm Fe} m_{\rm H} Z_{n,\odot} \int_0^R Z_{\rm depro}(r)\, n_{\rm H}(r)\, r^2 {\rm d}r, \quad (4)$$

where $A_{\rm Fe}$ is the atomic weight of iron, and $m_{\rm H}$ is the atomic unit mass. The hydrogen density $n_{\rm H}$ is derived from the gas density $n_{\rm gas}$ through the usual relation $n_{\rm gas} = (1 + n_{\rm e}/n_{\rm H})\, n_{\rm H} = 2.21 n_{\rm H}$, where $n_{\rm e}$ is the electron density; $n_{\rm gas}$ has been obtained through deprojection as detailed in Ettori et al. (2019) and Eckert et al. (2015).

In Fig. 13 we show all the cumulative iron mass profiles $M_{\rm Fe}(<R)$. At $R = R_{500}$ the iron mass $M_{\rm Fe,500} \equiv M_{\rm Fe}(<R_{500})$ for our systems ranges roughly between $10^{10}\,M_\odot$ and $10^{11}\,M_\odot$. Values of $M_{\rm Fe,500}$ for all clusters in our sample are listed in Table 3.

3.5. Mass-weighted abundances and iron-to-gas-mass ratio

From the deprojected profiles $Z_{\rm depro}$ derived in Sect. 3.4, we can compute mass-weighted abundances within a given radius,

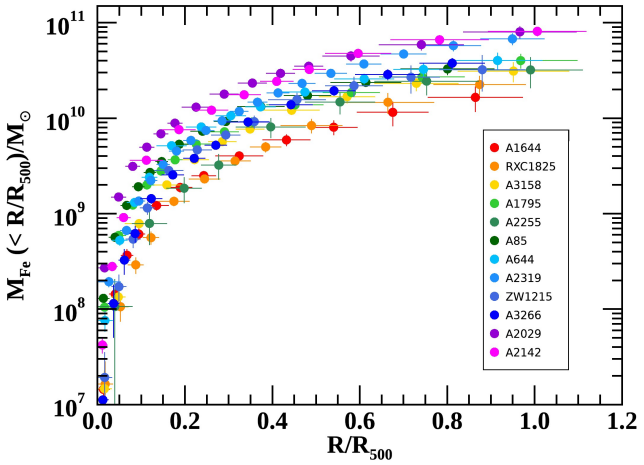

Fig. 13. Iron mass profiles as a function of R/R_{500}. Clusters in the inset are ordered as in Fig. 1.

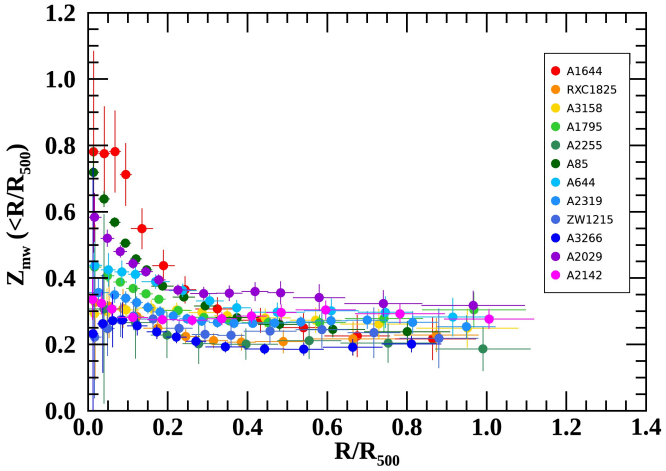

Fig. 14. Profiles of mass-weighted iron abundance within a given radius as a function of R/R_{500} for all systems in our sample. Clusters in the inset are ordered as in Fig. 1.

Table 3. Iron mass $M_{\rm Fe,500}$, gas mass $M_{\rm gas,500}$, hydrostatic mass M_{500}, and stellar mass $M_{\rm star,500}$ enclosed within R_{500} for the X-COP clusters.

Cluster	$M_{\rm Fe,500}$ ($10^{10}\,M_\odot$)	$M_{\rm gas,500}$ ($10^{13}\,M_\odot$)	M_{500} [(a)] ($10^{14}\,M_\odot$)	$M_{\rm star,500}$ ($10^{12}\,M_\odot$)
A1644	$2.01^{+0.57}_{-0.62}$	$4.77^{+0.09}_{-0.10}$	3.48 ± 0.20	–
RXC1825	$2.72^{+0.62}_{-0.59}$	$5.94^{+0.07}_{-0.07}$	4.08 ± 0.13	–
A3158	$3.28^{+0.81}_{-0.79}$	$6.73^{+0.08}_{-0.07}$	4.26 ± 0.18	–
A1795	$4.19^{+0.69}_{-0.75}$	$6.92^{+0.08}_{-0.09}$	4.63 ± 0.14	$3.02^{+0.28}_{-0.28}$
A2255	$3.23^{+1.21}_{-1.03}$	$8.80^{+0.14}_{-0.14}$	5.26 ± 0.34	–
A85	$4.24^{+0.60}_{-0.58}$	$9.17^{+0.09}_{-0.08}$	5.65 ± 0.18	$2.10^{+0.33}_{-0.33}$
A644	$4.40^{+0.89}_{-0.94}$	$8.05^{+0.10}_{-0.11}$	5.66 ± 0.48	$3.70^{+0.38}_{-0.38}$
A2319	$7.14^{+0.95}_{-0.90}$	$14.51^{+0.10}_{-0.10}$	7.31 ± 0.28	$5.11^{+0.46}_{-0.46}$
ZW1215	$3.61^{+1.56}_{-1.77}$	$8.72^{+0.08}_{-0.08}$	7.66 ± 0.52	$3.34^{+0.39}_{-0.39}$
A3266	$4.91^{+0.63}_{-0.57}$	$11.94^{+0.11}_{-0.09}$	8.80 ± 0.57	–
A2029	$8.33^{+1.13}_{-1.14}$	$13.34^{+0.16}_{-0.17}$	8.82 ± 0.35	$6.51^{+0.47}_{-0.47}$
A2142	$8.07^{+0.81}_{-0.84}$	$14.80^{+0.14}_{-0.13}$	8.95 ± 0.26	$6.97^{+0.51}_{-0.51}$

Notes. [(a)]Hydrostatic mass from Ettori et al. (2019).

$Z_{\rm mw}(<R)$. We define $Z_{\rm mw}(<R)$ as

$$Z_{\rm mw}(<R) = \frac{\int_0^R Z_{\rm depro}(r) n_{\rm H}(r) r^2 {\rm d}r}{\int_0^R n_{\rm H}(r) r^2 {\rm d}r}. \quad (5)$$

Mass-weighted abundances for all our systems are reported in Fig. 14. Working with integrated quantities has the advantage that the cumulative functions are more regular than density functions. Consequently the mass-weighted $Z_{\rm mw}(<R)$ profiles appear smoother and more regular than the $Z_{\rm depro}$ profiles. At R_{500} the mean value of the mass-weighted abundance is $Z_{\rm mw,500} = 0.247^{+0.013}_{-0.012}\,Z_\odot$. The total scatter of $Z_{\rm mw,500}$ around the mean value is small, $0.037 \pm 0.008\,Z_\odot$, which is ~15%. This finding has significant implications that will be discussed in Sect. 6.

3.6. Scaling relations

Having derived robust estimates of Fe masses within R_{500} for the first time, we investigate scaling relations between $M_{\rm Fe,500}$ and other ICM observables, namely the gas mass $M_{\rm gas,500}$ and the total mass M_{500} enclosed within R_{500}; $M_{\rm gas,500}$ is obtained as usual by integrating $n_{\rm gas}$, while the total masses M_{500} have been derived in Ettori et al. (2019), under the assumption of hydrostatic equilibrium. Values for $M_{\rm gas,500}$ and M_{500} for our sample are reported in Table 3.

In Fig. 15 we show $M_{\rm Fe,500}$ as a function of $M_{\rm gas,500}$ (top panel) and M_{500} (bottom panel) for all clusters in our sample. In both cases, we observe a very clear correlation without any difference between cool-core and non-cool-core clusters (marked with blue and red circles, respectively).

We model the relations with a standard power law,

$$\frac{M_{\rm Fe,500}}{\hat{M}_{\rm Fe,500}} = 10^{\alpha} \left(\frac{M_{\rm X}}{\hat{M}_{\rm X}}\right)^{\beta}, \quad (6)$$

where $M_{\rm X} = M_{\rm gas,500}$ or $M_{\rm X} = M_{500}$. We center the relation on the pivot values $\hat{M}_{\rm Fe,500} = 4.24 \times 10^{10}\,M_\odot$, $\hat{M}_{\rm gas,500} = 8.80 \times 10^{13}\,M_\odot$, and $\hat{M}_{500} = 5.66 \times 10^{14}\,M_\odot$ set at the median of the distributions of $M_{\rm Fe,500}$, $M_{\rm gas,500}$, and M_{500} respectively.

We fit our data by performing a linear regression analysis in the log-log space using the IDL package *linmix_err.pro* by Kelly (2007), based on Bayesian inference, which treats measurement errors in both variables and allows an intrinsic scatter around the regression line. Though we are aware of the unreliability of intrinsic scatter estimates for our data (see Sect. 2.3.1), we decided to include this quantity in our regression analysis, to account for some intrinsic spread. Conscious of the possible underestimation of this quantity (see Sect. 2.3.1) we will, conservatively, make use of the total scatter as a measurement of the dispersion of the data around the best fit value. The best-fitting values and the total scatter are listed in the first row of Table 4; in both cases the slope value is close to 1, meaning that the relation is consistent with being linear. No segregation between cool-cores and non-cool-cores is found in either relation. This suggests a lack of any causal connection between the mechanism responsible for the formation of a cool core and the overall enrichment of the ICM.

Focusing on the relation $M_{\rm Fe,500}-M_{\rm gas,500}$, the total scatter of the data around the best fit power-law relation is $\sigma = (15.0 \pm 3.2)\%$. It is worth pointing out that the tight relation between $M_{\rm Fe,500}$ and $M_{\rm gas,500}$ is not connected to the flatness of the Z profiles, but to the small scatter in the average abundance profile.

If the scatter in the average abundance profile were high, clusters with similar $M_{\rm gas,500}$ could have different $M_{\rm Fe,500}$, leading to a large scatter in the scaling relation. Scatter for the

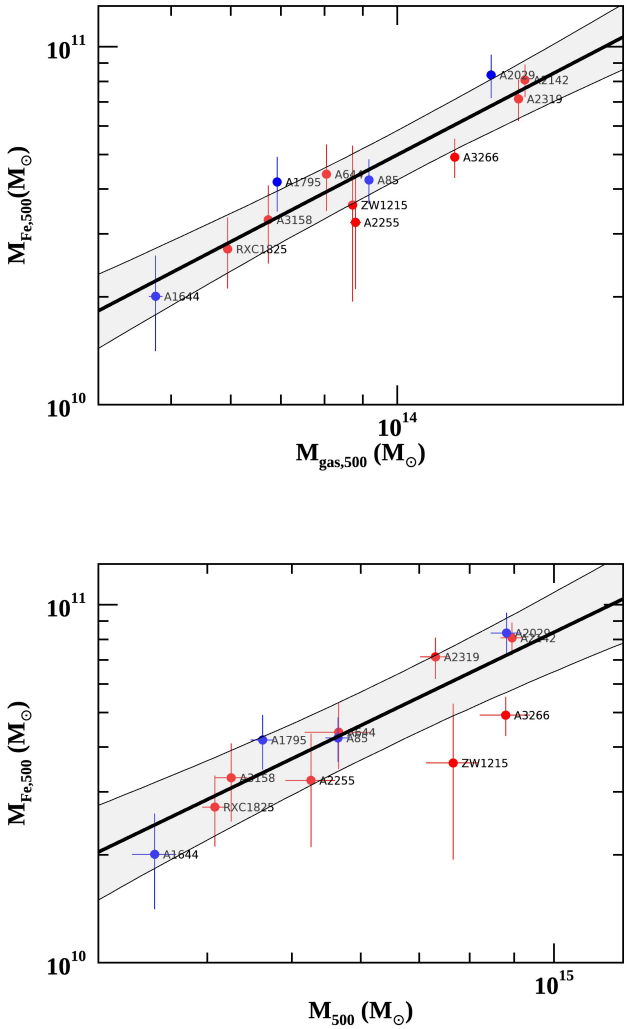

Fig. 15. *Top panel*: $M_{\rm Fe}$ within R_{500} vs. $M_{\rm gas}$ within R_{500}; *bottom panel*: $M_{\rm Fe}$ within R_{500} vs. $M_{\rm tot}$ within R_{500}. Blue and red mark cool-core and non-cool-core systems, respectively. Black lines are the best fits and shaded areas indicate the 1σ confidence regions, including intrinsic scatters.

$M_{\rm Fe,500}-M_{500}$ relation is slightly larger $\sigma = (22.3 \pm 4.7)\%$ than for the $M_{\rm Fe,500}-M_{\rm gas,500}$ relation.

Scatter values reported in Table 4 (see also Sect. 4.2) are very similar and reciprocally consistent within $1-2\sigma$. Nonetheless, the scatter for the $M_{\rm Fe,500}$ versus $M_{\rm gas,500}$ relation is the lowest and best constrained, making the correlation between gas and Fe mass the tightest amongst the scaling relations we have investigated. Henceforth, in the admittedly limited mass range covered by X-COP, $M_{\rm gas,500}$ can be used as a proxy for $M_{\rm Fe,500}$ and the latter can be estimated for any massive system from the former with an uncertainty of ∼15%. Of course, this is just another way of saying that our distribution of $Z_{\rm mw,500}$ measurements features a small scatter.

4. Optical data

The X-COP sample has seven clusters in common with the Multi Epoch Nearby Cluster Survey (MENeaCS). For these, we have deep optical imaging data, taken with the purpose of performing a weak-lensing analysis of the clusters (Herbonnet et al., in prep.) and their constituent galaxies (Sifón et al. 2018), and to study transient phenomena (particularly intra-cluster supernovae, Sand et al. 2011). van der Burg et al. (2015, hereafter vdB15) combined the original *g*- and *r*-band imaging with additional *u*- and *i*-band imaging and performed a study of the stellar mass content of these clusters. The optical analysis presented in this paper is based on their data set. We provide a brief summary here, and for more details we refer to vdB15.

The basis of the study are *g*- and *r*-band imaging data taken with MegaCam at the Canada-France-Hawaii Telescope (CFHT). Imaging data in the *u*- and *i*-bands were also acquired using the Wide Field Camera at the *Isaac Newton* Telescope. Sources are detected in the *r*-band, and aperture fluxes are measured in each filter stack from Point-Spread-Function-homogenised images using Gaussian weight functions. Typical 5σ aperture magnitude limits are 24.3, 24.8, 24.2, and 23.3 in the *ugri*-filters, respectively. The aperture fluxes form the basis for the spectral energy distribuion (SED) fitting, from which stellar mass to light ratios are estimated for each galaxy. The stellar population libraries from Bruzual & Charlot (2003) are used to model the SED. The star formation history is parametrized as $\mathrm{SFR} \propto e^{-t/\tau}$, where the time-scale τ is allowed to range between 10 Myr and 10 Gyr. We assume a Chabrier (2003) IMF, solar metallicity, and the Calzetti et al. (2000) dust extinction law. Using the total flux measured in the *r*-band, luminosities are converted into stellar masses. Initially each galaxy is assumed to be part of the cluster (to set the luminosity distance), and in a second step a statistical subtraction of fore- and background interlopers is performed using multi-band photometry available in the Cosmic Evolution Survey (COSMOS) field (Muzzin et al. 2013).

4.1. Stellar mass profiles

To accumulate stellar mass radial profiles, we center clusters on the X-ray centroids, as opposed to the BCGs in vdB15. The impact of how these profiles are cantered is generally small (the centers differ by at most 2% of R_{500}).

For the profiles we consider all galaxies with stellar masses in excess of $10^9\,M_\odot$[4]. We point out that the BCG is included in our mass computation, while the intra-cluster light (ICL) is not. For each galaxy we consider the statistical uncertainty on the estimated stellar masses, using the measured flux uncertainties. For this we bootstrap the flux measurements within their uncertainties, and perform 100 perturbations for each galaxy. When we combine the stellar masses of all galaxies, we account for these uncertainties on individual stellar masses. We account for and subtract the contribution of fore- and background galaxy interlopers as described in vdB15. Briefly, we make use of the COSMOS field and, for consistency, only consider data taken in the *ugri*-bands. The subtraction procedure introduces a Poisson noise term (statistical uncertainty), but also a systematic uncertainty related to how representative the reference field is of the true cluster field background. Since the COSMOS field is relatively small, there is a substantial uncertainty due to field-to-field (often called "cosmic") variance. We estimate this uncertainty using Moster et al. (2011) and we tested in vdB15

[4] We note that these galaxies make up the vast majority of the stellar mass component in a typical cluster galaxy population. To showcase this, we integrate the SMF, which was measured in van der Burg et al. (2018), to have a Schechter (1976) form with a low-mass slope of $\alpha \approx -1$ and a characteristic mass of $M^* = 10^{10.8}\,M_\odot$, down to infinitely low masses. If, instead, we only consider galaxies with $M_\star > 10^9\,M_\odot$, this would account for more than 98% of the stellar mass that is present in the entire population.

Table 4. Best fit parameters for scaling relations.

	$M_{\rm gas,500}$			M_{500}		
	α	β	$\sigma(\%)$	α	β	$\sigma(\%)$
$M_{\rm Fe,500}$	$0.01^{+0.03}_{-0.03}$	$1.10^{+0.20}_{-0.19}$	15.0 ± 3.2	$0.005^{+0.04}_{-0.04}$	$1.18^{+0.27}_{-0.26}$	22.3 ± 4.7
$M_{\rm star,500}$	$-0.03^{+0.08}_{-0.08}$	$1.11^{+0.60}_{-0.52}$	20.7 ± 5.9	$-0.07^{+0.08}_{-0.09}$	$1.39^{+0.68}_{-0.75}$	22.0 ± 6.1

Notes. Equations (6) and (7) define α and β, and σ is the intrinsic scatter around the best fit line in the log-log space.

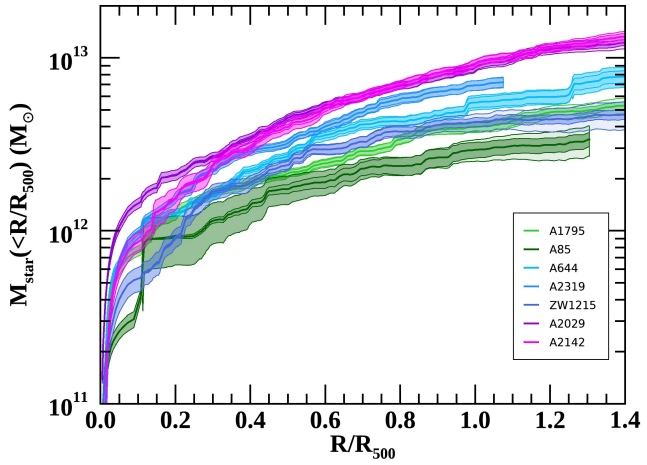

Fig. 16. Cumulative stellar mass profiles. Dark shades are statistical uncertainties, faint shades are cosmic variance uncertainties.

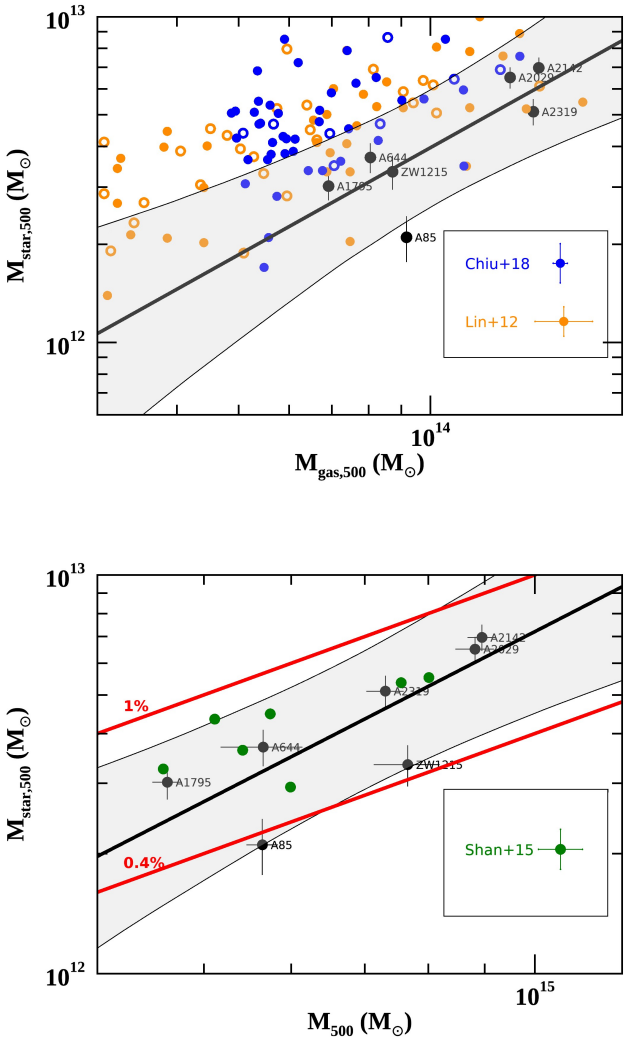

Fig. 17. Stellar mass within R_{500} as a function of the gas mass (*top panel*) and total mass (*bottom panel*) within R_{500} for the X-COP sample (black circles). Black lines are the best fits and shaded areas indicate the 1σ confidence regions, including intrinsic scatters. *Top panel*: we plot for comparison data from Chiu et al. (2018) (blue circles) and Lin et al. (2012) (yellow circles) samples. Low redshift clusters of both samples are marked with open circles. *Bottom panel*: we report data by Shan et al. (2015) (green circles). We omit error bars for these samples to avoid overcrowding the plot: typical error bars are shown in the legend. Red lines in the *bottom panel* represent two fixed levels (0.4% and 1%) of stellar fraction.

that the estimated variance is consistent with the scatter obtained when considering the four spatially-independent CFHT Legacy Deep fields used for the background subtraction. Stellar masses are integrated to derive the cumulative stellar mass profile for each cluster, $M_{\rm star}(<R)$. Profiles for $M_{\rm star}(<R)$ are shown in Fig. 16. The statistical uncertainty due to flux errors on individual galaxies and the systematic uncertainty on the profile due to field-to-field variance are shown separately in the cumulative stellar-mass profiles (see Fig. 16). It is worth pointing out that the cumulative stellar masses plotted in Fig. 16 are derived by integrating within a projected radius. To derive the stellar mass enclosed within a sphere of radius R_{500}, we perform a correction assuming a generalized Navarro-Frenk-White (gNFW) profile for the galaxy distribution (see vdB15), with a concentration parameter $c = 0.72$ and a slope $\alpha = 1.64$. We find that 75% of the mass obtained integrating along the line of sight lies within the sphere with radius R_{500}. We therefore multiply the stellar mass estimates by a factor of 0.75. The values of $M_{\rm star,500}$ corrected in this manner are listed in the last column of Table 3.

4.2. Scaling relations for stellar masses and stellar mass uncertainties

As previously done for $M_{\rm Fe,500}$, we investigate scaling relations for $M_{\rm star,500}$. In Fig. 17 we plot $M_{\rm star,500}$ versus $M_{\rm gas,500}$ (top panel) and $M_{\rm star,500}$ versus M_{500} (bottom panel). We fit the relations with a standard power law,

$$\frac{M_{\rm star,500}}{\hat{M}_{\rm star,500}} = 10^{\alpha} \left(\frac{M_{\rm X}}{\hat{M}_{\rm X}}\right)^{\beta}, \quad (7)$$

for both $M_{\rm X} = M_{\rm gas,500}$ or $M_{\rm X} = M_{500}$. We adopt the same pivot values for $\hat{M}_{\rm gas,500}$ and $\hat{M}_{500}$ used in Sect. 3.6 and set $\hat{M}_{\rm star,500} = 3.70 \times 10^{12}\,M_\odot$, at the median of the $M_{\rm star,500}$ distribution. Best-fitting values are listed in the second row of Table 4. In both cases the slope value is compatible with unity, meaning that the relation is consistent with being linear.

Table 5. Parameters for scaling relation $M_{\rm star,500}$ vs. $M_{\rm gas,500}$ (top panel in Fig. 17) for the samples by Chiu et al. (2018) and Lin et al. (2012).

Sample	Pivot masses ($M_\odot$)		Best-fit-values		Intrinsic scatter
	$\hat{M}_{\rm gas,500}$	$\hat{M}_{\rm star,500}$	α	β	$\sigma(\%)$
Chiu+2018	6.13×10^{13}	4.69×10^{12}	-0.011 ± 0.016	0.48 ± 0.12	32.3 ± 2.3
Lin+2012	4.51×10^{13}	4.71×10^{12}	-0.014 ± 0.015	0.64 ± 0.04	36.8 ± 2.7

Notes. We adopt $\hat{M}_{\rm gas,500}$ and $\hat{M}_{\rm star,500}$ as the pivot masses; α, β, and σ are the best fit parameters as in Table 4.

Interestingly, the best fit relations for $M_{\rm star,500}-M_{\rm gas,500}$ and $M_{\rm star,500}-M_{500}$ feature modest scatters, comparable to those for the corresponding relations for $M_{\rm Fe,500}$ plotted in Fig. 15. These findings suggest that processes underpinning metal enrichment proceed at essentially the same pace in all our objects.

Remarkably, A85 is an outlier in both the relations plotted in Fig. 17, having a rather low stellar content for its mass. This is in agreement with the findings presented by Shan et al. (2015, see their Fig. 7), where A85 features a low stellar fraction (comparable to ours), but at odds with estimates from Kravtsov et al. (2018) who find $M_{\rm star,500}\sim7-8\times10^{12}\,M_\odot$, which is extremely high and would cause A85 to be an outlier on the opposite side of our scaling relations. The reason for this discrepancy is not clear: as we will discuss below, comparisons between different stellar measurements are quite delicate and should be treated carefully. The impact of the eventual underestimation of the stellar mass of A85 will be discussed in Sect. 5.1.

Comparing stellar mass measurements from different cluster samples requires some caution, since stellar masses are estimated via complicated models, which include many physical processes and are based on different assumptions. For example, assuming a Salpeter (1955)-diet initial mass function (IMF) in deriving stellar masses from luminosities, provides values that exceed by a factor of ∼2 those obtained using a Chabrier (2003) mass function. Similarly, different choices for example, for stellar population synthesis (SPS), stellar evolution, star formation history (SFH), dust attenuation, and stellar mass function are sources of possible bias and can induce significant systematics on the final stellar mass estimation. Comparison between samples where different assumptions have been made can be used to provide an estimate of systematics affecting stellar mass measurements.

We compare our $M_{\rm star,500}$ versus $M_{\rm gas,500}$ measurements with those reported in Chiu et al. (2018) and Lin et al. (2012), hereafter C18 and L12. We omit error bars for these two samples to avoid overcrowding the plot. We do not plot results reported in Gonzalez et al. (2013) and Kravtsov et al. (2018) since their samples have little overlap with the mass range of XCOP clusters; it is worth noting that the stellar masses in both these works are higher than those plotted in Fig. 17. This is due to the different assumptions adopted to estimate stellar masses: both these works adopt the mass-to-light ratios calibrated by Cappellari et al. (2006), which are dynamical mass-to-light ratios, and might be biased high by up to 30% (Leauthaud et al. 2012). C18 present stellar masses for a sample of 91 clusters; we restricted the comparison to clusters with M_{500} in a range similar to ours, namely: $4.5\times10^{14}-10^{15}\,M_\odot$. C18 clusters cover a wide range in redshift ($0.25<z<1.25$). To highlight possible differences, we show low-redshift systems, $z<0.4$, as open circles. The L12 sample includes 94 clusters that span the redshift range 0−0.6; they assume a Kroupa (2001) IMF, consequently stellar masses for their sample (yellow circles in Fig. 17) have been rescaled by a factor of 0.76 (see C18), to bring them to Chabrier (2003) values. Low-redshift clusters ($z<0.1$) are marked with open circles. Unfortunately, L12 do not report M_{500} for their systems, so we do not apply any restriction to the masses of their sample. The bulk of both samples has stellar masses that are higher than ours, by a factor of 1.5−3, but there is some overlap between our data and theirs. Differences seem to diminish at higher gas (and stellar) masses. Discrepancies are still present when we restrict the comparison to the low-redshift subsamples (open circles), meaning they cannot be ascribed to the high redshift objects, nor be interpreted in any evolutionary framework. Both C18 and L12 data sets exhibit a significant scatter; to quantify differences between the samples, we fit C18 and L12 using our power law model (Eq. (7)) and compare best fit relations. L12 do not explicitly report error bars for their data, so we assign each point of their sample a typical error bar (see their Fig. 3) of 0.045 dex for $M_{\rm gas,500}$ and for $M_{\rm star,500}$ in logarithmic scale. Best fit values for C18 and L12, along with the total scatters and the assigned values of pivot masses, are reported in Table 5.

We note that the best fit for C18 does not match values reported by the authors in their paper, because we only included massive clusters.

Since differences between samples seem to diminish at higher masses, we choose to evaluate stellar mass discrepancies at three reference gas mass values: $M_{\rm gas,500}=[6.5,8.8,15]\times10^{13}\,M_\odot$, which are approximately the minimum, the median, and the maximum $M_{\rm gas,500}$ values for our sample. At these three reference gas masses, C18 (L12) estimates exceed ours by [90%, 58%, 12%] ([78%, 56%, 24%]), respectively. Values at the minimum reference gas mass should be taken with caution both for C18 and L12: data from C18 have been restricted to massive systems, with $M_{500}>4.5\times10^{14}\,M_\odot$, approximately the mass of A1795. However, as shown in the top panel of Fig. 17, many C18 points feature $M_{\rm gas,500}$ values that are significantly lower than that found for A1795. Indeed, while measuring gas masses is quite simple and straightforward, estimates of M_{500} are subject to bias, depending on the adopted method. The low gas-mass values for many C18 data instill some doubt that bias could be present and that we are also plotting (and comparing) clusters whose mass is below the $4.5\times10^{14}\,M_\odot$ threshold. Since the stellar fraction is well known to increase when M_{500} decreases, including these (possibly) less massive systems could increase the discrepancy with our measurements. Moreover, the L12 sample includes low-mass systems, since we could not apply any selection on mass, and the discrepancy in the minimum reference gas mass may also be overestimated in this case. As a consequence, at the minimum reference gas mass, the discrepancy should be regarded with some caution, while values at the median and maximum reference gas masses are more robust. Globally, assuming a systematic discrepancy of 50%−60% between our measurements and those reported in C18 and L12 seems reasonable.

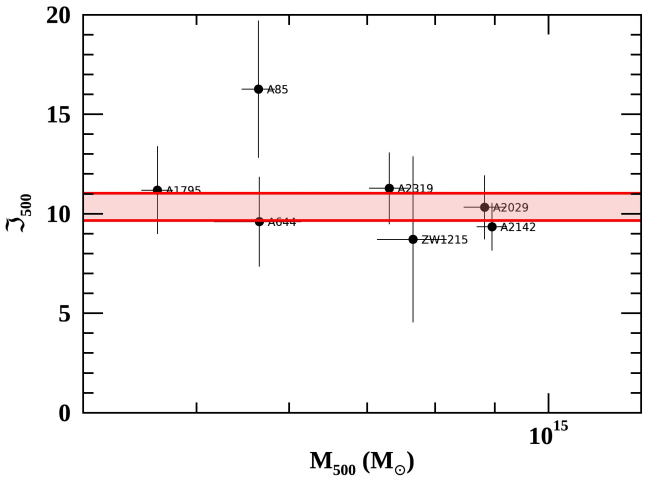

Fig. 18. Iron share for the clusters of our sample. The red band is the 68% confidence interval around the mean.

Let us now consider the $M_{\rm star,500}-M_{500}$. In the bottom panel of Fig. 17 we show two lines corresponding to two levels of stellar fraction, namely $M_{\rm star,500}/M_{500}=1\%$ and 0.4%. All our clusters are within this range (except for A85 which is slightly below). Interestingly, our stellar fractions are in agreement with those reported by Leauthaud et al. (2012), and almost one order of magnitude larger than those found by (Girelli et al. 2020). As pointed out in Leauthaud et al. (2012), stellar fraction estimates are subject to substantial systematic errors, similar to those we have estimated for the $M_{\rm star,500}-M_{\rm gas,500}$ relation.

5. Combining iron in the ICM and stars

Having derived estimates of the iron mass in the ICM for all our objects and of stellar masses for a sizeable fraction of them, we will now merge this information to address two fundamental questions. (1) How is iron shared between stars and the ICM? (2) Is the iron mass measured in our clusters consistent with expectations based on SN rates? We will start by addressing the first of these questions.

5.1. Iron share

The iron share $\mathfrak{I}$ is defined as the ratio between iron diffused into the ICM and locked into stars. We shall evaluate $\mathfrak{I}$ within R_{500}, that is,

$$\mathfrak{I}_{500}=\frac{M_{\rm Fe,500}}{M^{\rm star}_{\rm Fe,500}}. \tag{8}$$

The value $\mathfrak{I}_{500}$ quantifies how stars and intra-cluster plasma share the total cluster iron content. The iron diffused in the ICM, $M_{\rm Fe,500}$, has been derived from Eq. (4), while $M^{\rm star}_{\rm Fe,500}$ is determined from our stellar mass estimates (see Sect. 4.1), assuming on average solar metallicity, which is $Z^{\rm star}\simeq Z_{\rm m,\odot}$ (see Renzini & Andreon 2014)[5], where $Z_{\rm m,\odot}$ is the solar iron abundance by mass, $Z_{\rm m,\odot}=A_{\rm Fe}Z_{\rm n,\odot}X$, $A_{\rm Fe}$ is the atomic weight of iron (see Eq. (4)), and X the hydrogen mass fraction. Following the standard approach (see Renzini & Andreon 2014; Maoz & Graur 2017) we adopt the Asplund et al. (2009) values so that $Z_{\rm n,\odot}=3.16\times10^{-5}$ and $X=0.7$ provide $Z_{\rm m,\odot}=0.00124$. For brevity we will shorten $Z_{\rm m,\odot}$ using the standard notation $Z_\odot$. The iron share can then be rewritten as

$$\mathfrak{I}_{500}\equiv\frac{M_{\rm Fe,500}}{Z_\odot M_{\rm star,500}}. \tag{9}$$

Iron share measurements are plotted in Fig. 18 and listed in Table 6. As we can see, values vary between 8 and 11, reaching a maximum of about 16 for A85. The mean value for $\mathfrak{I}_{500}$ is 10.35 ± 0.69 with a total scatter $\sigma=1.48$, (i.e., ∼14%). In Sect. 4.2 we noted that A85 is particularly poor in stellar mass and that a broad range of values is found in the literature, some consistent with ours (Shan et al. 2015), others much larger (Kravtsov et al. 2018). A higher value for the A85 stellar mass could reconcile its iron share with the mean value of the sample, however, if we were to assume the value reported by Kravtsov et al. (2018), the iron-share for A85 would stand out in Fig. 18 on the opposite side of the red band.

Our measurements show that, on average, the iron locked in stars contributes only 10% to the total iron in clusters. It would therefore seem that feedback effects, independently of where and when they took place, are very efficient, in the sense that 90% of the iron in clusters, while originating from the stellar component, has managed to escape galaxies and pollute the hot gas. Thus, we can visualize stars as "iron factories" producing metal only to hand it over to the ICM.

Table 6. Iron shares $\mathfrak{I}_{500}$ and effective iron yields $\mathcal{Y}_{\rm Fe,\odot}$.

Cluster	$\mathfrak{I}_{500}$	$\mathcal{Y}_{\rm Fe,\odot}$
A1795	11.18 ± 2.18	7.07 ± 1.27
A85	16.26 ± 3.43	10.01 ± 1.99
A644	9.60 ± 2.23	6.15 ± 1.29
A2319	11.28 ± 1.78	7.12 ± 1.03
ZW1215	8.71 ± 4.14	5.63 ± 2.40
A2029	10.32 ± 1.59	6.57 ± 0.92
A2142	9.34 ± 1.17	6.00 ± 0.68

5.2. Effective iron yield

The effective iron yield $\mathcal{Y}_{\rm Fe}$ provides the efficiency with which stars produce iron in clusters. It is defined as the total Fe mass divided by the mass of the gas that went into stars:

$$\mathcal{Y}_{\rm Fe}=\frac{M^{\rm star}_{\rm Fe,500}+M_{\rm Fe,500}}{M_{\rm star,500}(0)}, \tag{10}$$

where $M_{\rm star,500}(0)$ is the mass of gas that went into stars whose present mass is reduced to $M_{\rm star,500}$ by the mass return from stellar mass loss, which is $M_{\rm star,500}(0)=r_oM_{\rm star,500}$, where r_o is the return factor. Following Renzini & Andreon (2014) and Maraston (2005), we shall assume $r_o=1/0.58$. By dividing $\mathcal{Y}_{\rm Fe}$ by the Fe solar abundance, $Z_\odot$ (Asplund et al. 2009), we can express our result in solar units as

$$\mathcal{Y}_{\rm Fe,\odot}\equiv\mathcal{Y}_{\rm Fe}/Z_\odot. \tag{11}$$

The effective iron yield in solar units for the objects in our sample is plotted in Fig. 19, and values are reported in Table 6. The

[5] Stars have a range of metallicity depending on where and when they were formed, and there is a known relationship between galaxy stellar mass and metallicity. On average the stellar metallicities (weighted by stellar mass) of cluster galaxies should be close to solar and perhaps even super-solar, as shown, for example, in Maoz et al. (2010) and references therein. This would, slightly, further raise the total iron mass in clusters, and increase the discrepancy with the expectations from SNe (see Sect. 5.2).

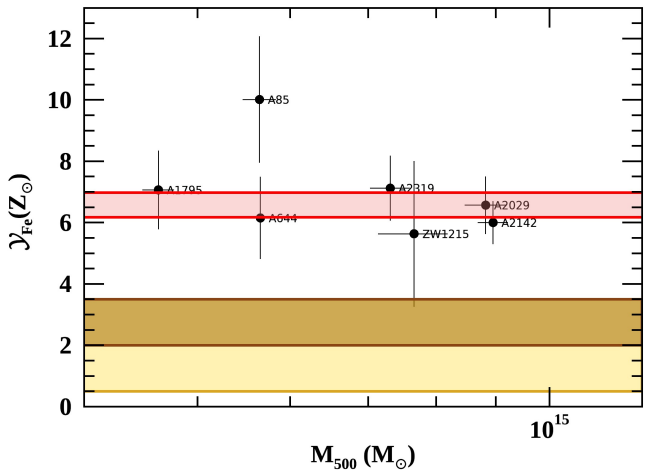

Fig. 19. Effective iron yield for the clusters of the sample. The red band is the 68% confidence interval around the mean. The yellow band shows the expected value computed through the SN yields derived from Maoz & Graur (2017) and Renzini & Andreon (2014) (see text for details); the brown band represents the expected value derived assuming a higher SNIa explosion rate in galaxy clusters than in the field, following Maoz & Graur (2017).

mean value for $\mathcal{Y}_{Fe}$ is 6.58 ± 0.40 and the total scatter is 0.87 (i.e., ∼13%).

Estimates of the expected $\mathcal{Y}_{Fe}$ have been derived by several authors, such as Renzini & Andreon (2014) and Maoz & Graur (2017). It is computed as the product of the Fe mass produced by a SN explosion, y, and the number of SN events produced per unit mass of gas turned into stars, k. Both contributions from Ia and core-collapse (CC) SN are considered as they are of the same order. Thus, $\mathcal{Y}_{Fe}$ can be written as

$$\mathcal{Y}_{Fe} = y_{Ia} \cdot k_{Ia} + y_{CC} \cdot k_{CC}, \quad (12)$$

where Ia and CC subscripts refer to the two different SN types. For Ia, following Maoz & Graur (2017), we assume $y_{Ia} = 0.7\,M_\odot$ and $k_{Ia} = 1.3 \times 10^{-3}\,M_\odot^{-1}$. Renzini & Andreon (2014) (see also Greggio & Renzini 2011), suggest that k_{Ia} could be as high as $2.5 \times 10^{-3}\,M_\odot^{-1}$. However, this high value is based on early measurements of the SN Ia rate in local Sb galaxies requiring some uncertain assumptions about the star formation histories of such galaxies. For CC SN, following Maoz & Graur (2017), we assume $y_{CC} = 0.074\,M_\odot$ and $k_{CC} = 1.0 \times 10^{-2}\,M_\odot^{-1}$. Substituting the above values in Eq. (12) and dividing by the solar abundance, we get

$$\mathcal{Y}_{Fe,\odot} = 0.93\,Z_\odot. \quad (13)$$

Given the simplified fashion in which this calculation was carried out, by assuming one average value for Ia and CC yields and SN rates, a rather large uncertainty should be associated with $\mathcal{Y}_{Fe}$. Although the factor of 2 error suggested in Renzini & Andreon (2014) might be too generous (see discussion above) we shall nonetheless assume it as part of our conservative approach in interpreting our measurements.

The expected effective Fe yield is shown in Fig. 19 as a yellow shaded region. Remarkably, measured values, which range from 6 to 10, are much higher, even when assuming a highly conservative factor of 2 uncertainty in the expected effective Fe yield.

Recently, it has been shown (see Maoz & Graur 2017; Friedmann & Maoz 2018) that type Ia SN explosions in galaxy clusters are more frequent than in the field. If, following Maoz & Graur (2017), we assume a SN_{Ia} rate per unit mass of $k_{Ia} = (5.4 \pm 2.3) \times 10^{-3}\,M_\odot^{-1}$, we derive $\mathcal{Y}_{Fe,\odot} = 2.6^{+0.9}_{-0.6}\,Z_\odot$ (gold shaded region in Fig. 19). This revised SN rate brings the expected effective yield closer but it is still well below the measured ones.

Table 7. Systematic uncertainties on iron share $\mathfrak{I}_{500}$ and effective iron yield $\mathcal{Y}_{Fe,\odot}$.

Variable	Sys. on var.	Direction	$\downarrow \mathfrak{I}_{500}$	$\downarrow \mathcal{Y}_{Fe,\odot}$
M_{star}	60%	↑	60%	50%
r_o	20%	↓	–	20%
ICL	50%	↑	50%	40%
$M_{Fe,500}$	15%	↓	15%	10%
$\mathcal{R}$	50(20)%	↑	50(20)%	40(15)%

Notes. The down facing arrows placed next to $\mathfrak{I}_{500}$ and $\mathcal{Y}_{Fe,\odot}$ indicate that all corrections have the effect of reducing estimates for these two variables.

5.3. Uncertainties on iron share and yield

Here we evaluate possible uncertainties in the calculation of the iron share and yield. We consider three distinct sources of errors: (1) those associated to the stellar mass estimates; (2) those associated to ICM mass estimates, and (3) those associated to comparing stellar and ICM measurements within a given radius. A summary of systematics is presented in Table 7. We list sources of systematics, the assumed percentual variation, whether it acts as an increase or a decrease, and the impact on iron share and yield.

Let us start with stellar mass estimates. We identify the following key points in the calculation that goes from the magnitudes of individual galaxies to the estimate of the iron share and yield: (1) the stellar luminosities of individual galaxies are converted into stellar masses and then summed up to derive the cluster stellar mass, M_{star}; (2) the stellar mass is converted into iron locked in stars, $M^{star}_{Fe,500}$, by assuming a certain value of the metallicity, and (3) an estimate of the stellar mass loss is used to compute the mass of gas that went into stars, $M_{star}(0)$. Systematic errors in the stellar mass estimates are discussed extensively in Sect. 4.2. Although we lack a clear understanding of the origin of different estimates of M_{star}, our comparison shows that we may be underestimating stellar masses by as much as 60%. If we correct for this systematic, we derive a reduction of the mean iron share by the same percentage and of the mean effective iron yield by about 50% (see also Table 7). The factor r_o, introduced to account for the stellar mass loss (see Sect. 5.2), is a further source of systematic uncertainty. The value we have adopted, $r_o = 1/0.58$, has been derived for a top heavy IMF; for a Salpeter IMF it would go down to 1/0.70, implying that we may be overestimating r_o by as much as 20%. Correcting for this systematic leads to a reduction of the mean effective iron yield by the same percentage.

Another source of uncertainty in our estimate of the iron share and yield is the contribution of ICL to the total cluster light. There have been several measurements of ICL on individual systems and a few on low redshift samples. We shall consider two studies based on the stacking of large samples of survey data, Sloan Digital Sky Survey (SDSS) and Dark Energy Survey (DES) respectively. In the first, Zibetti et al. (2005) found that the ICL contribution to the total optical emission in a cluster

is about 10%, while in the second, Zhang et al. (2019) derived a much larger value, about 50%[6]. An evaluation of which of these two measures is more robust is well beyond the scope of this paper. We adopt a conservative approach and assume the larger contribution of 50%. This results in a reduction of the mean iron share by the same percentage and of the mean effective iron yield by about 40% (see also Table 7).

Let us now turn to uncertainties on ICM Fe mass estimates. As evident from Eq. (4), the estimate on the Fe mass is based on two quantities: the gas mass and the Fe abundance. The former is one of the better measured in X-rays, with systematics less than a few percent (see discussion in Eckert et al. 2019) while the latter is the most challenging (e.g., Molendi et al. 2016). Thus, uncertainties on the ICM Fe mass will be dominated by uncertainties on the Fe abundance. An estimate of the systematic uncertainty on the Fe abundance has been presented in Sect. 3.2. Here we take the 15% uncertainty, conservatively assume that it takes the form of a reduction on the iron mass, and apply it directly to our ICM Fe mass estimates[7]. We find a reduction by the same amount, which is 15% of the iron share and of $\simeq$10% of the effective iron yield (see also Table 7).

Finally, let us consider systematics associated to the combination of our ICM and stellar mass estimates. Our measurements are integrated out to R_{500} both for stellar and ICM masses. When computing iron shares and effective yields from this data, the underlying assumption is that the radial distribution of stellar and ICM masses in cluster outskirts do not differ substantially. This is an approximation as the ICM is more compactly distributed than the stellar matter. From our data we estimate that the ratio of mass within R_{200} over mass within R_{500} is respectively $\sim$1.5 for the ICM and $\sim$1.8 for the stellar component, leading to a value of $\mathcal{R}$ defined as $\mathcal{R} \equiv (M_{\rm star,200}/M_{\rm star,500})/(M_{\rm gas,200}/M_{\rm gas,500})$ of 1.2. Thus, under the assumption that metal abundances in the ICM continue to remain constant beyond R_{500}, iron shares and effective yields would decrease by about $\sim$1.2 and 1.15 respectively, if we were to extend our measurements out to R_{200}. No data is available to correct masses beyond R_{200}, and we consider two different possibilities. If the stellar mass profile follows the dark matter profile, then, given that $f_{\rm gas} = M_{\rm gas}/M_{\rm tot}$ predicted by simulations at large radii is either constant or increasing with radius (Kravtsov et al. 2005; Planelles et al. 2013), the numbers derived above apply. If, conversely, we adopt a more conservative approach and extrapolate linearly the ratios for $M_{\rm star}$ and $M_{\rm gas}$ discussed above out to a very large radius of $1.5\,R_{200}$, we come up with a value of $\mathcal{R}$ of about 1.5, leading to a drop in iron shares and effective yields of $\sim$1.5 and 1.4 respectively. In Table 7 we report both estimates of the systematics, with the less conservative in parenthesis.

Having evaluated the various systematics affecting our estimates of iron shares and effective yields, we now step back and finish this section with some general comments. Of all our systematics, those affecting ICM iron mass estimates are the smallest. This is easily understandable: albeit challenging, X-ray measurements are fairly straightforward and have been carried out on a set of observations specifically designed for cluster outskirts. Conversely, stellar masses are characterized by the largest systematics. This is not surprising; $M_{\rm star}$ estimates are observationally challenging (e.g., intra-cluster light) and at the same time rely on sophisticated and non-unique modeling. The difference in radial profiles between stellar mass and the ICM may also lead to substantial systematics, particularly if the stellar mass does not follow the dark matter distribution. Inspection of Table 7 shows that individual sources of errors are insufficient to reconcile our measurements of the iron yield with expectations based on SN rates. However, if we combine all systematics affecting stellar mass estimates we end up with a value of $\sim$3.8 for $\mathcal{Y}_{\rm Fe,\odot}$, which is relatively close to the one expected if we assume the $\rm SN_{Ia}$ rate proposed by Maoz & Graur (2017) (see also Fig. 19). Clearly, combining the three systematics on stellar mass with others, particularly those on $\mathcal{R}$, will reconcile the mean measured yield with the expected one.

6. Discussion

We have analyzed the first representative sample of massive clusters for which iron abundance has been measured out to R_{500}. Building on significant work presented in previous X-COP papers (Ghirardini et al. 2018, 2019; Ettori et al. 2019; Eckert et al. 2019), we have taken a closer look at the method commonly employed to measure Fe abundances, identifying and correcting a major systematic error plaguing measurements in cluster outskirts.

Our analysis sheds light on the complicated interplay between instrumental background, effective area calibration, the Fe Kα line, and the Fe L-shell emission from which iron abundance measurements emerge. As expounded in Sect. 2.3.1, Fe Kα measurements are far less prone to systematic uncertainties than L-shell ones, which implies they are amenable to statistical treatments leading to the estimation of robust sample properties such as mean and scatter. Furthermore, as discussed in Sect. 3.2, working with hot systems ensures that fitting with a one temperature and one abundance model spectra that, to some extent, must be multi-temperature and multi-abundance, yields only modest systematic errors.

Our abundance measurements are unique when compared to those extracted from previous samples because: (1) they are based on a representative sample; (2) they are unaffected by a systematic error that has either limited (Leccardi & Molendi 2008a) or likely biased (Mernier et al. 2017; Lovisari & Reiprich 2019) other measurements; (3) according to our conservative estimates, residual systematics must be smaller than 15%, and (4) for virtually all systems, they extend to R_{500}.

Because of these properties, several important implications follow from our measurements. Our profiles flatten out at large radii, suggesting early enrichment and significant feedback, admittedly not a new result (e.g., Fabjan et al. 2010; Planelles et al. 2014), however, the radial range and representative nature of our sample extends its import well beyond previous findings. It is worth noting that the feedback responsible for the enrichment of the outskirts is different from that currently occurring in the center of cool cores (e.g., Gaspari et al. 2011) in terms of power and affected area (although not underlying physics). While the AGN feedback of the BCG is strongly injecting kinetic energy and redistributing metals in the central regions of cool cores, the enrichment of the outskirts likely occurs through smoother injection during the proto-cluster phase. The cumulative effect of multiple smaller proto-galaxies (with lower self-regulated AGN and/or stellar feedback power and smaller affected area) is thus more important for the outskirt evolution, rather than impulsive bursts from a BCG.

[6] Actually, Zhang et al. (2019) provide a measure of the ICL plus central galaxy light. The estimate we quote has been derived by subtracting, in an approximate and conservative fashion, the contribution of the central galaxy.

[7] This is a conservative approach because: (1) systematics have been estimated for the abundance in the outskirts, where they are larger, and we apply them to the integrated quantity, and (2) the systematic on the abundance profile is itself conservative (see discussion in Sect. 3.2).

We find no evidence of segregation between cool-core and non-cool-core systems beyond ~0.3 R_{500}. This shows that, as was found for thermodynamic properties (Ghirardini et al. 2019), the physical state of the core does not affect global cluster properties. Our mean abundance within R_{500} shows a very modest scatter of <15%, suggesting the enrichment process must be quite similar in all these massive systems. This is a new finding and has significant implications for feedback processes. Together with the results from thermodynamic properties, that is, renormalized entropy profiles (see Fig. 6 of Ghirardini et al. 2019), it affords a coherent picture where feedback effects do not vary significantly from one system to another. Another way of looking at the low scatter in mean abundance is through the $M_{\rm gas,500}$ versus $M_{\rm Fe,500}$ relation. The tight nature of the correlation implies that $M_{\rm Fe,500}$ can be estimated with good accuracy from $M_{\rm gas,500}$. It will be interesting to see if and how these properties extend to less massive systems although, as pointed out earlier in this section, deriving iron abundances for such objects may only be possible with the advent of high spectra resolution instruments such as those on board XRISM (Tashiro et al. 2018; Guainazzi & Tashiro 2018) and ATHENA (Nandra et al. 2013).

Before moving on to discussing optical measurements, we would like to address two questions on size namely: (1) how can a sample of only 12 systems prove so powerful in providing constraints on the cluster population? and (2) will increasing the sample by, say, two or tenfold lead to more stringent measurements? The answer to these questions goes as follows. Measurements are limited either by statistical or systematic errors. In the case at hand, although statistical errors on individual systems are larger than systematic ones, errors on the mean are not. For example (see Sect. 3.2), the statistical error on the mean abundance is actually smaller than the systematic one. This implies that future measurements on larger samples will provide more stringent measurements only if the ensuing reduction on statistical errors will be accompanied by a comparable decrement on systematic uncertainties. For this and other reasons, several of the authors of this paper are engaged in ensuring that the instrumental background on ATHENA features the smallest systematics of any imaging X-ray mission ever flown.

For a subsample of 7 of our 12 systems, we have secured stellar masses. We have found that $M_{\rm star,500}$ and $M_{\rm gas,500}$ correlate well. The scatter, within the limits of the available data, is comparable with the one between $M_{\rm Fe,500}$ and $M_{\rm gas,500}$, reinforcing the concept that enrichment must be quite similar in all our systems. By combing stellar with ICM measurements, we have been able to take an inventory of Fe in clusters. We find that the amount diffused in the ICM with respect to that locked in stars, the so-called iron share, is very high, about 10 times, and features a moderate scatter around the mean value. The implication is that the bulk of Fe produced by stars is expelled in the ICM and that this process proceeds essentially at the same pace in all systems. Like the iron share, the effective yield features a modest scatter of ~13% around the mean value. This has significant implications for the approach taken when computing the iron masses expected from SN rates. Indeed, a description of the enrichment process through a simple equation (i.e., Eq. (12)), based on average properties such as mean SN rates and mean iron mass per SN, could not be justified if the distribution of measured iron yields in the cluster population were to be characterized by a large scatter or a multi-modal distribution. Thus, while the actual numbers we adopt could be incorrect, the procedure is most likely sound.

By comparing the measured effective iron yield with the expected one, we find that the efficiency for Fe production in cluster galaxies must be higher than predicted. Similar claims have been made before (e.g., Renzini & Andreon 2014), however previous measurements were based on non-representative samples and extended to significantly smaller radii. As pointed out in Molendi et al. (2016), this made such claims at the very least premature. The thorough analysis of systematic uncertainties conducted on X-ray (see Sect. 3.2) shows that ICM iron masses cannot explain the discrepancy between measured and predicted effective yields. The only possible alternatives are that either SN yields or stellar masses, or both, have been significantly underestimated.

As discussed in Sect. 5.2, recent estimates have revised upwards the SN Ia Fe yield in galaxy clusters, however the ensuing increase in predicted effective yield is insufficient to bring it into agreement with the measured one. The remaining option is that stellar masses must be underestimated. In Sects. 4.2 and 5.3 we have seen this may be happening for a number of reasons. Amongst them some, such as the conversion of multi-band optical and infrared data to stellar masses, are related to technical issues and others, such as a substantial contribution from ICL or from a stellar population lying in the outskirts, are of a more astrophysical and perhaps appealing nature. However, whatever the explanation for the discrepancy turns out to be, since the scatter in the correlations found in all our scaling relations is much smaller than the ratio between measured and predicted iron share, the correction factor will have to be very similar in all clusters. This could either imply that we need to correct the conversion adopted to go from multi-band optical and infrared data to stellar masses by a constant factor or, more intriguingly, that a sizeable but roughly constant fraction of ICL or stellar mass is missing from our inventory.

To summarize, although current estimates suggest, with some strength, that the measured iron mass in clusters is well in excess of the predicted one, systematic errors prevent us from making a definitive statement. Further advancements will only be possible when systematic uncertainties, principally those associated to the estimate of stellar masses, both within and beyond R_{500}, can be reduced.

7. Summary

We have measured iron abundance for the X-COP sample of massive clusters, these are our main findings.

- Our measurements are unprecedented for four reasons: (1) they are based on a representative sample; (2) they are unaffected by a systematic error that has plagued previous measurements; (3) they feature residual systematics smaller than 15%, and (4) for virtually all systems they extend to R_{500}.
- Our profiles flatten out at large radii suggesting early enrichment and significant feedback, admittedly not a new result (e.g., Fabjan et al. 2010; Planelles et al. 2014), however the radial range and representative nature of our sample extends its import well beyond previous findings.
- We find no evidence of segregation between cool-core and non-cool-core systems beyond ~0.3 R_{500}. This shows that, as was found for thermodynamic properties (Ghirardini et al. 2019), the physical state of the core does not affect global cluster properties.
- Our mean abundance within R_{500} shows a very modest scatter, <15%, suggesting the enrichment process must be quite similar in all these massive systems. This is a new finding and has significant implications on feedback processes. Together with results from thermodynamic properties, that is, renormalized entropy profiles (see Fig. 6 of Ghirardini et al. 2019),

it affords a coherent picture where feedback effects do not vary significantly from one system to another.
– Another way of looking at the low scatter in mean abundance is through the $M_{\rm gas,500}$ versus $M_{\rm Fe,500}$ relation. The tight nature of the correlation, σ <15%, implies that $M_{\rm gas,500}$ can be used as a robust proxy for $M_{\rm Fe,500}$.

For a subsample of 7 of our 12 systems we have secured stellar masses. By combing stellar with ICM measurements, we have derived the following results.

– The amount of Fe diffused in the ICM with respect to that locked in stars, the so-called iron share, is very high, about 10 times, and features a moderate scatter around the mean value. The implication is that the bulk of Fe produced by stars is expelled in the ICM and that this process proceeds at a similar pace in all systems.
– Although current estimates suggest, with some strength, that the measured iron mass in clusters is well in excess of the predicted one, systematic errors prevent us from making a definitive statement.

Further advancements will only be possible when systematic uncertainties, principally those associated to the estimate of stellar masses, both within and beyond R_{500}, can be reduced.

Acknowledgements. We warmly thank F. Mernier, L. Lovisari, and D. Maoz for providing feedback on an early version of this paper. We acknowledge financial contribution from contract ASI-INAF n. 2017-14-H.0. This research is based on observations obtained with *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). M.G. is supported by the Lyman Spitzer Jr. Fellowship (Princeton University) and by NASA *Chandra* GO8-19104X/GO9-20114X and HST GO-15890.020-A grants. VB acknowledges support by the DFG project nr. 415510302.

References

Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
Arnaud, K. A. 1996, in Astronomical Data Analysis Software and Systems V, eds. G. H. Jacoby, & J. Barnes, ASP Conf. Ser., 101, 17
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Barret, D., den Herder, J. W., Piro, L., et al. 2013, ArXiv e-prints [arXiv:1308.6784]
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Biffi, V., Planelles, S., Borgani, S., et al. 2017, MNRAS, 468, 531
Biffi, V., Planelles, S., Borgani, S., et al. 2018, MNRAS, 476, 2689
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buote, D. A. 2000, MNRAS, 311, 176
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Cavagnolo, K. W., Donahue, M., Voit, G. M., & Sun, M. 2009, ApJS, 182, 12
Chabrier, G. 2003, PASP, 115, 763
Chiu, I., Mohr, J. J., McDonald, M., et al. 2018, MNRAS, 478, 3072
Cucchetti, E., Clerc, N., Pointecouteau, E., Peille, P., & Pajot, F. 2019, A&A, 629, A144
De Grandi, S., Ettori, S., Longhetti, M., & Molendi, S. 2004, A&A, 419, 7
De Grandi, S., Santos, J. S., Nonino, M., et al. 2014, A&A, 567, A102
De Luca, A., & Molendi, S. 2004, A&A, 419, 837
Dolag, K., Hansen, F. K., Roncarelli, M., & Moscardini, L. 2005, MNRAS, 363, 29
Eckert, D., Roncarelli, M., Ettori, S., et al. 2015, MNRAS, 447, 2198
Eckert, D., Ettori, S., Pointecouteau, E., et al. 2017, Astron. Nachr., 338, 293
Eckert, D., Ghirardini, V., Ettori, S., et al. 2019, A&A, 621, A40
Ettori, S., Fabian, A. C., Allen, S. W., & Johnstone, R. M. 2002, MNRAS, 331, 635
Ettori, S., Ghirardini, V., Eckert, D., et al. 2019, A&A, 621, A39
Fabjan, D., Borgani, S., Tornatore, L., et al. 2010, MNRAS, 401, 1670
Friedmann, M., & Maoz, D. 2018, MNRAS, 479, 3563
Gaspari, M., Melioli, C., Brighenti, F., & D'Ercole, A. 2011, MNRAS, 411, 349
Ghirardini, V., Ettori, S., Eckert, D., et al. 2018, A&A, 614, A7
Ghirardini, V., Ettori, S., Eckert, D., & Molendi, S. 2019, A&A, 627, A19
Ghizzardi, S., Molendi, S., Pizzolato, F., & De Grandi, S. 2004, ApJ, 609, 638
Girelli, G., Pozzetti, L., Bolzonella, M., et al. 2020, A&A, 634, A135
Gonzalez, A. H., Sivanandam, S., Zabludoff, A. I., & Zaritsky, D. 2013, ApJ, 778, 14
Greggio, L., & Renzini, A. 2011, Stellar Populations. A User Guide from Low to High Redshift (Weinheim: Wiley-VCH Verlag)
Guainazzi, M., & Tashiro, M. S. 2018, Proceedings of the IAU Symposium IAUS342 [arXiv:1807.06903]
Kelly, B. C. 2007, ApJ, 665, 1489
Kravtsov, A. V., Nagai, D., & Vikhlinin, A. A. 2005, ApJ, 625, 588
Kravtsov, A. V., Vikhlinin, A. A., & Meshcheryakov, A. V. 2018, Astron. Lett., 44, 8
Kriss, G. A., Cioffi, D. F., & Canizares, C. R. 1983, ApJ, 272, 439
Kroupa, P. 2001, MNRAS, 322, 231
Leauthaud, A., George, M. R., Behroozi, P. S., et al. 2012, ApJ, 746, 95
Leccardi, A., & Molendi, S. 2007, A&A, 472, 21
Leccardi, A., & Molendi, S. 2008a, A&A, 487, 461
Leccardi, A., & Molendi, S. 2008b, A&A, 486, 359
Leccardi, A., Rossetti, M., & Molendi, S. 2010, A&A, 510, A82
Lin, Y.-T., Stanford, S. A., Eisenhardt, P. R. M., et al. 2012, ApJ, 745, L3
Loewenstein, M. 2013, ApJ, 773, 52
Lovisari, L., & Reiprich, T. H. 2019, MNRAS, 483, 540
Maoz, D., & Graur, O. 2017, ApJ, 848, 25
Maoz, D., Sharon, K., & Gal-Yam, A. 2010, ApJ, 722, 1879
Maraston, C. 2005, MNRAS, 362, 799
Matsushita, K. 2011, A&A, 527, A134
McLaughlin, D. E. 1999, AJ, 117, 2398
Mernier, F., de Plaa, J., Kaastra, J. S., et al. 2017, A&A, 603, A80
Mernier, F., Biffi, V., Yamaguchi, H., et al. 2018, Space Sci. Rev., 214, 129
Molendi, S., Eckert, D., De Grandi, S., et al. 2016, A&A, 586, A32
Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJS, 206, 8
Nandra, K., Barret, D., Barcons, X., et al. 2013, ArXiv e-prints [arXiv:1306.2307]
Nevalainen, J., David, L., & Guainazzi, M. 2010, A&A, 523, A22
Planelles, S., Borgani, S., Dolag, K., et al. 2013, MNRAS, 431, 1487
Planelles, S., Borgani, S., Fabjan, D., et al. 2014, MNRAS, 438, 195
Pratt, G. W., Arnaud, M., Piffaretti, R., et al. 2010, A&A, 511, A85
Rasia, E., Mazzotta, P., Bourdin, H., et al. 2008, ApJ, 674, 728
Rasia, E., Borgani, S., Murante, G., et al. 2015, ApJ, 813, L17
Read, A. M., Guainazzi, M., & Sembay, S. 2014, A&A, 564, A75
Renzini, A., & Andreon, S. 2014, MNRAS, 444, 3581
Salpeter, E. E. 1955, ApJ, 121, 161
Sand, D. J., Graham, M. L., Bildfell, C., et al. 2011, ApJ, 729, 142
Schechter, P. 1976, ApJ, 203, 297
Shan, Y., McDonald, M., & Courteau, S. 2015, ApJ, 800, 122
Sifón, C., Herbonnet, R., Hoekstra, H., van der Burg, R. F. J., & Viola, M. 2018, MNRAS, 478, 1244
Simionescu, A., Allen, S. W., Mantz, A., et al. 2011, Science, 331, 1576
Simionescu, A., Werner, N., Mantz, A., Allen, S. W., & Urban, O. 2017, MNRAS, 469, 1476
Smith, R. K., Brickhouse, N. S., Liedahl, D. A., & Raymond, J. C. 2001, ApJ, 556, L91
Snowden, S. L., Mushotzky, R. F., Kuntz, K. D., & Davis, D. S. 2008, A&A, 478, 615
Tashiro, M., Maejima, H., Toda, K., et al. 2018, in SPIE Conf. Ser., Proc. SPIE, 10699, 1069922
Truong, N., Rasia, E., Biffi, V., et al. 2019, MNRAS, 484, 2896
Urban, O., Simionescu, A., Werner, N., et al. 2014, MNRAS, 437, 3939
Urban, O., Werner, N., Allen, S. W., Simionescu, A., & Mantz, A. 2017, MNRAS, 470, 4583
Urdampilleta, I., Mernier, F., Kaastra, J. S., et al. 2019, A&A, 629, A31
van der Burg, R. F. J., Hoekstra, H., Muzzin, A., et al. 2015, A&A, 577, A19
van der Burg, R. F. J., McGee, S., Aussel, H., et al. 2018, A&A, 618, A140
Werner, N., Urban, O., Simionescu, A., & Allen, S. W. 2013, Nature, 502, 656
Zhang, Y., Yanny, B., Palmese, A., et al. 2019, ApJ, 874, 165
Zibetti, S., White, S. D. M., Schneider, D. P., & Brinkmann, J. 2005, MNRAS, 358, 949

Appendix A: Abundance profiles for X-COP clusters

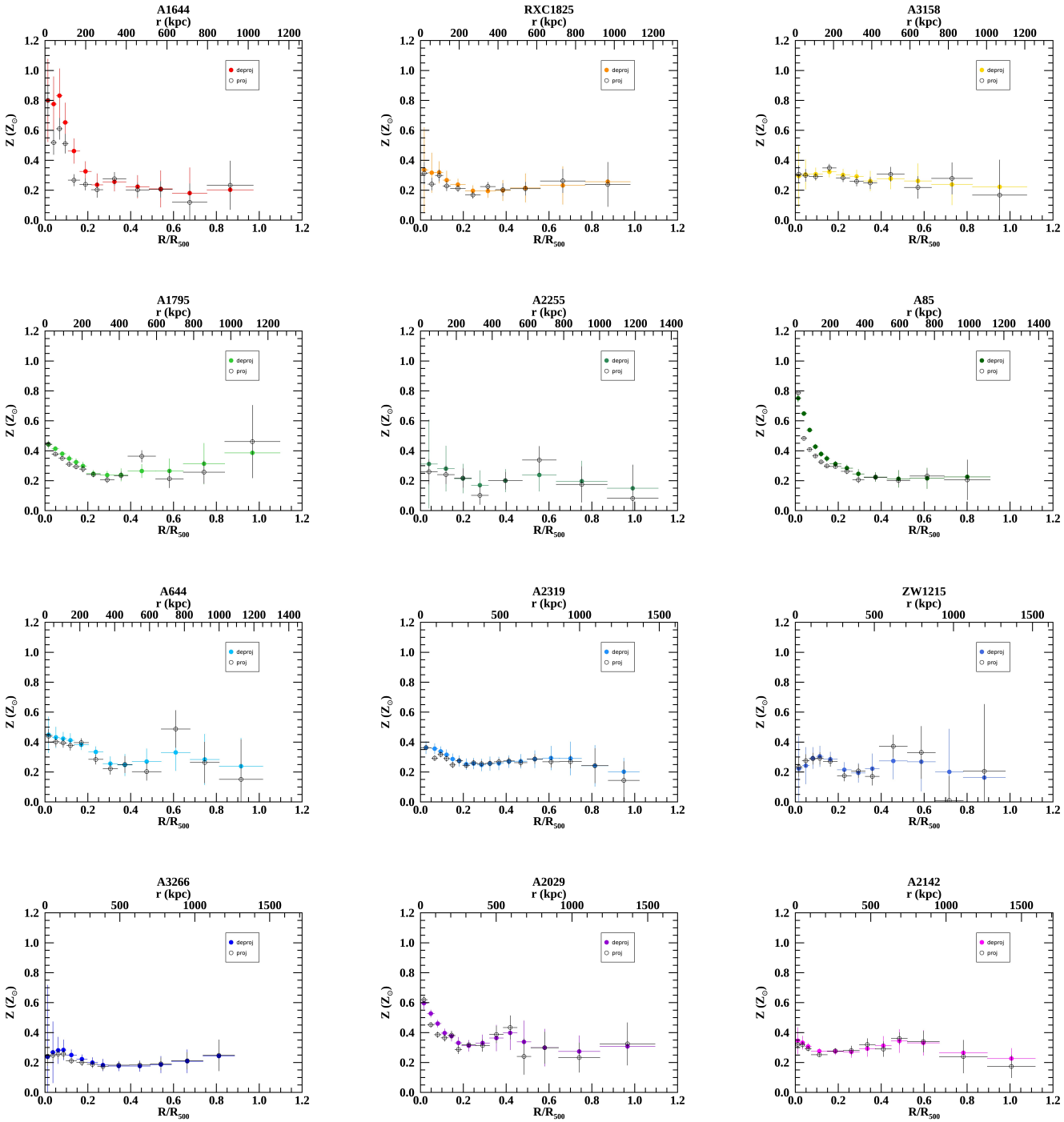

Fig. A.1. Projected (open black circles) and deprojected (coloured closed circles) abundance profiles for all X-COP clusters.

Appendix B: Investigating the L-shell bias

In this appendix we investigate the difference between metal abundance measurements derived when including and excluding the Fe L-shell in spectral fits. We aim at providing a plausible explanation; determination of the actual explanation is beyond not just this paper but also 20 years of efforts in the calibration of effective areas, and the instrumental and celestial backgrounds of the EPIC instrument. Let us start by describing in a little more detail what we are trying to explain. In Sect. 2.2 we found that as we move out to large radii measurements of the Fe abundance performed while including the Fe L-shell spectral range sometimes drop to a low value. In these instances, the measurement typically features a small error. Excluding the L-shell range from the fit returns a substantially larger value of the abundance, with larger errors. The fact that the drop in abundance is not found in all bins beyond a certain radius, but only in a few of them, suggests that it results from more than one cause and that the particular combination responsible for the drop must be an infrequent one. We shall assume the drop in Fe L-shell measurements is the result of a combination of systematic errors in effective areas and background modeling. Since we require plausibility, the magnitude of the errors will be limited by our knowledge of systematics.

We simulate spectra for a region at large radii, where the background is dominant above ∼2 keV. For simplicity we consider only the MOS2 detector. The X-ray and instrumental background modeling is similar to the one used in our analysis (see Sect. 2.2), the source model is taken from one of the outer bins of A644, and the metal abundance is set at $0.25\,Z_{\odot}$. We perturb the value of the normalization of the CXB and the foreground component by up to 10% and 30% respectively, and of the slope of the CXB by $\Delta\Gamma = 0.025$, in agreement with current knowledge of fore- and background variations. The effective area is perturbed with respect to the real one by ∼5−10%, in agreement with findings from the analysis of systematics (Read et al. 2014). The perturbed model is used to generate simulated data, and we do include Poisson statistics but we impose unrealistically long exposure times (10 Ms) to ease inspection of residuals. The simulated data is then fit with the unperturbed model and effective areas, which are the same as those adopted for the analysis of our real data. We experimented with different values of perturbations for model parameters and effective areas finding that, in most instances, the fit to the simulated data returned values of the metal abundance consistent with the simulated one. In a few cases, typically where the effective area used in the simulation featured a drop of 5−10% around 1.1 keV and the simulated CXB and foreground normalizations were 5−10% lower and 20−30% higher respectively than the ones adopted in the spectral analysis, the best-fitting value of the metal abundance, as measured using the full energy band, was significantly smaller than the simulated one, $\sim 0.10-0.15\,Z_{\odot}$. In these cases, when we excluded the 0.9−1.3 keV range from the fitting procedure, the value of the metal abundance rose to higher values, just as we had found in the real data, and became compatible with the simulated one. Moreover, as for the spectral analysis conducted on real data, the error on the metal abundance computed using the L-shell data was substantially smaller, by about a factor of 2, than the one derived when we excluded it; the temperature did not vary significantly between the two fits (5−10%).

To gain some insight into what leads to the L-shell bias, we show the simulated data, best-fitting model and residuals in the form of the ratio of data over model, when fitting the full energy range (Fig. B.1) and when excluding the L-shell region (Fig. B.2)

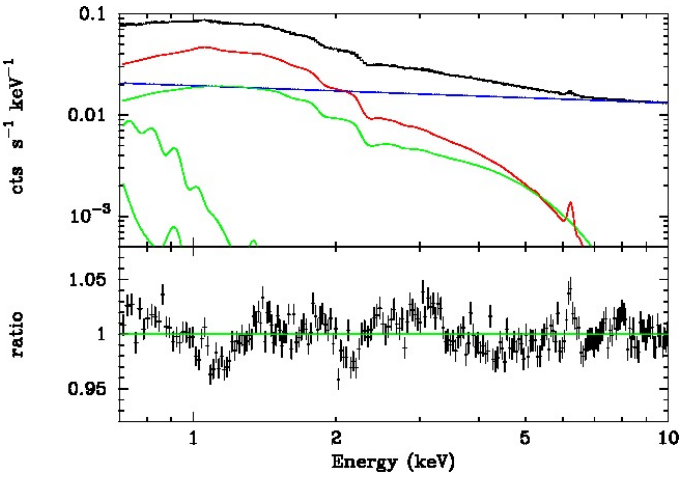

Fig. B.1. *Top panel*: simulated data and best fit model when making use of the full spectral range. The source is shown in red, the sky fore- and background in green, and the instrumental background in blue. *Bottom panel*: residuals in the form of a data over model ratio. All residual are below 5%, and are due to differences in the effective area and fore- and background parameters used in the simulation with respect to those adopted for the fit. We note how the low abundance value forced by the data around the Fe L-shell results in residuals around the Fe Kα line.

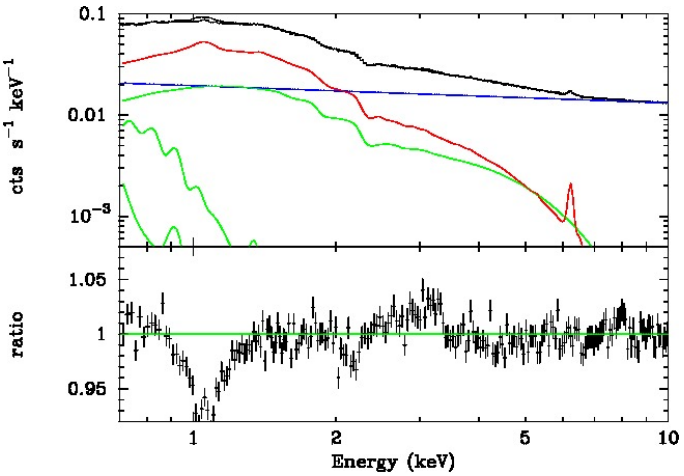

Fig. B.2. Same as for Fig. B.1 except that the L-shell range (0.9.13 keV) is excluded from fitting. We note how the residuals around the Fe Kα line are no longer present.

for one of the cases where L-shell and Kα measurements do not agree. As we can see from Fig. B.1, where the L-shell is included in the fitting procedure, the combination of the perturbation of the effective area around 1.1 keV and the normalization of the cosmic background and foreground components leads to a reduction of the simulated data with respect to the model at this energy; the fitting procedure adjusts by decreasing the metal abundance. As a results of this decrement, the model produces an Fe Kα line that is below the simulated one. This is clearly visible when inspecting residuals around 6 keV. An important point is that the fitting algorithm resolves the tension between reproducing the L-shell emission and the Kα line in favor of the former because the statistics around 1 keV is higher than around 6 keV. Inspection of Fig. B.2 shows that if the fitting routine does not have to reproduce the L-shell emission, it can fit adequately the Kα line and return an unbiased estimate of the metal abundance.

Appendix C: Biases in average metallicity measurements

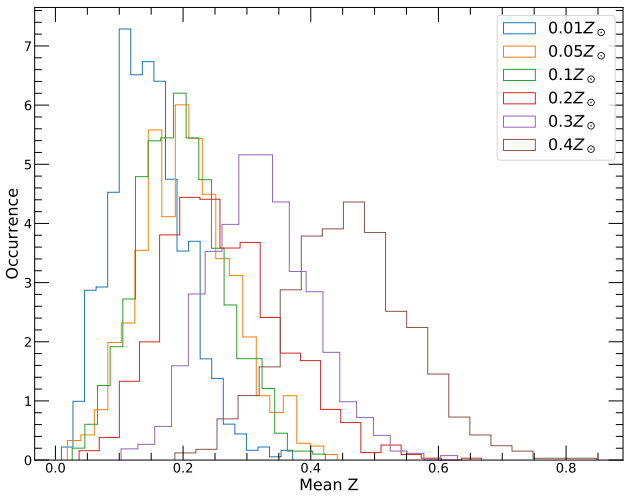

Fig. C.1. Distributions of simulated sample means in the outermost radial bin for various input metallicities and for 100 X-COP-like samples. The input metallicities range from 0.01 (blue) to 0.4 $Z_\odot$ (brown).

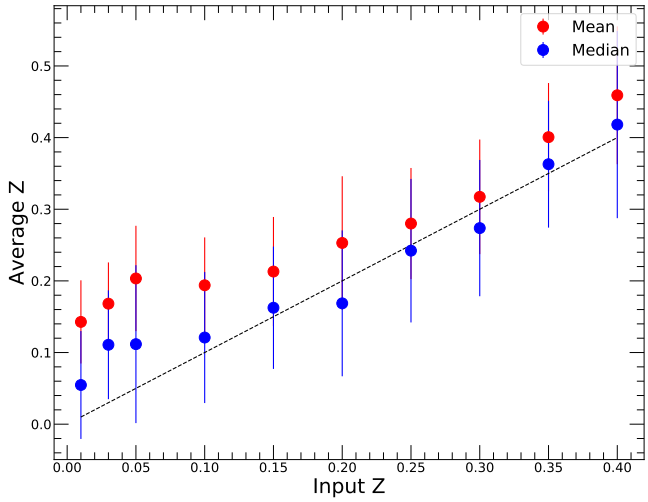

Fig. C.2. Most probable average sample values (filled circles) and standard deviation of the expected sample averages (error bars) as a function of input metallicity for sets of 100 simulated spectra. The data points show the expected weighted mean (red) and median (blue). The black dashed line is the one-to-one relation.

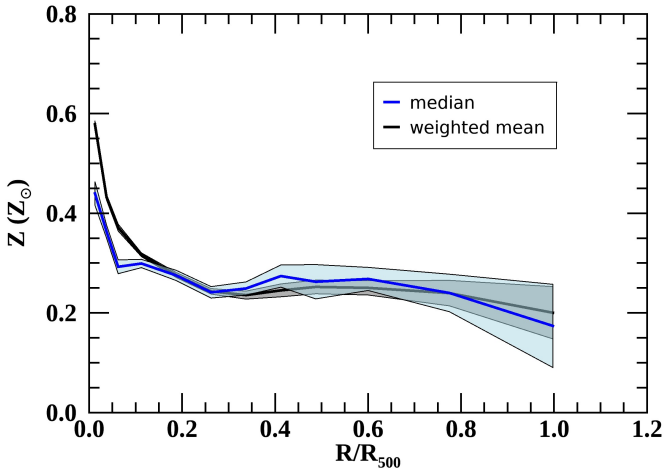

Fig. C.3. Comparison between the weighted-mean abundance profile (black line), as derived in Sect. 3.1, and the median abundance profile (blue line) for the X-COP sample. The shaded areas represent the statistical errors of the profiles. The two profiles exhibit an excellent agreement.

The mean metallicity profile presented in Sect. 3 was computed by performing a weighted mean of the individual iron abundance measurements. This procedure assumes Gaussian posterior distributions for the individual measurements, which is not necessarily verified in the case of the metal abundance since it is a positive-definite quantity. Especially in the case of low abundance and large uncertainty, we expect the posterior distribution to be strongly skewed toward high values, and the mean of the posterior to be biased high.

To verify whether our measurements can be affected by this bias, especially in the outermost regions, we performed Monte Carlo simulations of spectra for various input metallicities. We simulated data based on the observed spectra in the outermost radial bin of a typical X-COP cluster (A3158) and fitted the simulated spectra using the X-COP procedure. For a grid of metal abundances spanning the range [0.01−0.4] $Z_\odot$, we generated 100 realizations of the spectra and studied the distribution of fitted metal abundance values. As expected, we find that when the uncertainties are large, the distributions are highly non-Gaussian and skewed towards high values. Then, from the generated metal abundance measurements, we randomly picked a sample of 12 observed values and computed the weighted mean to mimic the X-COP sample selection. We repeated the selection 100 times to study the distribution of expected mean values for the sample.

In Fig. C.1 we show the distributions of expected mean values for various input metallicities. We can see that for very low metallicity ($Z = 0.01\,Z_\odot$) the most probable mean value is ∼0.12 and recovering the true value is very unlikely. For higher metal abundance values the peak of the distribution gets progressively closer to the true value, although it remains biased, even for $Z = 0.4\,Z_\odot$. However, when performing the same exercise with the median of the distribution instead of the weighted mean, we find values that are much closer to the true value. In Fig. C.2 we show the most probable sample values as a function of metallicity for the weighted mean and the median. The error bars show the standard deviation of the distribution of expected mean values. While at very low metallicity ($Z < 0.1\,Z_\odot$) the median is biased as well, above this threshold the median accurately recovers the true value, whereas the weighted mean remains biased. Therefore, the comparison between the mean and the median value can tell us whether our measurement is affected by a strong bias.

To this aim, we computed the median profile of X-COP metal abundance measurements and compared it with the weighted mean (see Fig. C.3). The median profile is consistent with the mean profile at each radius, implying that our measurements are robust. However, we caution that in other contexts averaging metal abundance values can introduce important biases when the non-Gaussianity of posterior metal abundance values is ignored.