Athena X-IFU synthetic observations of galaxy clusters to probe the chemical enrichment of the Universe

E. Cucchetti, E. Pointecouteau, P. Peille, N. Clerc, E. Rasia, V. Biffi, S. Borgani, L. Tornatore, K. Dolag, M. Roncarelli, et al.

A&A 620, A173 (2018)
https://doi.org/10.1051/0004-6361/201833927

Astronomy & Astrophysics

Athena X-IFU synthetic observations of galaxy clusters to probe the chemical enrichment of the Universe

E. Cucchetti[1], E. Pointecouteau[1], P. Peille[2], N. Clerc[1], E. Rasia[3], V. Biffi[3,4], S. Borgani[3,4,5], L. Tornatore[3], K. Dolag[6,7], M. Roncarelli[8,9], M. Gaspari[11,⋆], S. Ettori[9,10], E. Bulbul[12], T. Dauser[13], J. Wilms[13], F. Pajot[1], and D. Barret[1]

[1] IRAP, Université de Toulouse, CNRS, CNES, UPS, Toulouse, France
e-mail: edoardo.cucchetti@irap.omp.eu
[2] CNES, 18 Avenue Edouard Belin, 31400 Toulouse, France
[3] INAF, Osservatorio Astronomico di Trieste, Via Tiepolo 11, 34131 Trieste, Italy
[4] Dipartimento di Fisica dell'Università di Trieste, Sezione di Astronomia, Via Tiepolo 11, 34131 Trieste, Italy
[5] INFN – National Institute for Nuclear Physics, Via Valerio 2, 34127 Trieste, Italy
[6] University Observatory Munich, Scheinerstr. 1, 81679 Munich, Germany
[7] Max Plank Institut für Astrophysik, Karl-Schwarzschield Strasse 1, 85748 Garching bei Munchen, Germany
[8] Dipartimento di Fisica e Astronomia, Università di Bologna, Via Gobetti 93, 40127 Bologna, Italy
[9] INAF, Osservatorio di Astrofisica e Scienza dello Spazio, Via Pietro Gobetti 93/3, 40129 Bologna, Italy
[10] INFN, Sezione di Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy
[11] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[12] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[13] Dr. Karl Remeis-Observatory and Erlangen Centre for Astroparticle Physics, Sternwartstr. 7, 96049 Bamberg, Germany

Received 23 July 2018 / Accepted 24 September 2018

ABSTRACT

Answers to the metal production of the Universe can be found in galaxy clusters, notably within their intra-cluster medium (ICM). The X-ray Integral Field Unit (X-IFU) on board the next-generation European X-ray observatory *Athena* (2030s) will provide the necessary leap forward in spatially-resolved spectroscopy required to disentangle the intricate mechanisms responsible for this chemical enrichment. In this paper, we investigate the future capabilities of the X-IFU in probing the hot gas within galaxy clusters. From a test sample of four clusters extracted from cosmological hydrodynamical simulations, we present comprehensive synthetic observations of these clusters at different redshifts (up to $z \leq 2$) and within the scaled radius R_{500} performed using the instrument simulator SIXTE. Through 100 ks exposures, we demonstrate that the X-IFU will provide spatially resolved mapping of the ICM physical properties with little to no biases ($\lesssim$5%) and well within statistical uncertainties. The detailed study of abundance profiles and abundance ratios within R_{500} also highlights the power of the X-IFU in providing constraints on the various enrichment models. From synthetic observations out to $z = 2$, we have also quantified its ability to track the chemical elements across cosmic time with excellent accuracy, and thereby to investigate the evolution of metal production mechanisms as well as the link to the stellar initial mass-function. Our study demonstrates the unprecedented capabilities of the X-IFU of unveiling the properties of the ICM but also stresses the data analysis challenges faced by future high-resolution X-ray missions such as *Athena*.

Key words. galaxies: clusters: intracluster medium – galaxies: abundances – galaxies: fundamental parameters – techniques: imaging spectroscopy – methods: numerical – X-rays: galaxies: clusters

1. Introduction

Metals and other heavy elements in the intra-cluster medium (ICM) represent a fossil record of the chemical evolution of the Universe. Trapped in the dark matter (DM) potential of galaxy clusters (White et al. 1993), they remain unaltered within the optically-thin, collisionless thermal plasma. Elements originate within stars or through supernovæ (SN), before being spread by stellar winds or by the SN explosions. Hence, the chemical enrichment of a given cluster relates to the integrated star formation history of the cluster, as well as to the overall stellar initial mass function (IMF). The abundances and spatial distribution of metals in the ICM can also be connected to its dynamical history and to the mechanical action of active galactic nuclei (AGNs) outflows or jets (e.g. Gaspari et al. 2011).

Most of the low-mass elements (C, O, Mg, Si, and S) are produced by end-of-life massive stars ($\geq$10 $M_\odot$) undergoing core-collapse supernovæ (SN_{cc}; see Nomoto et al. 2013, for a review). The evolution of SN_{cc}-related enrichment through time is dictated by the initial mass and metallicity of the progenitor star. High-mass elements, from Si-like elements (Al, Si, S, Ca, and Ar) to Fe and Ni, are on the other hand the result of thermonuclear reactions occurring during the explosion of white dwarfs (type Ia supernovæ – SN_{Ia}; Hillebrandt et al. 2013). Although the mechanisms of these explosions – either via accretion of a companion star onto the white dwarf (Whelan & Iben 1973) or via mergers of binary systems (Webbink 1984) – is still poorly understood (see Maoz et al. 2014), the timescale of these events, related to longer-living low-mass stars, suggests a later enrichment across cosmic time. Traces of other elements (C, N, Ne, and Na) can also be produced when low- and

⋆ *Einstein* and *Spitzer* Fellow.

intermediate-mass stars (typically $\leq 6\,M_{\odot}$) enter their asymptotic giant branch (AGB) phase (Iben & Renzini 1983). The individual study of these phenomena based on detailed observations of nearby SN is difficult as they are very rare. Rather than a direct study on stellar populations, the detailed spectroscopic study of the ICM is an interesting alternative probe to test metal production models up to the early periods of the Universe.

Beyond the first steps in high-resolution X-ray spectroscopy (Canizares et al. 1979, 1982) and despite the lack of spatial resolution (Peterson & Fabian 2006), the advent of high-resolution grating instruments such as *XMM-Newton*/RGS (den Herder et al. 2001) and *Chandra*/HETG (Canizares et al. 2005) drastically changed our view of the ICM enrichment, by giving access for the very first time to a large number of atomic lines (de Plaa et al. 2007; de Plaa 2013; Molendi et al. 2016; Werner et al. 2007). Clusters have, therefore, become excellent laboratories in which to test plasma physics and the chemical enrichment models up to the present epoch (see Werner et al. 2008, for a review). Despite limited spectral resolutions, instruments based on charged coupled devices (CCDs) have also been pushed to the maximum of their abilities to benefit of their spatial resolution in investigating the spatial distribution of chemical elements in the ICM (de Grandi & Molendi 2009; Mernier et al. 2016a,b, 2017).

The perspective of micro-calorimeter-based imaging spectrometers, such as the soft X-ray spectrometer (SXS) on board *Hitomi* (Takahashi et al. 2016), has opened new possibilities in studying the ICM: from the spatial scales of the enrichment (sources of production, processes of mixing and dispersion) to the kinematics of the hot gas (turbulence, shocks Hitomi Collaboration 2016, 2018a,b,c), which complement the indirect estimates via surface brightness and warm gas tracers (e.g. Churazov et al. 2012; Gaspari & Churazov 2013; Hofmann et al. 2016; Gaspari et al. 2018). Unfortunately, the short lifetime of the SXS gave only a glimpse of its potential. These renewed capabilities in galaxy-cluster observation now rely on future missions, such as the X-ray Recovery Imaging and Spectroscopy Mission (XRISM; Ishisaki et al. 2018) or the Advanced Telescope for High-ENergy Astrophysics (*Athena*; Nandra et al. 2013). Namely, the X-ray Integral Field Unit (X-IFU) on board the future European X-ray observatory (Barret et al. 2016; Pajot et al. 2018), will provide narrow-field observations (5′ in equivalent field-of-view diameter) over the 0.2–12 keV bandpass, with a required 5″ spatial resolution and an unprecedented spectral resolution of 2.5 eV (required up to 7 keV).

Investigating the chemical enrichment of the Universe is one of *Athena*'s prime science objectives (Ettori et al. 2013; Pointecouteau et al. 2013) which drives top-level performances of the telescope . In addition to the spectral resolution of the X-IFU, which will allow to resolve faint atomic lines of less abundant elements, this science objective drives the need for a high effective area of the telescope along with a well-calibrated low energy band, required to accurately resolve lines of light elements such as C ($\geq$0.2 keV). Number of breakthroughs on the study of chemical species and their evolution should in fact come from measurement in the low-energy band, where the effective area is the highest. The fine spectroscopic capabilities of the X-IFU in this energy band will probe the production and circulation of metals within galaxy clusters across cosmic time, up to a redshift of $z \leq 2$ and a distance of R_{500}[1] from the cluster's centre. By accurately measuring the abundances of the most common elements (e.g. O, Si, S, and Fe), the X-IFU will be capable of constraining the number of time-integrated SN_{Ia} and SN_{cc} products. For the first time, the spatially-resolved measurements of less abundant elements (e.g. C, Al, S, and Ca) as well as rare elements (e.g. Mn, Cr, and Ti) will provide insights on the initial metallicity of the SN_{Ia} progenitors, and therefore on their formation mechanisms. The science of the chemical enrichment is a driver of the performance of the instrument, which needs to be assessed before launch.

In this paper, we investigate the feasibility of recovering the physical parameters of the ICM through X-IFU observations. Careful attention is given to the different enrichment mechanisms and their evolution over time. We used a sample of four simulated galaxy clusters with different masses studied at different redshifts, obtained via hydrodynamical cosmological simulations (Rasia et al. 2015; Biffi et al. 2017). These objects are passed as input to a dedicated end-to-end (E2E) simulation pipeline of the X-IFU instrument, based on the simulator SIXTE (Wilms et al. 2014). In Sect. 2, we present the properties of the sample of simulated clusters. This is followed by a detailed description of our simulation pipeline (Sect. 3). The data analysis, post-processing procedures and results validation are in turn described (Sect. 4). The outputs of our synthetic observations obtained through the pipeline for the four local clusters are then used (Sect. 5) to infer the main properties of the sample and study its enrichment. This investigation is also extended to higher redshift values (Sect. 6) to look into the X-IFU abilities to capture the evolution of abundances through cosmic time. Finally, results and outcomes of our study are discussed (Sect. 7).

2. Generation of the cluster sample

The sample of four clusters of galaxies analysed in this study is taken from Biffi et al. (2018) and includes two massive and two smaller systems, to bracket a broad mass range across the considered redshift values (Table 1). In both mass bins, we choose a cool-core (CC) and a non-cool-core cluster (NCC), defined based on their pseudo-entropy profiles as described in Leccardi et al. (2010). This small sample gives a view of part of the expected cluster population planned to be investigated by the X-IFU. The objects are part of a larger set of 29 Lagrangian regions extracted from a parent cosmological DM-only simulation and re-simulated at higher resolution including baryons (see Bonafede et al. 2011). The parent cosmological volume is $1\,h^{-1}$ Gpc per side and adopts a Λ-CDM cosmological model with $\Omega_{\rm M} = 0.24$, $\Omega_{\rm b} = 0.04$, $H_0 = 72\,{\rm km\,s^{-1}\,Mpc^{-1}}$ (i.e. $H_0 = h \times H_{100}$, where $h = 0.72$ and $H_{100} = 100\,{\rm km\,s^{-1}\,Mpc^{-1}}$), $\sigma_8 = 0.8$ and $n_{\rm s} = 0.96$, consistent with WMAP-7 constraints given in Komatsu et al. (2011).

The magnification simulations were performed with a version of the tree-PM smoothed particle hydrodynamics (SPH) code GADGET-3 (Springel 2005), including an improved hydrodynamical scheme (Beck et al. 2016) and a variety of physical processes describing the evolution of the baryonic component (see Rasia et al. 2015, for more details). Briefly, these comprise metallicity-dependent radiative cooling (Wiersma et al. 2009), star formation and stellar feedback (thermal supernova feedback and galactic winds, see Springel & Hernquist 2003), cold and hot gas accretion onto super-massive black holes powering AGN thermal feedback (Steinborn et al. 2015; modelling the action of cold accretion Gaspari & Sdowski 2017), and metal enrichment (Tornatore et al. 2004, 2007)

[1] R_{500} is the radius including a density contrast of 500 times the critical density of the Universe, $\rho_c = 3H(z)^2/8\pi G$, at the given redshift z.

Table 1. Properties of the simulated clusters at different redshift values in their evolution.

Name Type		C1 CC	C2 NCC	C3 CC	C4 NCC
$z = 0.105$	R_{500} (kpc h^{-1})	723	799	1027	1009
	M_{500} ($M_\odot$ h^{-1})	2.39×10^{14}	3.22×10^{14}	6.86×10^{14}	6.51×10^{14}
	T_{500} (keV)	3.22	4.20	6.47	6.36
$z = 0.5$	R_{500} (kpc h^{-1})	552	676	694	715
	M_{500} ($M_\odot$ h^{-1})	1.55×10^{14}	2.84×10^{14}	3.08×10^{14}	3.38×10^{14}
	T_{500} (keV)	2.88	4.06	4.58	4.84
$z = 1.0$	R_{500} (kpc h^{-1})	389	396	446	570
	M_{500} ($M_\odot$ h^{-1})	0.92×10^{14}	0.97×10^{14}	1.38×10^{14}	2.89×10^{14}
	T_{500} (keV)	2.41	2.47	3.12	4.41
$z = 1.48$	R_{500} (kpc h^{-1})	269	289	345	351
	M_{500} ($M_\odot$ h^{-1})	0.50×10^{14}	0.62×10^{14}	1.07×10^{14}	1.12×10^{14}
	T_{500} (keV)	1.76	2.10	3.07	3.08
$z = 2.0$	R_{500} (kpc h^{-1})	174	181	220	215
	M_{500} ($M_\odot$ h^{-1})	0.23×10^{14}	0.25×10^{14}	0.45×10^{14}	0.42×10^{14}
	T_{500} (keV)	1.19	1.55	2.24	1.86

from SN_{Ia}, SN_{cc}, and AGB stars. Specifically, we assumed the IMF by Chabrier (2003), the mass-dependent lifetimes by Padovani & Matteucci (1993) and stellar yields by Thielemann et al. (2003) for SN_{Ia}, Woosley & Weaver (1995) and Romano et al. (2010) for SN_{cc}, and Karakas (2010) for AGB stars.

In our model of chemical enrichment, we follow the production and evolution of 15 chemical species: H, He, C, Ca, O, N, Ne, Mg, S, Si, Fe, Na, Al, Ar, and Ni. These elements are the individual species traced in the simulations. Although these do not cover the full spectrum of interest (lacking e.g. Mn or Cr, which are important tracers of the enrichment as recently shown in Hitomi Collaboration 2017; Simionescu et al. 2018), the variety of abundances provides a good starting point for a meaningful study on the ICM and the demonstration of the X-IFU capabilities in this view. For every gas particle in the simulation, we traced the chemical composition and the fraction of each metal that is produced by the three enrichment sources (i.e. SN_{Ia}, SN_{cc}, AGB; see Biffi et al. 2017, 2018, for further detail). Each object is analysed at different redshifts, $z = 0.105$, 0.5, 1, 1.48, and 2, to assess the enrichment through time. Table 1 provides the characteristic radius, R_{500}, along with the mass, M_{500}, and the mass-weighted temperature, T_{500}, of the associated sphere of radius R_{500} for the entire cluster sample.

For each SPH particle, the output quantities provided by `GADGET-3` are used as input for our simulation. These include the position of the particle, $\underline{x}$, its 3D velocity in the observer's frame, $\underline{v}$, its mass density, ρ, its mass, m, its temperature, T, and the individual masses of the 15 individual chemical species X, μ_X, tracked in the simulations. The gas density n of each SPH particle is obtained by dividing ρ by m. The mass of each element of atomic mass number A_X is converted into abundances Z_X, expressed in solar metallicity units assuming the solar fractions $Z_{\odot,X}$ from Anders & Grevesse (1989). Abundances can be therefore written as

$$Z_X = \frac{1}{Z_{\odot,X}} \times \frac{\mu_X}{\mu_H \times A_X}, \quad (1)$$

with μ_H the hydrogen mass of the particle.

3. End-to-end simulations

In this section, we describe in detail the set-up of the pipeline used for the synthetic X-IFU observations, as well as the physical assumptions made in the simulations.

3.1. Synthetic X-IFU observations

Simulations of the cluster data set are carried out using the X-IFU end-to-end (E2E) simulator SIXTE[2] (Wilms et al. 2014), which creates realistic X-IFU observations. SIXTE uses as an input a specific SIMPUT file (Schmid et al. 2013) containing either all the emission spectra of the particles or directly a photon list, with the time, coordinates on the sky and energy of the emitted photons. This second approach is preferred for our simulations, as it exerts a lower computational demand, induced in the former by the unparallelised random generation of photons currently implemented in SIXTE (an example of the first approach is given in Roncarelli et al. 2018). SIXTE outputs are generated not only considering the instrumental spatial and spectral responses, but also incorporate other features from the detectors such as their geometry, vignetting and internal particle background.

3.1.1. Photon list generation

Each simulated cluster comes as a list of SPH particles, which may emit X-ray photons. To generate the photon list used in the E2E simulation, the particle emission is modelled by a collisional diffuse thermal plasma using the `APEC` code (Smith et al. 2001). More specifically, the `vvapec` model on `XSPEC` (Arnaud 1996) is adopted, as it can be parametrised according to the particles physical properties listed above, notably the individual abundances of each element. The corresponding atomic database used for the emission model is derived from `ATOMDB v3.0.9`. For the galactic absorption, the `wabs` model (Morrison & McCammon 1983) is preferred for computational speed, although more accurate absorption models do exist (e.g. `TBabs`, Wilms et al. 2000). For all four clusters

[2] `http://www.sternwarte.uni-erlangen.de/research/sixte/`

we fixed the column density to $n_{\rm H} = 0.03 \times 10^{22}$ cm^{-3}, which is a representative value for the latitudes at which most clusters shall be observed with X-IFU (Kalberla et al. 2005). Abundances are set to solar as per Anders & Grevesse (1989) and atomic cross-sections are taken as per Verner et al. (1996). The overall flux, F (in counts s^{-1} cm^{-2}), of each particle is computed using the vvapec normalisation $\mathcal{N}$ (emission-measure-weighted by the distance in units of cm^{-5}):

$$\mathcal{N} = \frac{10^{-14}}{4\pi[D_{\rm A}(1+z)]^2} \int n_{\rm e} n_{\rm p} {\rm d}V, \quad (2)$$

where $D_{\rm A}$ is the angular distance of the particle computed from its redshift z (derived from the speed of the particle and the cluster mean redshift), and V the particle volume. We considered a full ionisation of the intra-cluster gas with $n_{\rm e} = 1.2 n_{\rm p}$ ($n_{\rm e}$ and $n_{\rm p}$ being the densities of electrons and protons, respectively). These emission spectra are considered as probability density distribution function and normalised accordingly over the instrumental energy bandpass (i.e. 0.2–12 keV). For a fixed exposure time Δt, photons are drawn from the afferent probability distribution following a Poisson statistic of parameter $F \Delta t A$, where A is the total mirror area (taken at 1.4 m^2 at 1 keV, energy dependence of the effective area is included later on in SIXTE via the ancillary response function – ARF as explained below). Each newly created photon is added to the photon list with the sky coordinates of its parent particle (right ascension and declination).

The output product of this stage is a "complete" photon list (with their true energy and position) at the entrance of the telescope. This list is computed once for each cluster, and contains a large number of simulated photons (≥1 Ms). It is then sampled randomly by SIXTE to achieve smaller lists for more typical exposure times (e.g. 100 ks).

3.1.2. Observational setup

For each simulation, we consider an exposure time $\Delta t = 100$ ks over the entire X-IFU field-of-view. The complete photon list is used as input for the `xifupipeline` function of SIXTE, which samples the photon list accounting for the energy-dependence of the effective area to create the event list seen by the X-IFU detector over Δt. The pipeline accounts for the most up-to-date responses of the current baseline of the telescope (i.e. 15-row mirror modules corresponding to a mirror effective area of 1.4 m^2 at 1 keV[3]) and for a hexagonal detector array of 3832 micro-calorimeter pixels, more specifically Large Pixel Array 2 (LPA2) pixel configuration, developed for the X-IFU and described in Smith et al. (2016). Pile-up, telescope point spread function, vignetting and detector geometry effects are also included as function of the pixels corresponding off-axis angles. Finally, we verified that given the low count rates of our clusters (≤1 cts s^{-1} pix^{-1}), pile-up and cross-talk over the observation can be neglected (see den Hartog et al. 2018; Peille et al. 2018).

For each cluster, we simulated enough pointings to map the cluster spatially up to at least R_{500} (as required in the current science objectives for the X-IFU). This translates, for local clusters, into at least seven pointings. The corresponding event lists are then merged during post-processing to obtain a single event file.

[3] RMF: `athena_xifu_rmf_v20171107.rmf` | ARF: `athena_xifu_15row_onaxis_pitch249um_v20171107.arf`

Table 2. Parametrisation of the galactic foreground model used in the simulation with a `apec + phabs*apec` model.

Model	Parameter	Unit	Value
apec	T	keV	0.099
apec	Z		1
apec	z		0
apec	Norm	cts s^{-1} amin^{-2}	1.7×10^{-6}
phabs	$n_{\rm H}$	10^{22} cm^{-3}	0.018
apec	T	keV	0.225
apec	Z		1
apec	z		0
apec	Norm	cts s^{-1} amin^{-2}	7.3×10^{-7}

3.2. Foreground and background components

In addition, we accounted for the contribution of different foreground and background sources to ensure more representative observations.

3.2.1. Astrophysical foreground

The foreground emission is caused by the X-ray emission of the local bubble in which the solar system is embedded and by the Milky Way hot gaseous content. This component can be modelled by the sum of a non-absorbed and absorbed thermal plasma emission as specified in McCammon et al. (2002) and parametrised as per (Lotti et al. 2014, see Table 2). An additional normalisation constant over the entire model is used for versatility purposes, resulting in a total foreground model reading as `constant*(apec+phabs*apec)` in XSPEC. This component is folded into SIXTE using a SIMPUT file.

3.2.2. Cosmic X-ray Background

The cosmic X-ray background (CXB) component is due to the contributions of AGNs, star forming galaxies and active stars along the line-of-sight (Lehmer et al. 2012). A fraction of these sources will be resolved by the instrument as a function of its spatial resolution, and will be excised from the observations. Given the requirement on the spatial resolution for *Athena*/X-IFU (5″), 80% of the total flux of these point sources in terms of the integral of their log(N)/log(S) distribution should be resolved by the instrument (Moretti et al. 2003). For 100 ks exposure times, this translates into limiting fluxes of $\sim 3 \times 10^{-16}$ ergs s^{-1} cm^{-2} for the X-IFU. As the number of star forming sources is at least an order of magnitude lower at this flux, we only considered the AGN contribution in this study. The unresolved fraction of these point sources results in a diffuse background component, which we classically fitted using an absorbed power-law model during post-processing (McCammon et al. 2002).

We did not include AGN point sources in the inputs derived from the hydrodynamical simulations. Instead, to generate realistic CXB data, we drew a list of AGN sources with associated X-ray spectra by sampling the luminosity function of Hasinger et al. (2005) in the luminosity-redshift space, given the boundary conditions $L_{\rm X} \geq 10^{42}$ erg s^{-1} unabsorbed 0.5–2 keV rest-frame luminosity, $0 < z < 5$ and the size of the cosmological volume encompassed within a field-of-view. Each source is associated with a spectral energy distribution following templates described in Gilli et al. (2007), according to a distribution

of power-law indexes and intrinsic absorption column densities related to various levels of obscuration, as described in Gilli et al. (1999, 2007). Spatial distributions are fully random in the sky plane (no clustering). Further details about the procedure can be found in Clerc et al. (2018). This component is also included in the simulations using a SIMPUT file.

3.2.3. Instrumental background

The instrumental particle background is caused by interactions of high-energy cosmic rays and protons with the instrument structure, which create secondary particles in the soft X-ray band. Both primary and secondary particles can hit the detectors and be recorded as regular events. The X-IFU design includes an onboard cryogenic anti-coincidence detector (Macculi et al. 2016), which will ensure the required level of 5×10^{-3} counts s^{-1} keV^{-1} cm^{-2} over the 2–10 keV energy band (Lotti et al. 2017). The generation of this component is directly implemented within SIXTE.

For each of these three background components, we associate a flag on the photons in the event file. Thus, these specific events can be respectively masked to study background effects on the observations (notably of the internal particle background) or selected exclusively to generate background maps. Throughout this study, we assume that these background components have no systematic uncertainty. Systematic effects of the background knowledge on the observations are discussed Sect. 7 and are considered in more ample detail elsewhere (Cucchetti et al. 2018).

4. Data processing

In this section we describe the post-processing approach used in our simulations and its validation.

4.1. Source contamination

From the event list output of SIXTE, a first selection is made on the grading of the events, which is conditioned by the frequency of detection in a given pixel and defines the spectral resolution of the event. In practice, the grading procedure will occur in-flight using the onboard event processor depending on the time separation between events in the same pixel (Peille et al. 2018), similarly to the strategy implemented on the SXS (Seta et al. 2012). In this case, grading occurs automatically within the simulator and is available in the event list. Only high-resolution events corresponding to $\Delta E = 2.5$ eV, the required spectral resolution for X-IFU, were used. As the count rates of the clusters are low over the entire field-of-view ($\leq$1 count s^{-1} pix^{-1}), almost all events (throughput $\geq$99%) are high-resolution photons. Events with lower grading values are discarded in the rest of the study.

Using the selected events, we reconstructed raw brightness maps in counts as presented in Fig. 1. Beside the ICM, emission from other sources either present in the hydrodynamical simulations (i.e. strongly-emitting particles or clumps) or in the CXB can be observed. For the CXB, following the assumption by Moretti et al. (2003), we start by selecting the brightest simulated sources (with fluxes above $\sim 3 \times 10^{-16}$ ergs s^{-1} cm^2, see Sect. 3.2.2). After simulating CXB-only pointings to find their coordinates on the detector, these sources were excised automatically from the brightness maps by finding all the corresponding pixels above a 2σ threshold in counts with respect to the average count in neighbouring pixels in the event list. This way, the diffuse emission of the cluster is accounted for when masking a pixel. The average cut-off flux of the sources will be higher than $\sim 3 \times 10^{-16}$ over the full map. The lower the emission (e.g. in the outskirts), the closer the limiting flux of the excised point-sources will be to the threshold flux. Although this process would not be possible for real event files, we adopted this strategy to avoid potential biases related to specific point-source detection algorithms. Once these sources are removed, a final visual inspection was performed to remove any residual unexcised AGNs as well as any remaining visible point-like source which may be related to the hydrodynamical simulations.

4.2. Spatial binning

For the considered exposure time of 100 ks, single pixels do not always capture enough photons to allow the measurement of chemical abundances. We therefore grouped them into regions to increase the signal-to-noise ratio (S/N) adequately.

Two methods were considered to spatially bin the pixels: the `contbin` tool developed by Sanders (2006) and an adapted 2D Voronoï tessellation method by Cappellari & Copin (2003). Both methods were tested and a comparison showed no visible difference in accuracy of the pipeline (see also Appendix A or Sanders 2006 for a more detailed comparison). Unlike the Voronoï tessellation, contour binning offers better results in describing the radial-like shape of the cluster emission, and was therefore retained for this study. The `contbin` binning scheme can be directly applied to our cluster count maps to compute regions of equal S/N. The real exposure map (constant here) and the spatial mask of excised sources are also used as inputs of the algorithm. In addition, we chose to fix the aspect-ratio of the binning regions (so-called constrain filling-factor) to two, to avoid long radial filamentary regions, which could artificially mix spatially-uncorrelated structures (especially in low count rates areas such as the outskirts of the cluster).

The binning procedure operates on the count maps, which are dominated by the ICM bremsstrahlung emission, that is, the continuum of the spectral energy distribution. We further optimised our pixel binning in view of our scientific objective of measuring chemical abundances and to account for smaller local surface brightness variations atop the bulk of the cluster emission due to for example, clumps or bubbles. To do so, we first divided the surface brightness map into annuli centred on the brightest part of the cluster and containing roughly the same number of counts ($\sim$300 000). For each of these annular regions, the total spectrum of the annulus is fitted over the entire bandpass (0.2–12 keV) with a continuum-only `vvapec` model (metal abundances set to zero) to estimate the number of counts in the continuum for a given annulus $C_{\text{continuum only}}$. Although slightly overestimating counts (due to the presence of lines), this simple approach converges quickly. When abundances are left free then set to zero after the fit, strong lines are sometimes spaced by values of the order of the energy resolution of the instrument (due to strong bulk motion or clumping within the cluster), causing fits to converge on local minima or not at all. Even after convergence, differences from the previous method do not exceed $\leq$3%. The total number of counts on a pixel i (C_i) in the annulus is then re-scaled as follows:

$$\bar{C}_i = \left(1 - \frac{C_{\text{continuum only}}}{C_{\text{total annulus}}}\right) C_i \tag{3}$$

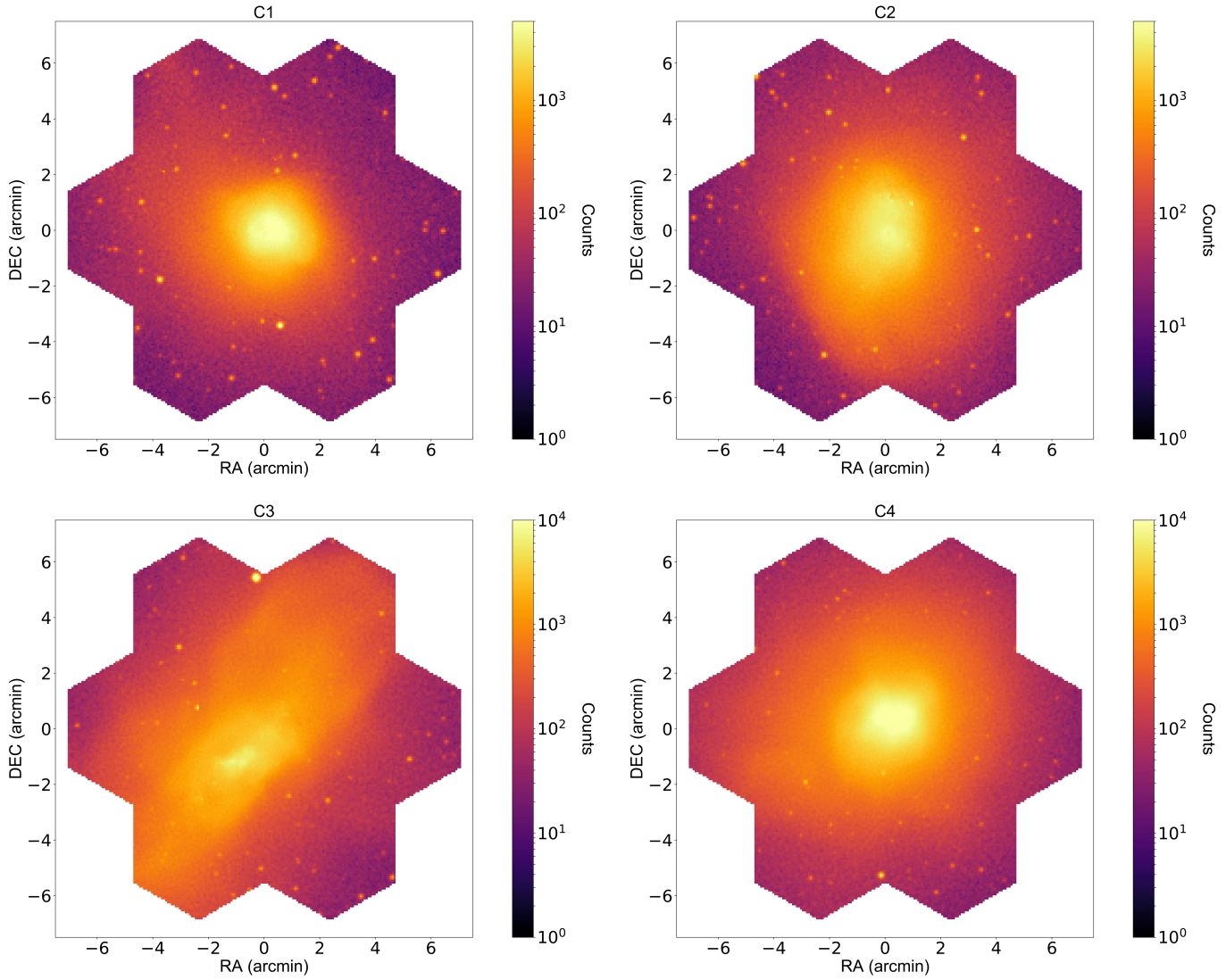

Fig. 1. Maps in number of counts per X-IFU pixel (249 μm pitch) for our clusters 1 to 4 (see Table 1) at redshift $z = 0.1$. Each mosaic is made of 7 X-IFU pointings of 100 ks each.

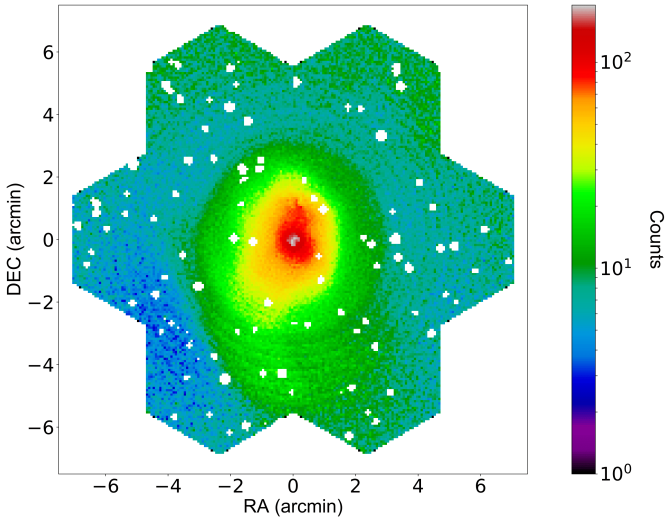

Fig. 2. Example of a continuum subtracted count map for cluster C2 used for spatial binning. Each of the ~27 000 pixels was re-scaled as explained in Sect. 4.2 to enhance density contrasts in the cluster.

The resulting template image for cluster C2 can be seen Fig. 2. This approach allows to define regions over a continuum subtracted image, enhancing the brightness fluctuations with respect to the azimuthal continuum model due to chemical elements emission and other local density contrasts. This template image is used exclusively for spatial binning. Accounting for its statistics, we requested a S/N level of 30 (900 counts in regions) to `contbin`. The resulting spatial regions are used over the full count maps to compute the corresponding spectra. Given a count ratio of ~100 between the template map and the full surface brightness map, we ensure that each spectrum has a high statistical significance, that is, $S/N \sim 300$. This represents a total of at least 90 000 counts per region, in other words three counts in each instrumental channel, allowing high significance regions for the fits.

This binning approach has been used throughout this paper when statistics allows it, that is, in the case of local clusters ($z \leq 0.5$). For high-redshift clusters, count rates are often insufficient to ensure statistically independent regions with a high-enough S/N. In this case, we followed the formulation of the *Athena* science objectives on chemical evolution of the Universe (specified in the introduction) and considered two radial annular regions, over 0–0.3 R_{500} and over 0.3–1 R_{500}.

4.3. Accounting for vignetting effects

Despite the narrow field-of-view of the X-IFU, the *Athena* telescope will introduce a slight vignetting effect in the observations (between 4 and 8% of the counts over the energy bandpass). Vignetting is simulated within SIXTE using tabulated values derived from ray-tracing simulations of the mirror assemblies (Willingale et al. 2014) and needs to be accounted for before attempting a spectral fit. To do so, for each binned region, the baseline X-IFU effective area of every single pixel is multiplied over the energy bandpass by the same vignetting function implemented in SIXTE accounting for the off-axis position of the pixel. The specific response function of the entire region is then obtained by averaging, in each energy channel, the response of each of pixels weighted by their respective number of counts.

4.4. Temperature and metal maps

The resulting unbinned spectra were fitted within XSPEC, using a log-normal likelihood minimisation (i.e. C-statistics adapted from Cash 1979), which is well adapted for Poissonian data sets (i.e. channels with low number of counts). An absorbed single-temperature thermal plasma model (i.e. `wabs*vvapec`) was fitted to the spectrum accumulated in each region. The column density $n_{\rm H}$ is fixed to its input value, while temperature, abundances of metals traced in the input numerical simulations, (see Sect. 2), redshift and normalisation are left free. As a first approach, all parameters (metallicity for each of the 12 chemical elements $Z_{\rm X}$, temperature, redshift and normalisation) are fit simultaneously over the 0.2–12 keV energy band (broad-band fit). The background components are accounted as an additional model, as described in Sect. 3.2. Their spectral shapes is assumed to be perfectly known whilst the normalisation of each three component is set free (either the model norm for the CXB and instrumental background, or the multiplicative constant factor for the foreground, see Sect. 3.2.1). Using the best fit results provided by XSPEC, we were able to construct spatial maps for each physical parameter over each full synthetic observation.

4.5. Input parameter maps

To estimate the goodness of the fit, the output maps were compared to weighted input maps, reconstructed from the input numerical simulations using the same spatial binning regions, j, than the outputs. The adopted weighting scheme depends on the considered parameters. For instance, emission-measure-weighted quantities are expected to be more representative than mass-weighted schemes (Biffi et al. 2013), especially for abundances. The value of the input emission-measure-weighted parameter P in the region j therefore reads as

$$P_j = \frac{\Sigma_i P_i \mathcal{N}_i}{\Sigma_i \mathcal{N}_i}, \tag{4}$$

where $\mathcal{N}_i$ is given by Eq. (2) for each SPH particle i contributing to the spatial region j.

For the temperature, it has been long known that emission-measure- or mass-weighted schemes do not match fitted quantities (Gardini et al. 2004; Rasia et al. 2006, 2008). A known method to account for this bias is to use the so-called spectroscopic temperature weighting, introduced by Mazzotta et al. (2004), which translates here into

$$T_j = \frac{\Sigma_i \mathcal{N}_i T_i^{+0.25}}{\Sigma_i \mathcal{N}_i T_i^{-0.75}}. \tag{5}$$

Notably, we verified that the use of spectroscopic temperature input maps indeed reduced the biases of the fitted temperature with respect to the emission-measure-weighted input maps (see also Appendix B). We note that this method is particularly suited to high-temperature regions (≥3 keV), well represented in the central parts of our clusters (see also Fig. 3, upper central panel), but may be limited towards the cluster outskirts or in the case of cooler systems. Although the low-temperature regions could be processed more accurately with the extended method presented in Vikhlinin et al. (2006), since our outer regions are very large (small number of counts per pixel) and relative differences remain within statistical error bars of the XSPEC fit, we use the implementation of the spectroscopic temperature presented in Mazzotta et al. (2004) for all our regions.

4.6. Assessment of systematics

For a physical parameter P, the goodness of the reconstruction is evaluated using two methods. First, the deviation of the fitted value is evaluated in terms of its relative error distribution with respect to the weighted input value (i.e. $\Delta P_j = (P_j - P_{{\rm in},j})/P_{{\rm in},j}$) over the various spatial regions j. A priori, if no systematic effect is present in the pipeline, the relative error distribution for P should be Gaussian (if the number of regions is sufficiently high) and centred, with a standard deviation $\sigma_{\Delta P}$ of the same order as the averaged normalised value of the statistical error, $\mu_{\rm fit}$, derived from the XSPEC fits (see Appendix B for a more generic estimation for non-Gaussian cases). This approach is mainly used to determine the presence of biases in the reconstruction by fitting the relative error distribution using a Gaussian, accounting for the corresponding errors σ_j derived from XSPEC. For this fit, clear outliers with relative errors above 100% (one or two regions overall) are removed for consistency. As a second test, we compute in each region j the value $\chi_j = (P_j - P_{{\rm in},j})/\sigma_j$, which shows the goodness of the fits in terms of the fitting errors and should follow a χ distribution. By computing the reduced value of the distribution, $\chi^2_{\rm red}$, we can estimate the overall goodness of the fit with respect to the statistical errors derived from XSPEC.

Using the first test, we noticed that the broad-band fit initially used introduced biases on the temperature and the abundances recovered for local clusters (∼10%), when compared to emission-measure-weighted quantities. The use of spectroscopic temperatures (Eq. (5)) decreases the biases on temperatures. However, biases on abundances are not accurately corrected by this new approach, underlining a bias in the overall fitting procedure. For our data analysis therefore, we switched to a multi-band fit of the spectrum (following Rasia et al. 2008) including a velocity broadening component to the lines to account for variability and mixing along the line-of-sight (`bvvapec` model on XSPEC). The fit is performed as follows for each region:

- As a first step, since the cluster sample is relatively hot (≥4 keV in the centre) the temperature is recovered from the high energy band (3.5–12 keV), then fixed.
- Iron metallicity, $Z_{\rm Fe}$, is recovered by a subsequent broad-band fit (i.e. 0.2–12 keV) to estimate the contribution from both the K and L complex, then fixed.
- Metallicities of other elements are then computed by fitting specific energy bands (for redshift $z \leq 0.1$):
  - Abundances for C, N, O, Ne and Na are fitted over the 0.2–1.2 keV bandpass.
  - Abundances for Mg, Si, Ar and Ca are fitted over the 1.2–3.5 keV bandpass.

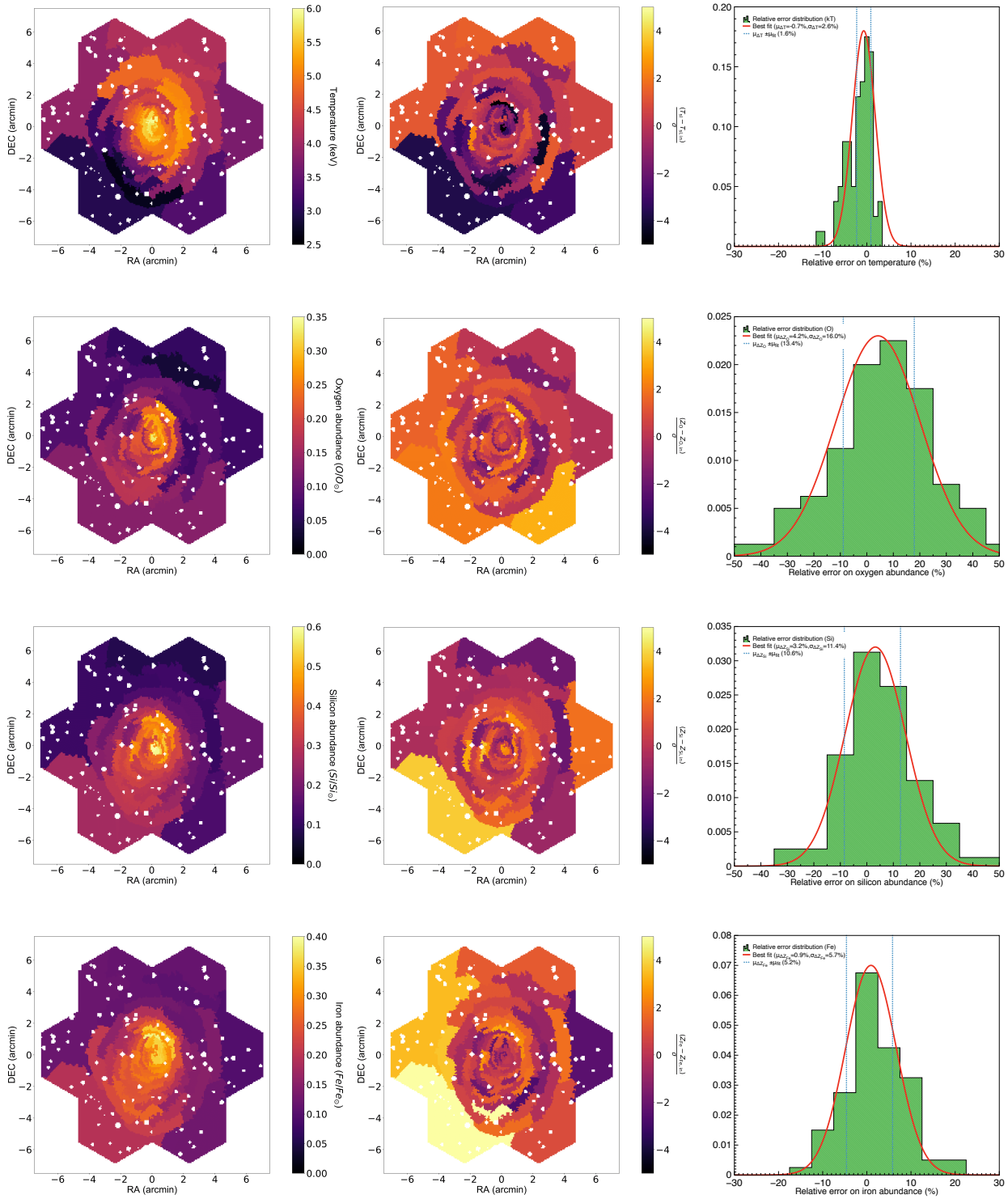

Fig. 3. *Left*: reconstructed ICM parameters maps for C2 at $z \sim 0.1$ using the multi-band fit presented in Sect. 4.6 with $S/N \sim 300$ (~90 000 counts per spatial bin). *Middle*: distribution within the regions j of $(P_j - P_{\mathrm{in},j})/\sigma_j$ indicating the goodness of the fit for each region in terms of σ_j (see Sect. 4.6). *Right*: relative error distribution across all spectral regions (green histogram). The red solid line pictures the Gaussian best fit. The vertical blue dashed lines are set at the mean value of the fit errors (see Sect. 4.6). *From top to bottom*: spectroscopic temperature T_{sl} (in keV), emission-measure-weighted abundances of oxygen (O), silicon (Si) and iron (Fe) (with respect to solar).

  - Abundance of Ni is finally fitted over the 0.2–12 keV bandpass.
- With all other parameters fixed, redshift, velocity broadening and normalisation are recovered with a broad-band fit.

This multi-band post-processing approach is retained in the rest of the paper for local clusters (notably to create Figs. 3 and 4), along with the comparison of the fitted output temperature with spectroscopic input temperature maps. This technique is used under the caveat that despite some fitted parameters are fixed, errors are propagated correctly throughout the fit. Although not perfectly true, this effect remains very limited with respect to the total systematics and to our level of statistics, ensuring a safe

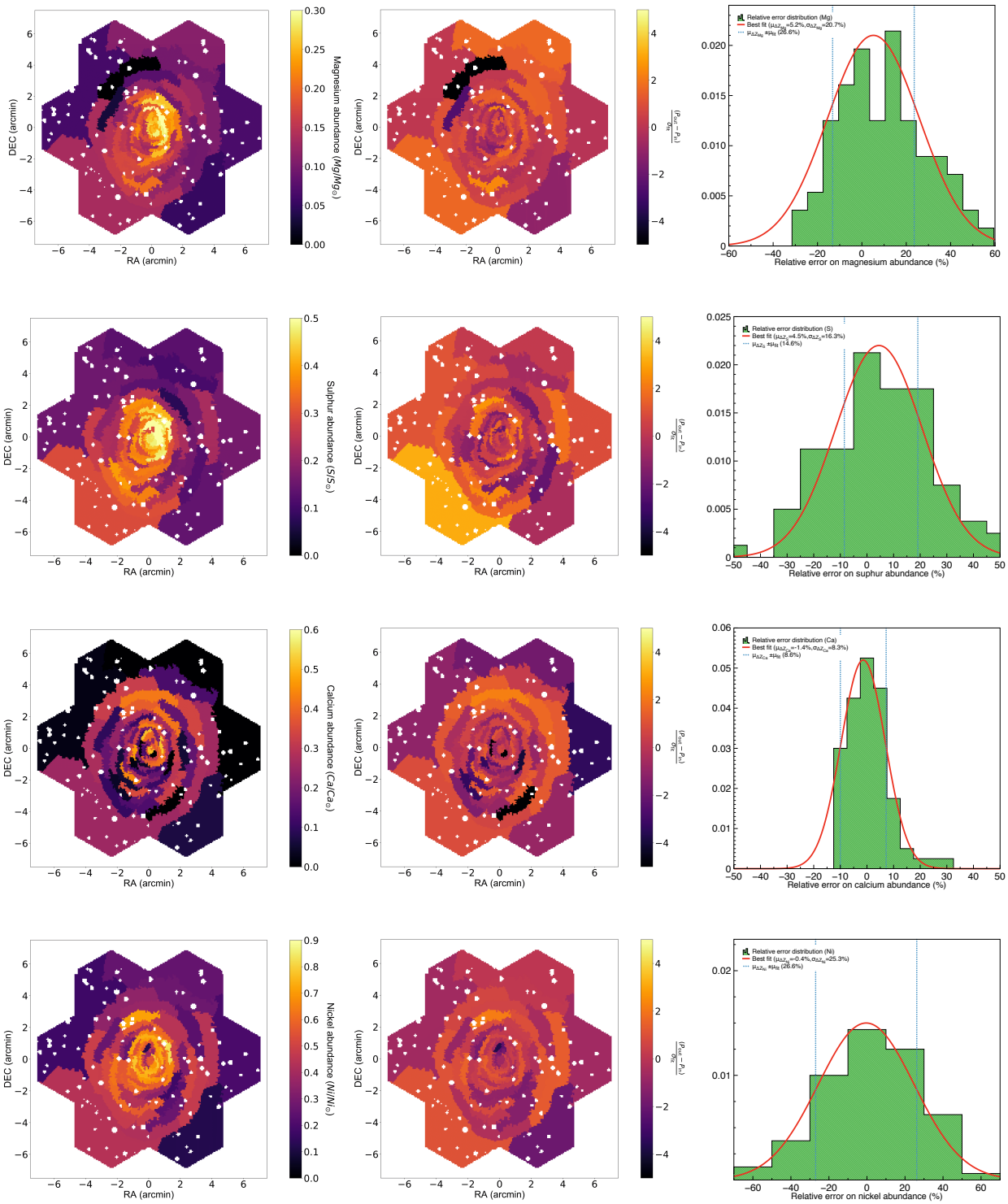

Fig. 4. As Fig. 3, *from top to bottom*: Magnesium, sulphur, calcium, and nickel abundances (with respect to solar).

application of this method. For all the fitted parameters, we computed the mean, $\mu_{\Delta P}$, and standard deviation, $\sigma_{\Delta P}$, of the relative errors between the output and input values to the mean of the XSPEC fit errors, $\mu_{\rm fit}$ (see right panels of Figs. 3 and 4 for an illustration in the case of cluster C2) and by computing $\chi^2_{\rm red}$ for each parameter. Results of this comparison for the main parameters are given in Table 3. In all cases, the mean fit error is consistent with the standard deviation of the error distribution $\sigma_{\Delta P}$, thus excluding large systematic effects. However, despite these changes, small biases of the order of a few percent ($\lesssim$5%) are still visible. This is particularly true for the normalisation and the abundances of low mass elements (e.g. O, Si), in which an underestimate of the normalisation directly results in an overestimate of these abundances (see also in Fig. 3). All these systematics are however well inside the statistical deviations. Using the second test, we notice that some of the $\chi^2_{\rm red}$ are not consistently recovered, especially for the normalisation. This can be partially explained by the presence of some outlier regions in the

Table 3. Values of the mean $\mu_{\Delta P}$ and standard deviation $\sigma_{\Delta P}$ of the best fit of the relative error distributions for the main physical parameters in cluster C2, over 80 spatial regions, with the average XSPEC error of the fits $\mu_{\rm fit}$.

	$\mu_{\Delta P}$ (%)	$\sigma_{\Delta P}$ (%)	$\mu_{\rm fit}$ (%)	$\chi^2_{\rm red}$	$\tilde{\chi}^2_{\rm red}$	$N_{\rm bad}$
$T_{\rm sl}$	−0.71	2.60	1.64	3.93	1.38	4
z	0.00	0.04	0.05	1.54	1.09	3
O	4.23	16.0	13.4	1.50	1.50	2
Mg	5.22	20.7	18.4	2.84	1.01	1
Si	3.22	11.4	10.6	1.78	1.37	3
S	4.48	16.3	14.6	1.52	1.41	1
Ca	−1.42	8.27	8.62	2.69	1.06	3
Fe	0.94	5.73	5.22	2.71	1.49	2
Ni	0.39	25.3	26.6	0.99	0.92	1
$\mathcal{N}$	−6.5	3.68	3.74	4.11	2.83	7

Notes. The corresponding values of $\chi^2_{\rm red}$ and the corrected $\tilde{\chi}^2_{\rm red}$ are also given, along with the number of outliers regions $N_{\rm bad}$ removed to compute $\tilde{\chi}^2_{\rm red}$.

fit, which strongly affect this computation. When the $N_{\rm bad}$ outliers regions for which the parameter is outside a $3\sigma_{\rm fit}$ level, are removed, the new reduced chi-squared, $\tilde{\chi}^2_{\rm red}$, shows more consistency (Table 3), indicating a good agreement between the fitted maps and the input distributions.

Small errors remaining after post-processing can be attributed to some degeneracy between fitted parameters (≥15 here) and the spatial binning, which regroups in a relatively large region (a few arcmin^2 in the outskirts) many different physical structures (see also the discussion in Appendix B). The choice of a single temperature model could also introduce some bias accounting for the complex structure over the line-of-sight. Two- or multi-temperature models could bring slight improvements, especially towards the centre of the object. Such schemes were investigated over single regions but showed no significant improvement over the entire pointings. The assessment of the improvements introduced by a multi-temperature scheme over the entire temperature distribution as well as the use of line-ratio techniques (Hitomi Collaboration 2018c) will be the object of future improvements of our pipeline, and shall be addressed in future studies. Finally, we also found that small effects related to statistics, to the XSPEC fitting procedure, and to the weighting schemes of the input maps partially explain these residual deviations (refer to Appendix B for a more ample and detailed discussion on the pipeline validation).

5. Properties of the ICM of local clusters

In this section we present the results on the ICM properties recovered by the X-IFU, starting from the interpretation of the raw output maps to larger studies involving the entire cluster sample.

5.1. Physical parameters maps

We show in Fig. 3 the reconstructed maps for cluster C2 at $z \sim 0.1$ for the spectroscopic temperature $T_{\rm sl}$ and the abundances of Fe, Si, and O. Those for Mg, S, Ca, and Ni are shown in Fig. 4. Similar maps for the other three clusters are provided in Appendix C. Beyond the recovery of abundances, the physical parameter maps and their combination provide a wealth of information on the dynamics of the cluster. For instance, we see from the temperature map of cluster C2 the presence of a hot bubble on the western part of the cluster and a cold arc in the south-eastern region. Interestingly, we also notice the correlation between the presence of low-mass elements (e.g. O, Si) and the temperature of the ICM. Several clumps and small groups are also visible on some of the clusters and cluster C3 exhibits a merging activity with a very bright central object. After post-processing, the redshift of each region is also recovered with excellent accuracy from the fit over a large number of lines and spectral features. The redshift map, once converted into velocities using the mean redshift of the cluster, provides a projected map of the bulk motions within the ICM. Likewise, the velocity broadening of the lines is recovered in the fit. Measuring both bulk motion and turbulent velocities through line shifts and line broadening respectively, is another main scientific objective for X-IFU. This, however, goes beyond the scope of this paper and we redirect the reader to Roncarelli et al. (2018) for an illustration.

5.2. Metallicity profiles of the ICM

The hierarchical formation of galaxy clusters along with processes of production and dispersion of chemical elements within stars and galaxies should lead to self-similar global abundance profiles. The value of the metallicity in the outskirts will depend notably on the enrichment of the intergalactic medium prior to the halo formation (see, e.g. Biffi et al. 2017).

To investigate the capabilities of the X-IFU in determining abundance profiles over the cluster data set, we computed the radial profiles at $z \sim 0.1$ for some of the main chemical elements: O, Mg, Si, and Fe (Fig. 5). We find that the values of abundance profiles are consistent across the sample, showing a peak near the centre (i.e. up to $0.1\,R_{500}$) and a decrease towards a constant value between $\sim 0.1/0.2\,Z_\odot$ out to R_{500}. Cluster C3 shows however values of metallicities systematically higher than the others, which could be caused by the ongoing merging activity visible in Fig. 1 (Bottom left) and possibly due to AGN feedback. This dynamic activity also creates a more difficult line-of-sight distribution of the parameters, making the XSPEC fit less accurate. This is particularly the case near the outskirts, where background contributions become relevant and regions are large, thus increasing the deviations from the input maps (especially for iron). These measured profiles, and notably their constant metallicity value in the outskirts, suggest as discussed in Biffi et al. (2017), Truong et al. (2018), that the enrichment of the cluster in the hydrodynamical simulations pre-dates the large infall towards the central object and is mainly determined by the early enrichment mechanisms of the Universe. Other independent hydrodynamical simulations (Vogelsberger et al. 2018) and recent observational results (e.g. Ezer et al. 2017; Mernier et al. 2017; Simionescu et al. 2017) also argue in favour of this paradigm.

Overall, the recovered metallicities are consistent with the input metallicity profiles within 3σ (most well within 1σ), showing the power of the X-IFU in recovering the properties of the ICM even for typical 100 ks exposures. This analysis can be compared to a similar study performed by simulating 200 ks cluster observations with *XMM-Newton*/EPIC MOS1 and MOS2 for the same chemical species (see Rasia et al. 2008 for more details, notably Fig. 4) or to current observational data using the CHEERS catalogue (Mernier et al. 2017). For typical exposure times, the X-IFU will provide accurate measurements of the main metallic content of the ICM, enough to reduce the uncertainties on current observations, even for less abundant elements. In addition to the individual profiles, the overall

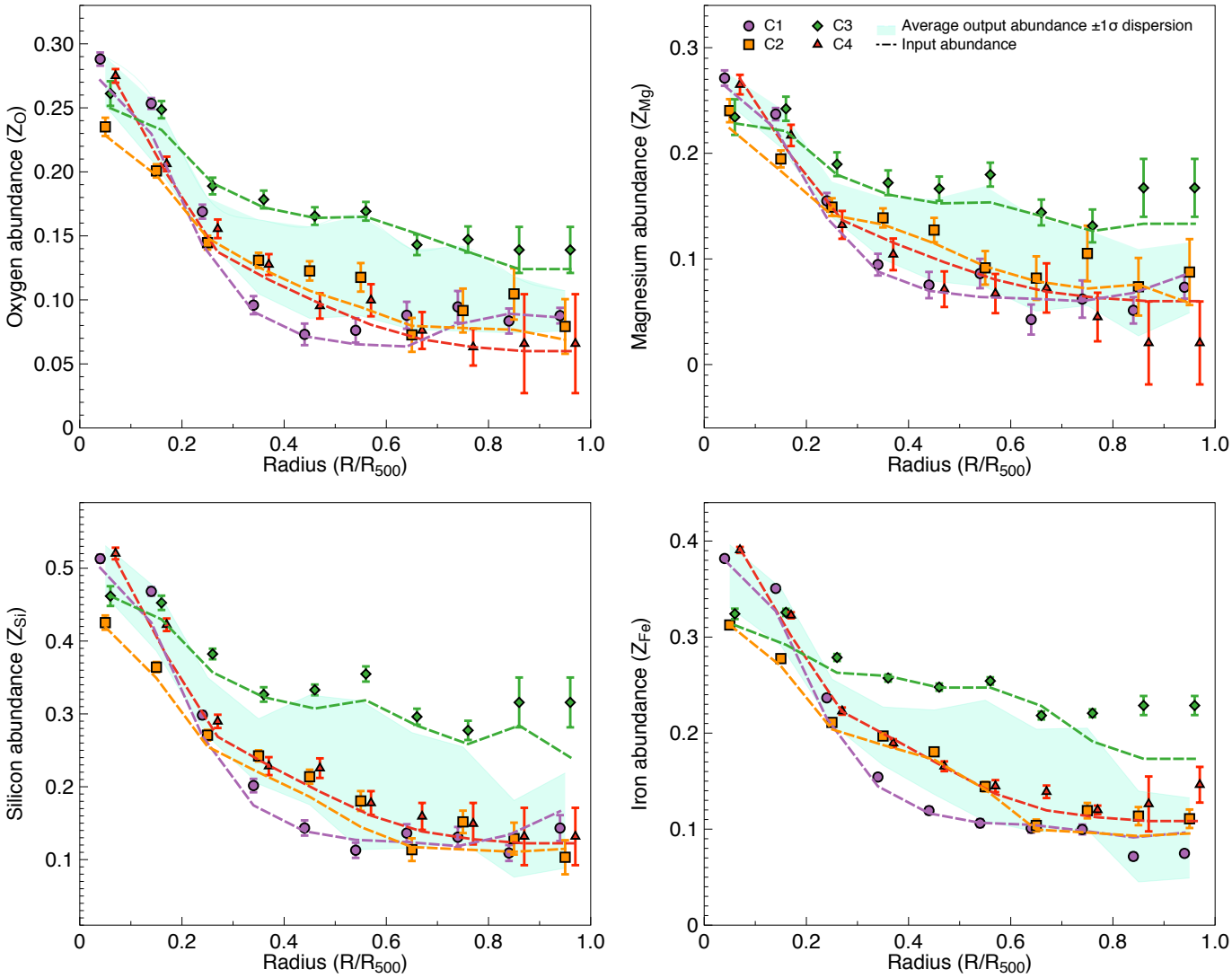

Fig. 5. Best-fit values of the metallicity as a function of radius, up to R_{500} ($0.1\,R_{500}$ bins) for the entire sample (C1 – purple dots, C2 – orange squares, C3 – green diamonds, C4 – red triangles). *From top left to bottom right*: oxygen, magnesium, silicon and iron abundances with respect to solar. The dashed lines represent instead the profile of the emission-measure-weighted input abundances using the same colours. The cyan shaded envelope represents the $\pm1\sigma$ dispersion of the recovered output metallicity for the entire sample. Points are slightly shifted for clarity.

dispersion across the sample (Fig. 5) is consistent with the average emission-measure-weighted input distribution. However, the sample considered here is relatively small, and the scatter of our results remains significant. Namely, we see that the dynamic behaviour of clusters C3 affects the overall scatter of the sample, otherwise similar for the other three objects (C1, C2, and C4). The values of iron abundance in the outskirts found in this study (between $0.1/0.2\,Z_\odot$) are somewhat lower than measured iron abundances in the outskirts, which range typically around $0.2\,Z_\odot$ (Werner et al. 2013; Mantz et al. 2017; Urban et al. 2017). Values remain however consistent with the hydrodynamical inputs, demonstrating that the X-IFU is able to recover the intrinsic physical parameters used for the clusters. The projection scheme adopted here also provides lower results than for example, emission-weighted schemes, in which the strongest emission regions (hotter and/or with more metal) will enhance the overall contribution.

Once in orbit, the X-IFU will probe a much larger number of galaxy clusters (≥10 per mass and redshift bins), therefore reducing the sample variance of these profiles even further, especially near the outskirts of the clusters. A more accurately constrained scatter will provide important information on the metallicity distribution of the ICM and firm observational confirmation of the nature of the enrichment scenario during the early phases of the Universe. These results highlight the sensitivity of the X-IFU to constrain with high accuracy the chemical enrichment pattern in cluster outskirts, and, therefore, to fully exploit its potential as a fossil record of the star formation history and feedback in the proto-cluster ecosystem.

5.3. Constraints on the chemical enrichment model

As chemical elements are trapped within the ICM, they represent a fossil record of the integrated history of chemical enrichment of the cluster. Strong constraints on the relative contribution of the various enrichment mechanisms (notably SN_{cc} and SN_{Ia}) could be given by accurate measurements of abundance

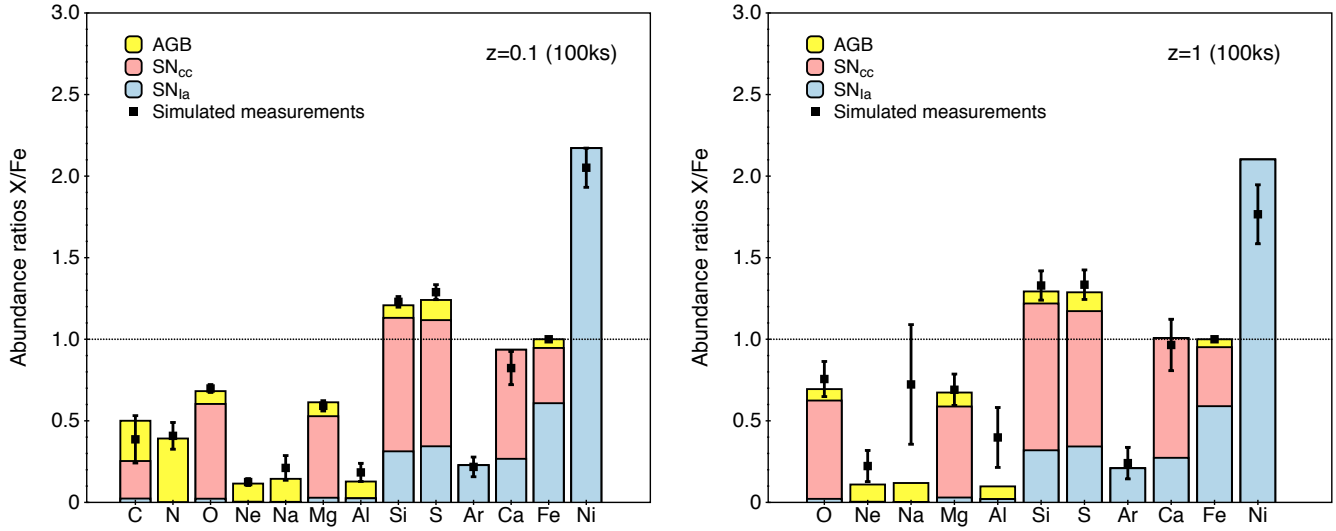

Fig. 6. Average abundance ratio with respect to iron within R_{500} at $z \sim 0.1$ over the cluster sample for $z = 0.1$ (*left*) and $z = 1$ (*right*) recovered using 100 ks observations. The input abundance ratio are shown as histogram bars filled with the respective contributions of SN_{Ia} (blue), SN_{cc} (magenta) and AGB stars (yellow), computed from the outputs of the hydrodynamical simulation presented in Sect. 2. For $z = 1$, carbon (C) and nitrogen (N) are not shown as the lines are outside the energy bandpass of the instrument (not fitted).

Table 4. Mean abundance ratio with respect to iron within R_{500} at $z \sim 0.1$ for each cluster in the sample.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
C/Fe	0.48 ± 0.34	0.35 ± 0.24	0.43 ± 0.28	0.48 ± 0.31
N/Fe	0.32 ± 0.12	0.40 ± 0.15	0.52 ± 0.30	0.57 ± 0.20
O/Fe	0.73 ± 0.04	0.68 ± 0.04	0.67 ± 0.06	0.69 ± 0.06
Ne/Fe	0.10 ± 0.03	0.12 ± 0.04	0.14 ± 0.06	0.20 ± 0.06
Na/Fe	0.17 ± 0.12	0.34 ± 0.24	0.20 ± 0.13	0.21 ± 0.17
Mg/Fe	0.63 ± 0.05	0.58 ± 0.06	0.59 ± 0.08	0.53 ± 0.08
Al/Fe	0.16 ± 0.08	0.18 ± 0.10	0.21 ± 0.18	0.20 ± 0.13
Si/Fe	1.24 ± 0.06	1.18 ± 0.05	1.27 ± 0.08	1.27 ± 0.09
S/Fe	1.33 ± 0.08	1.30 ± 0.08	1.26 ± 0.12	1.18 ± 0.12
Ar/Fe	0.19 ± 0.09	0.22 ± 0.11	0.20 ± 0.15	0.30 ± 0.17
Ca/Fe	0.94 ± 0.20	0.71 ± 0.16	0.71 ± 0.24	0.99 ± 0.25
Ni/Fe	2.12 ± 0.20	1.91 ± 0.21	2.11 ± 0.29	2.12 ± 0.32

ratios of elements within clusters. Using the small sample at our disposal, we estimated the capabilities of the X-IFU in recovering this information in the input hydrodynamical simulations within a radius of R_{500}. A first noticeable result is that for the entire cluster sample, recovered abundance ratios are consistent for all the elements (see Table 4). Using the average abundance ratio profile and the corresponding production yields of each element in the input hydrodynamical simulations, we computed the input fraction of each of the enrichment mechanisms (SN_{cc}, SN_{Ia}, or AGB) and compared it to the overall output abundance ratio given by our E2E simulations (Fig. 6, left). For this study a small fraction of the outlier regions ($\lesssim$5%) was excluded, as results clearly showed incompatible values with respect to emission-measure-weighted inputs for rare elements (N, Na, and Al). All of the elements present in the simulations are comparable to the inputs within their statistical error bars. Most of all, the ratios of the main elements of the ICM (e.g. O/Fe, Mg/Fe, and Si/Fe) are very accurately recovered with a significance of the detection $\geq$10 (i.e. the ratio between the value and the error). Rarer elements (typically Ne, Ar, S, Ca) are also very consistent with the hydrodynamical simulations. Less abundant elements (Na and Al) have looser constraints in the fitted regions due to the considered exposure time (low S/N of the lines) and seem slightly overestimated in their reconstruction.

With the small sample used here, we have demonstrated that the X-IFU is able to provide robust estimations of the abundance ratios. Given the low errors on the measurements, these results can be used to distinguish the contributions of the various mechanisms at play in the cosmological simulations, by comparing them to metal production theoretical models. Notably, using elements produced by single mechanisms (e.g. Ne, Na for AGB, or Ar, Ni for SN_{Ia}), accurate measurements of abundance ratios will provide strong constraints to the IMF and the contribution of each mechanism at local redshift. The ability to recover the corresponding supernovæ yield and to distinguish between multiple other models shall be addressed in a forthcoming study. The current observational strategy of the X-IFU plans to use at least 40 clusters of galaxies to investigate the chemical enrichment of the Universe. With such a large sample and by giving unprecedented information on other rare elements (e.g. Mn, Co, that could not be tested here since these metals are not separately traced in the hydrodynamical simulations), the X-IFU will undoubtedly provide new constraints to the global chemical enrichment models.

6. Chemical enrichment through cosmic time

Element abundances in local clusters embed the integrated chemical enrichment of the Universe up to this day. However, to understand how and when the ICM was enriched, the evolution of production sources with time and how the overall enrichment processes relate (e.g. the star formation history and the initial mass function) must be assessed. To do so, we analysed synthetic observations of the same four clusters taken at different stages of their evolution, hence considering five redshift values up to $z = 2$. We chose to keep a realistic exposure time fixed to 100 ks, regardless of the redshift. This exercise tested the capabilities of the X-IFU in a regime of low statistics, thus preventing the full spatial analysis presented above. For the highest redshift

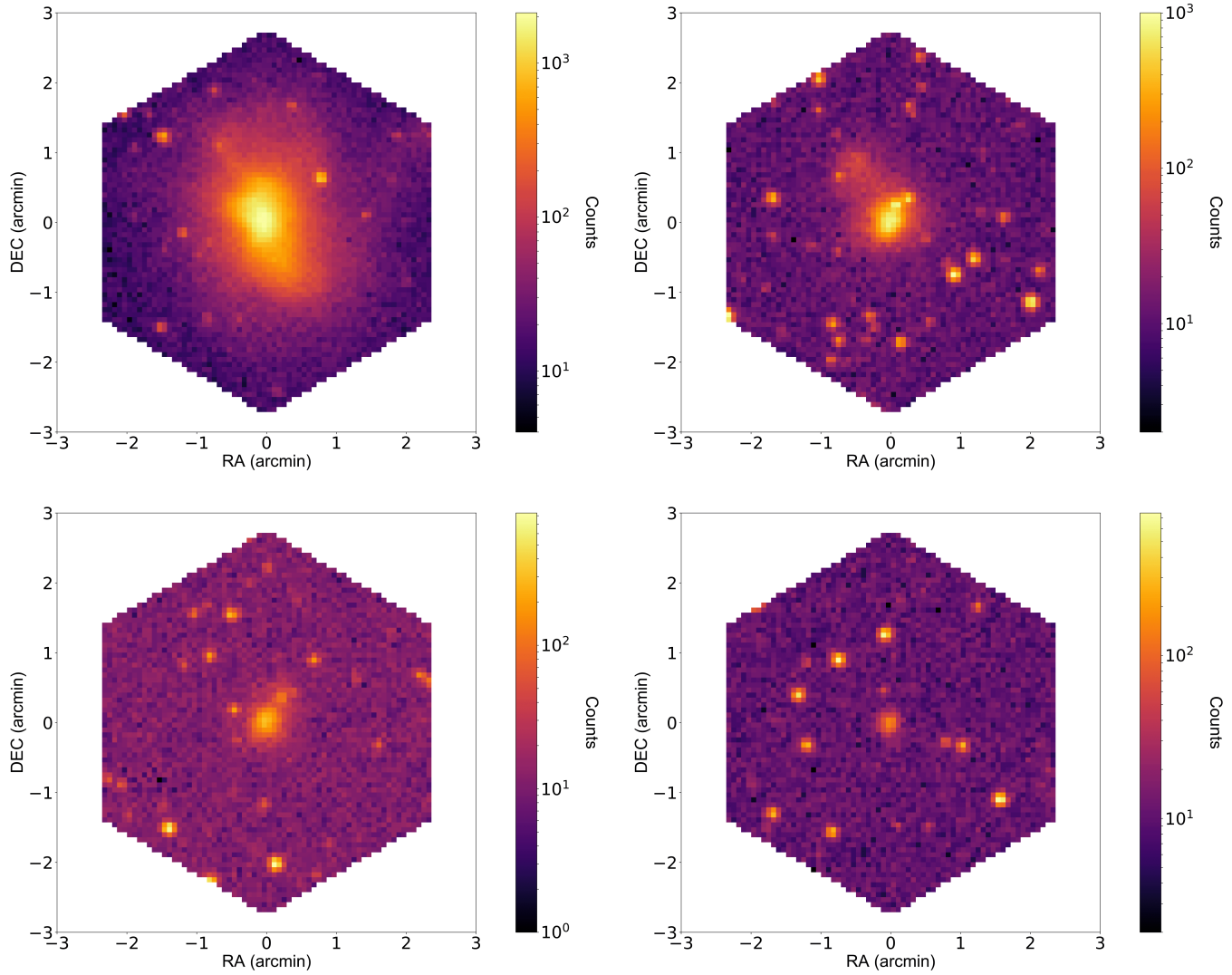

Fig. 7. *From top left to bottom right*: maps in number of counts per X-IFU pixel (249 μm pitch) for cluster C2 (see Table 1) simulated with the end-to-end simulator SIXTE for redshift $z \sim 0.5$ (*top left*), 1 (*top right*), 1.5 (*bottom left*) and 2 (*bottom right*), for an exposure time of 100 ks.

clusters we complied strictly to the definition of the *Athena* science case on chemical abundances and performed measurement in two different annuli from the cluster centre between 0–0.3 R_{500} and 0.3–1 R_{500}.

Figure 7 shows the mock surface brightness maps for C2 at the various redshift snapshots and illustrates the assembly history of halos through, for example, merging events. Figure 8 shows the evolution of the mean cluster abundance over our whole sample as a function of the redshift in the two aforementioned annular regions for O, Mg, Si, and Fe. Despite the lower source-to-background level of some objects, input metallicity values are recovered accurately within the statistical uncertainties of the measurements even for high-redshift clusters, although for $z \geq 1$, error bars start to be significant for elements such as O and Mg. In the case of low-mass elements, abundances are not measurable up to a redshift of $z = 2$, as lines are redshifted outside the instrument energy band (e.g. O and Mg for $z \geq 1.5$) or are too weak to be disentangled from the foreground and the background (e.g. Si at $z = 2$). As expected, measurements in the central parts of the cluster are more accurate, due to the higher level of background in the outskirts with respect to the cluster emission, especially for $z \geq 1.5$.

Through these measurements, we find that in the central parts of the cluster, metallicity hardly changes across time, even at a redshift of $z = 2$, once again consistently with the analysis by Biffi et al. (2017). This indicates that most of the enrichment occurs in the early days of the cluster. Interestingly, we notice that the iron abundance in the centre of the cluster slightly decreases with redshift, which could be explained for instance by an increase in time of iron production mechanisms (e.g. SN_{Ia}) or by the time delay with which long-lived SN_{Ia} release Fe. Abundances in the outskirts show a similar trend, with near-constant values up to local redshift values. Similar observational evidences, as for example, reported in Ettori et al. (2015) for iron abundance are consistent with these conclusions. The dynamic history of the clusters (mergers, shocks) is visible over time (data points are taken from the same cluster at different time steps), displaying local peaks of abundance, for example, C3 at $z = 1$ or C4 at $z = 0.5$. Given the sparse number of redshift points, turbulence or mixing within the cluster (whose eddy turn over timescale is of the order of a few Gyr over scales of ~1 Mpc for typical ~500 km s^{-1} velocities) create a more homogeneous distribution of metals in the structure, returning abundances to typical values after mergers.

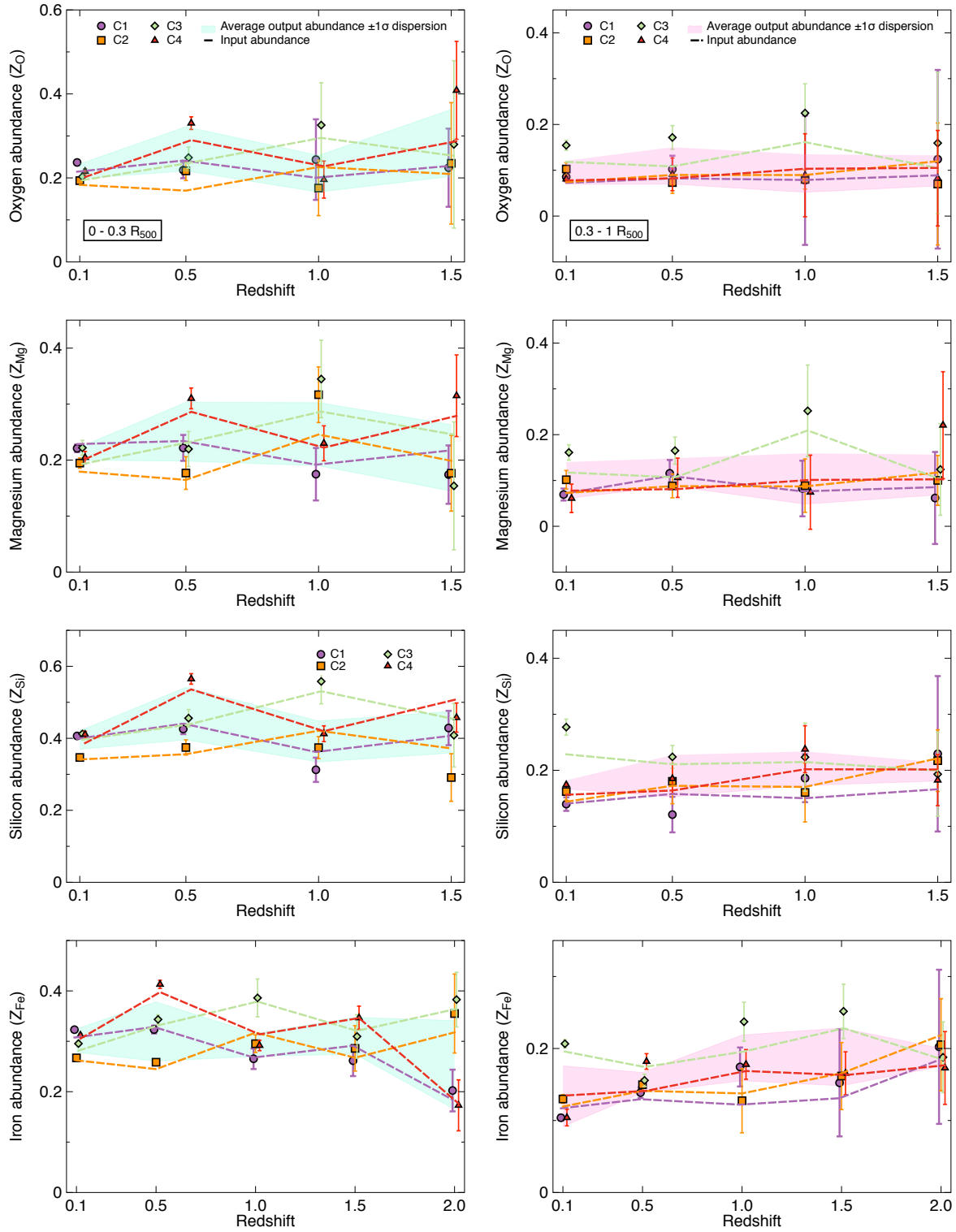

Fig. 8. Evolution of the average abundance of the cluster sample (C1 – purple dots, C2 – orange squares, C3 – green diamonds, C4 – red triangles) recovered via XSPEC as a function of the redshift between 0–0.3 R_{500} (*left*) and 0.3–1 R_{500} (*right*). *From top to bottom*: oxygen, magnesium, silicon and iron abundances with respect to solar. The dashed lines represent the profile of the emission-measure-weighted input abundances using the same colours. The cyan (resp. magenta) shaded envelope represents the $\pm 1\sigma$ dispersion of the output metallicity over the sample. Points are slightly shifted for clarity.

Using the accuracy of the X-IFU abundance measurements for high-redshift objects, we can also analyse changes in the metal enrichment mechanisms by performing a similar study as the one presented in Sect. 5.3. In the case of redshift $z = 1$ (see Fig. 6, right) for 100 ks observations, we find that the X-IFU will still be capable of accurately recovering abundance ratios within R_{500} with excellent accuracy. Most of the main elements have in fact no significant changes between both redshift values, consistently with our previous conclusions. In the case of high-redshift objects however, low-mass elements such as carbon and nitrogen

can no longer be detected (lines outside the energy bandpass) and rarer elements (e.g. Ne, Na, and Al) have large uncertainties due to the low S/N of the observations. Ni also tends to be underestimated (mostly in the outskirts) likely due to the low S/N of the line with respect to the high-energy background. This calls for better-adapted exposure strategies to optimise the results for distant objects and further investigate the chemical enrichment across cosmic time.

7. Summary and discussion

In this paper, we have addressed the feasibility of constraining the chemical enrichment of the Universe, which will be one of the key science objectives and a main driver of the performances of the future mission *Athena*. Notably, we investigated and quantified the capabilities of the X-IFU in accurately recovering metal abundances of the ICM across cosmic time. To this end, we developed a full end-to-end pipeline, which creates synthetic X-IFU observations using the instrument simulator SIXTE. We used as input of this pipeline a sample of four clusters generated using state-of-the-art cosmological simulations presented in Rasia et al. (2015) and Biffi et al. (2017) to create realistic event lists. All the relevant instrumental effects such as the convolution of spectra with the instrument spatial and spectral responses, realistic sources of foreground and background, and detector geometry were also included to obtain observations as realistic as possible.

The sample of four clusters was simulated at five different redshift values, for a fixed exposure time of 100 ks in order to achieve abundance measurements out to R_{500}. The accuracy of the pipeline was quantified by comparing our synthetic observations to weighted inputs quantities (e.g. spectroscopic temperature, emission-measure-weighted abundances). We find that a straightforward approach of a broad-band fit created systematic biases above 10% in a number of physical parameters. Rather a multi-band energy fitting procedure ensured more accurate recovery by optimising the extraction of the several chemical abundances and other physical parameters of interest (notably temperature). After post-processing, distributions are accurately recovered (almost always within the 3σ of the measurement error) with little to no systematic biases (of the order of a 5%, see Sect. 4.6) found mainly between the low-mass element abundances (e.g. O, Si) and the normalisation. The comparison of the relative distribution between outputs and inputs with respect to the XSPEC statistical error also showed reduced chi-squared values $\tilde{\chi}^2_{\rm red}$ close to 1 when a small fraction of outlier regions ($\leq$5%) is removed indicating a good accuracy in the fits. Remaining errors and biases can be linked to correlation between parameters (notably the normalisation), to the choice of the input weighting scheme, and to mixing effects along the line-of-sight in view of the single plasma temperature model used here. We also find that most of these errors decrease when statistics are strongly increased (biases below 2% at 1 Ms for the same spatial regions), suggesting that these effects may simply be related to statistics (see Appendix B). Studies conducted by decreasing the statistics of the runs (typically by decreasing the S/N of the regions to 50 or 100) provided equally encouraging results. Despite larger statistical errors (up to 10% higher), the main parameters (temperature, redshift, O, Si, or Fe abundance) were accurately recovered. Some fainter lines (e.g. Na, N, or Al) become however very difficult to constrain in this case.

For local clusters ($z \sim 0.1$), we demonstrate the power of the X-IFU in accurately recovering spatially resolved parameter maps, along with abundance profiles (Sect. 5.2) and abundance ratios (Sect. 5.3). The study was then extended at different redshift values, up to $z \sim 2$. By probing the chemical enrichment for very distant clusters and despite the lack of an optimised observation strategy (i.e. non-optimised exposure time), we also show the power of the X-IFU in investigating the ICM properties and the chemical enrichment of the distant Universe.

The binning and fitting procedures used here comprise "classical" approaches to X-ray data analysis, using S/N binning and fits through instrumental response matrices in XSPEC. Despite our efforts, the fitting procedures remain slightly biased ($\leq$5%) and small changes in the fitting approach can impact the overall results of the simulation (of the order of a few %). More accurate results may be achieved using, for example, Monte-Carlo (MC) fitting approaches, but unfortunately remain computationally cumbersome to be used on our large set of spatial regions. More optimised binning techniques could also be investigated for future applications (Kaastra & Bleeker 2016). The access to high-resolution spectra will provide new proxies to estimate quantities such as the temperature by using for example, line-ratio techniques. Eventually, hyper-spectral methods (e.g. blind source separation algorithms) or machine-learning-based fitting techniques (see, e.g. Ichinohe et al. 2018) could open new perspectives for the post-processing of high-resolution X-ray spectra. We would like to emphasise that, even though not applicable in our simulation case, the expected level of spectral resolution of the X-IFU will challenge our current knowledge accuracy of the spectral lines (centroid energies and intrinsic widths). This is critical to allow a meaningful interpretation of the results (as demonstrated in Hitomi Collaboration 2018d, for line ratios) and to disentangle fine spectroscopic effects (such as resonant scattering, Hitomi Collaboration 2018b). This emphasises the need for dedicated tools able to process and analyse future X-IFU high-resolution spectroscopy data-cube. In this regard, the *Athena* mission will certainly benefit from the advances expected in processing tools, fitting methods and atomic databases, from the future XRISM mission (Ishisaki et al. 2018).

Not only do these E2E simulations allow us to explore the capabilities of the future X-IFU instrument, but they also give crucial information on the effect of instrumental parameters in science observations. In this study for instance, the spectral shape of all the foreground and background components were assumed to be perfectly known. For the more local and massive clusters however, the field-of-view of the X-IFU will easily be encompassed within the angular extension of R_{500}. Cluster emission-free regions might thus be unavailable for local background calibration. The spectral resolution of the X-IFU will help mitigate this effect, by allowing us to disentangle various components through the characteristics of their spectral energy distribution. The instrument background may also contaminate the observation of faint sources, as the level of precision to which X-IFU is expected to perform requires its accurate and reproducible knowledge in flight. This may be achieved through, for example, in-flight cross-correlation with the WFI or the X-IFU cryogenic anti-coincidence detector (Cucchetti et al. 2018). Future developments could take advantage of this simulation pipeline to test other realistic instrumental effects (e.g. stray-light for galaxy cluster outskirts observations). More detailed studies of the abundance ratios recovered here will also be at the centre of a forthcoming study to highlight the capabilities of the X-IFU in constraining the ICM chemical enrichment, and notably to disentangle between the contributions of the various mechanisms of chemical enrichment (e.g. SN, AGB) throughout cosmic time.

Our study underlines the revolutionary capabilities brought by the X-IFU in future X-ray spectroscopy. With typical routine observations, the X-IFU will drastically change our

understanding of ICM mechanisms and provide a quantum leap forward in X-ray astronomy.

Acknowledgements. V. Biffi, S. Borgani and E. Rasia acknowledge financial contribution from the contract ASI-INAF n.2017-14-H.0. E. Rasia acknowledges the ExaNeSt and Euro Exa projects, funded by the European Union's Horizon 2020 research and innovation programme, under grant agreement No 754337. S. Borgani and L. Tornatore acknowledge support from the EU H2020 Research and Innovation Programme under the ExaNeSt project (Grant Agreement No. 671553). S. Borgani also acknowledge support from the INFN INDARK grant. S. Ettori acknowledges financial contribution from the contracts NARO15 ASI-INAF I/037/12/0, ASI 2015-046-R.0 and ASI-INAF n.2017-14-H.0. M. Gaspari is supported by NASA through Einstein Postdoctoral Fellowship Award Number PF5-160137 issued by the Chandra X-ray Observatory Center, which is operated by the SAO for and on behalf of NASA under contract NAS8-03060. Support for this work was also provided by Chandra grant GO7-18121X. The authors would like to extend the thanks to the anonymous referee for the suggestions and helpful comments.

References

Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
Arnaud, K. A. 1996, ASP Conf. Ser., 101, 17
Barret, D., Lam Trong, T., den Herder, J. W., et al. 2016, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray, Proc. SPIE, 9905, 99052F
Beck, A. M., Murante, G., Arth, A., et al. 2016, MNRAS, 455, 2110
Biffi, V., Dolag, K., & Böhringer, H. 2013, MNRAS, 428, 1395
Biffi, V., Planelles, S., Borgani, S., et al. 2017, MNRAS, 468, 531
Biffi, V., Planelles, S., Borgani, S., et al. 2018, MNRAS, 476, 2689
Bonafede, A., Dolag, K., Stasyszyn, F., Murante, G., & Borgani, S. 2011, MNRAS, 418, 2234
Canizares, C. R., Clark, G. W., Markert, T. H., et al. 1979, ApJ, 234, L33
Canizares, C. R., Clark, G. W., Jernigan, J. G., & Markert, T. H. 1982, ApJ, 262, 33
Canizares, C. R., Davis, J. E., Dewey, D., et al. 2005, PASP, 117, 1144
Cappellari, M., & Copin, Y. 2003, MNRAS, 342, 345
Cash, W. 1979, ApJ, 228, 939
Chabrier, G. 2003, PASP, 115, 763
Churazov, E., Vikhlinin, A., Zhuravleva, I., et al. 2012, MNRAS, 421, 1123
Clerc, N., Ramos-Ceja, M. E., Ridl, J., et al. 2018, A&A, 617, A92
Cucchetti, E., Pointecouteau, E., Barret, D., et al. 2018, Space Telescopes and Instrumentation 2018: Ultraviolet to Gamma Ray, Proc. SPIE, 10699
de Grandi, S., & Molendi, S. 2009, A&A, 508, 565
de Plaa, J. 2013, Astron. Nachr., 334, 416
de Plaa, J., Werner, N., Bleeker, J. A. M., et al. 2007, A&A, 465, 345
den Hartog, R., Kirsch, C., de Vries, C. 2018, J. Low Temp. Phys., 193, 533
den Herder, J. W., Brinkman, A. C., Kahn, S. M., et al. 2001, A&A, 365, L7
Ettori, S., Pratt, G. W., de Plaa, J., et al. 2013, ArXiv e-prints [arXiv:1306.2322]
Ettori, S., Baldi, A., Balestra, I., et al. 2015, A&A, 578, A46
Ezer, C., Bulbul, E., Nihal Ercan, E., et al. 2017, ApJ, 836, 110
Gardini, A., Rasia, E., Mazzotta, P., et al. 2004, MNRAS, 351, 505
Gaspari, M., & Churazov, E. 2013, A&A, 559, A78
Gaspari, M., & Sdowski, A. 2017, ApJ, 837, 149
Gaspari, M., Brighenti, F., D'Ercole, A., & Melioli, C. 2011, MNRAS, 415, 1549
Gaspari, M., McDonald, M., Hamer, S. L., et al. 2018, ApJ, 854, 167
Gilli, R., Comastri, A., Brunetti, G., & Setti, G. 1999, New Astron., 4, 45
Gilli, R., Comastri, A., & Hasinger, G. 2007, A&A, 463, 79
Hasinger, G., Miyaji, T., & Schmidt, M. 2005, A&A, 441, 417
Hillebrandt, W., Kromer, M., Röpke, F. K., & Ruiter, A. J. 2013, Front. Phys., 8, 116
Hitomo Collaboration (Aharonian, F., et al.) 2016, Nature, 535, 117
Hitomi Collaboration (Aharonian, F., et al.) 2017, Nature, 551, 478
Hitomi Collaboration (Aharonian, F., et al.) 2018a, PASJ, 70, 9
Hitomi Collaboration (Aharonian, F., et al.) 2018b, PASJ, 70, 10
Hitomi Collaboration (Aharonian, F., et al.) 2018c, PASJ, 70, 11
Hitomi Collaboration (Aharonian, F., et al.) 2018d, PASJ, 70, 12
Hofmann, F., Sanders, J. S., Nandra, K., Clerc, N., & Gaspari, M. 2016, A&A, 585, A130
Iben, Jr., I., & Renzini, A. 1983, ARA&A, 21, 271
Ichinohe, Y., Yamada, S., Miyazaki, N., & Saito, S. 2018, MNRAS, 475, 4739
Ishisaki, Y., Ezoe, Y., Yamada, S., et al. 2018, J. Low Temp. Phys., submitted
Kaastra, J. S., & Bleeker, J. A. M. 2016, A&A, 587, A151
Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775
Karakas, A. I. 2010, MNRAS, 403, 1413
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Leccardi, A., Rossetti, M., & Molendi, S. 2010, A&A, 510, A82
Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, ApJ, 752, 46
Lotti, S., Cea, D., Macculi, C., et al. 2014, A&A, 569, A54
Lotti, S., Mineo, T., Jacquey, C., et al. 2017, Exp. Astron., 44, 371
Macculi, C., Argan, A., D'Andrea, M., et al. 2016, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray, Proc. SPIE, 9905, 99052K
Mantz, A. B., Allen, S. W., Morris, R. G., et al. 2017, MNRAS, 472, 2877
Maoz, D., Mannucci, F., & Nelemans, G. 2014, ARA&A, 52, 107
Mazzotta, P., Rasia, E., Moscardini, L., & Tormen, G. 2004, MNRAS, 354, 10
McCammon, D., Almy, R., Apodaca, E., et al. 2002, ApJ, 576, 188
Mernier, F., de Plaa, J., Pinto, C., et al. 2016a, A&A, 592, A157
Mernier, F., de Plaa, J., Pinto, C., et al. 2016b, A&A, 595, A126
Mernier, F., de Plaa, J., Kaastra, J. S., et al. 2017, A&A, 603, A80
Molendi, S., Eckert, D., De Grandi, S., et al. 2016, A&A, 586, A32
Moretti, A., Campana, S., Lazzati, D., & Tagliaferri, G. 2003, ApJ, 588, 696
Morrison, R., & McCammon, D. 1983, ApJ, 270, 119
Nandra, K., Barret, D., Barcons, X., et al. 2013, ArXiv e-prints [arXiv:1306.2307]
Nomoto, K., Kobayashi, C., & Tominaga, N. 2013, ARA&A, 51, 457
Padovani, P., & Matteucci, F. 1993, ApJ, 416, 26
Pajot, F., Lam Trong, T., den Herder, J. W., Piro, L., & Cappi, M. 2018, J. Low Temp. Phys., submitted
Peille, P., Dauser, T., Kirsch, C., et al. 2018, J. Low Temp. Phys., submitted
Peterson, J. R., & Fabian, A. C. 2006, Phys. Rep., 427, 1
Pointecouteau, E., Reiprich, T. H., Adami, C., et al. 2013, ArXiv e-prints [arXiv:1306.2319]
Rasia, E., Ettori, S., Moscardini, L., et al. 2006, MNRAS, 369, 2013
Rasia, E., Mazzotta, P., Bourdin, H., et al. 2008, ApJ, 674, 728
Rasia, E., Borgani, S., Murante, G., et al. 2015, ApJ, 813, L17
Romano, D., Karakas, A. I., Tosi, M., & Matteucci, F. 2010, A&A, 522, A32
Roncarelli, M., Gaspari, M., Ettori, S., et al. 2018, A&A, 618, A39
Sanders, J. S. 2006, MNRAS, 371, 829
Schmid, C., Smith, R., & Wilms, J. 2013, SIMPUT - A File Format for Simulation Input (HEASARC, Cambridge (MA): Tech. report)
Seta, H., Tashiro, M. S., Ishisaki, Y., et al. 2012, IEEE Trans. Nucl. Sci., 59, 366
Simionescu, A., Werner, N., Mantz, A., Allen, S. W., & Urban, O. 2017, MNRAS, 469, 1476
Simionescu, A., Nakashima, S., Yamaguchi, H., et al. 2018, MNRAS, submitted [arXiv:1806.00932]
Smith, R. K., Brickhouse, N. S., Liedahl, D. A., & Raymond, J. C. 2001, ApJ, 556, L91
Smith, S. J., Adams, J. S., Bandler, S. R., et al. 2016, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray, Proc. SPIE, 9905, 99052H
Springel, V. 2005, MNRAS, 364, 1105
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289
Steinborn, L. K., Dolag, K., Hirschmann, M., Prieto, M. A., & Remus, R.-S. 2015, MNRAS, 448, 1504
Takahashi, T., Kokubun, M., Mitsuda, K., et al. 2016, Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray, Proc. SPIE, 9905, 99050U
Thielemann, F. K., Argast, D., Brachwitz, F., et al. 2003, in From Twilight to Highlight: The Physics of Supernovae, eds. W. Hillebrandt, & B. Leibundgut, 331
Tornatore, L., Borgani, S., Matteucci, F., Recchi, S., & Tozzi, P. 2004, MNRAS, 349, L19
Tornatore, L., Borgani, S., Dolag, K., & Matteucci, F. 2007, MNRAS, 382, 1050
Truong, N., Rasia, E., Mazzotta, P., et al. 2018, MNRAS, 474, 4089
Urban, O., Werner, N., Allen, S. W., Simionescu, A., & Mantz, A. 2017, MNRAS, 470, 4583
Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, ApJ, 465, 487
Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691
Vogelsberger, M., Marinacci, F., Torrey, P., et al. 2018, MNRAS, 474, 2073
Webbink, R. F. 1984, ApJ, 277, 355
Werner, N., Böhringer, H., Kaastra, J. S., et al. 2007, in Heating versus Cooling in Galaxies and Clusters of Galaxies, eds. H. Böhringer, G. W. Pratt, A. Finoguenov, & P. Schuecker, 309
Werner, N., Durret, F., Ohashi, T., Schindler, S., & Wiersma, R. P. C. 2008, Space Sci. Rev., 134, 337
Werner, N., Urban, O., Simionescu, A., & Allen, S. W. 2013, Nature, 502, 656
Whelan, J., & Iben, Jr., I. 1973, ApJ, 186, 1007
White, S. D. M., Navarro, J. F., Evrard, A. E., & Frenk, C. S. 1993, Nature, 366, 429
Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009, MNRAS, 393, 99
Willingale, R., Pareschi, G., Christensen, F., et al. 2014, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray, Proc. SPIE, 9144, 91442E
Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914
Wilms, J., Brand, T., Barret, D., et al. 2014, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray, Proc. SPIE, 9144, 91445X
Woosley, S. E., & Weaver, T. A. 1995, ApJS, 101, 181

Appendix A: Spatial binning algorithm comparison

Spatial binning of the data is used to increase the S/N of the considered regions, to have higher significance spectra. This is of particular interest whenever fine structures (e.g. line ratios, line doublets, absorption features) need to be observed in a spectrum. Multiple methods can be used to bin spatially, in this study two of them were considered:

– `contbin` tool developed by J. Sanders (Sanders 2006). The contbin scheme was run for a constrain fill value of two, which represents the maximum ratio length/width of a region.
– Voronoï tessellation, as defined by (Cappellari & Copin 2003)

Both these methods were tested on the same surface brightness map to estimate their performances. Various criteria were compared, such as their ability to reproduce spatial features of the cluster or the mean S/N ratio of the regions created. Figure A.1 shows a comparison of the cluster 2 regions at $z = 0.105$ without any foreground/background component (to investigate purely the binning effects) computed using either of these methods. Visually, we notice that `contbin` provides very similar results to Voronoï whenever the aspect ratio of the region is constrained to $C_{\rm fill} = 1$ (Fig. A.1, right). When a slightly higher value of the aspect ratio is allowed (Fig. A.1, bottom left), we notice that `contbin` is able to reproduce more accurately the radial contours of the cluster, notably the cold arc visible in the south-east corner of input data or the hot bubble rising west of the cluster (Fig. A.1, top left). Further, no difference is found on the average S/N of the regions, which is always above the required level. Finally, the methods give very close results in terms of number of regions (85 for Voronoï vs 87 for `contbin` for a S/N of 300), thus being equivalent computationally. For our purposes, `contbin` tool provides a more suitable binning algorithm than Voronoï. More ample tests also show no significant difference in the recovery of the physical parameters between both techniques.

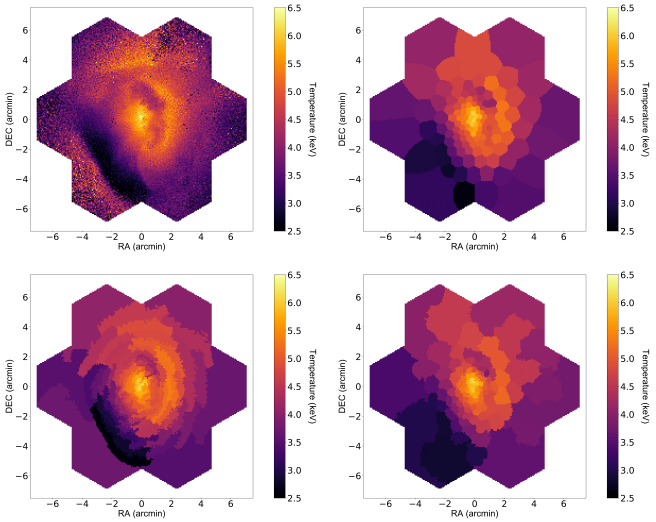

Fig. A.1. Spatial binning scheme comparison for cluster 4 spectroscopic temperature map (keV) without background, at redshift $z = 0.105$. *Top left*: unbinned raw input map from the hydrodynamical simulation. *Top right*: voronoi tessellation map using the algorithms described in (Cappellari & Copin 2003). *Bottom left*: contour map using `contbin` tool (Sanders 2006) with an aspect ratio constraint, $C_{\rm fill} = 2$. *Bottom right*: same as *bottom left*, with $C_{\rm fill} = 1$.

Appendix B: Validation of the simulation pipeline

Test hypothesis

The accuracy of the simulation pipeline needs to be verified using the cluster inputs provided. These inputs are 3D cubes of data, which needs to be projected along the line-of-sight of the instrument to be compared to the outputs of the end-to-end simulations. Ideally, this projection should be deterministic and give an unequivocal results. However, since the parameter distribution along the line-of-sight cannot be perfectly integrated (we only measure discrete number of counts, affected by statistics and background sources), multiple schemes exist to compare inputs and outputs depending on the physical quantity we wish to compare. Among those, the most widely used include:

– Emission-weighted projection, using the product $\rho^2\sqrt{T}$ of each element along the line-of-sight.
– Mass-weighted projection, using the mass of each element
– Emission-measure-weighted projection, using the emission-measure of a given line derived from Eq. (2)
– Spectroscopic schemes, as defined in (Mazzotta et al. 2004)

The accuracy of the simulated measurements was first tested by taking as estimator the relative error distribution of the output map, using emission-weighted-input maps as proxy for the input parameters. Being each region much larger than the telescope PSF they are considered independent on a strict statistical term. The relative error is assumed to follow a Gaussian

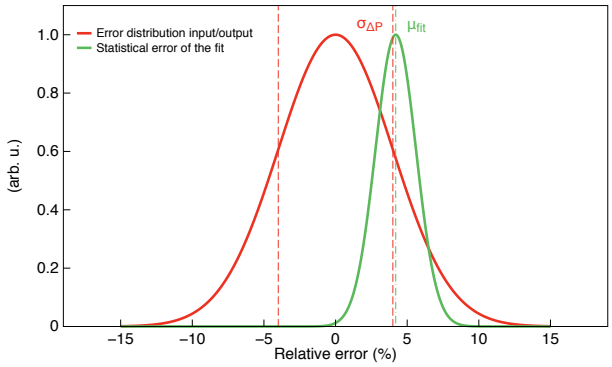

Fig. B.1. Comparison between the relative error distribution of the parameters (red) and the fitting error distribution (green – symmetric but not shown here for clarity). Ideally, if no biases are present relative error distribution should be centred and its standard deviation $\sigma_{\Delta P}$ should be very close to the mean value of the fitting error distribution $\mu_{\rm fit}$.

distribution given the sufficiently high number of regions considered (≥80), with a mean $\mu_{\Delta P} = 0$ (if no biases are present) and a standard deviation of $\sigma_{\Delta P}$, which indicates the total error on the parameter. The fitting error returned by XSPEC should also be Gaussian, and centred on given a value $\mu_{\rm fit}$ which depends on the exposure time and the emission model parameters. For an accurate measurement, the value of $\mu_{\rm fit}$ should be comparable to $\sigma_{\Delta P}$ do all parameter (Fig. B.1). A second test can also be performed using as estimator the ratio between the relative difference and the XSPEC error $\sigma_{\rm fit}$ for each region, i.e. $\chi_j = (P_{{\rm fit},j} - P_{{\rm in},j})/\sigma_{{\rm fit},j}$ and the corresponding reduced chi-square. Using the emission-measure-weighted input and the output distributions, let us take as null hypothesis (H_0): "The measurements obtained with the pipeline are consistent with the statistical errors for a given exposure time" and (H_1): "The measurements are biased" with a threshold $p_\alpha = 5\%$ (i.e. 97.5% of the Gaussian distribution, or $\sim 2.5\sigma$).

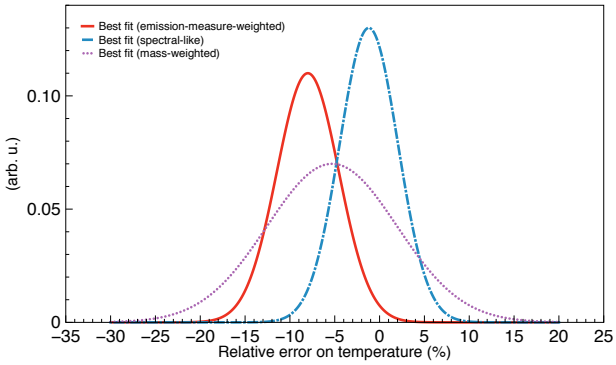

Fig. B.2. Gaussian best fits of the normalised relative error distribution on the measured temperature for different input map weighting scheme for cluster C4. The red solid curve indicates the emission-weighted best fit of the distribution ($\mu_{\Delta T} = -8.0\%$, $\sigma_{\Delta T} = 3.4\%$). The blue dash-dotted line indicates the spectroscopic temperature best fit ($\mu_{\Delta T} = -1.2\%$, $\sigma_{\Delta T} = 2.2\%$), while the dotted violet line indicates the best fit for a mass-weighted input ($\mu_{\Delta T} = -5.2\%$, $\sigma_{\Delta T} = 7.2\%$).

H_0: (H_0) can be rejected if the value of $\sigma_{\Delta P}$ is outside $\mu_{\rm fit} \pm 2.5\sigma_{\rm fit}$ for all the parameters. This can also be seen on the χ_j maps, if the number of regions with $|\chi_j| \geq 2.5$ is high. For all runs, the value of the error dispersion is always within this threshold and the number of outliers regions is small (see also Appendix C and Table 3), indicating that the (H_0) is valid and errors are consistent with the corresponding statistics.

H_1: (H_1) can be rejected if $\mu_{\Delta P} \sim 0$, within ~2.5 times the mean standard error of the distribution (due to the finite size of the sample). The samples are generally composed of $N_{\rm reg} \sim 80$ regions and the standard error on $\mu_{\Delta P}$ is given by $\sigma_{\Delta P}/\sqrt{N_{\rm reg}}$. Under these assumptions, clear biases are visible in the reconstructed emission-weighted and mass-weighted temperature maps, which presented a systematic underestimation of ~5–10% (Fig. B.2). This bias is explained by mixing effects along the line-of-sight and complexity to disentangle multi-temperature plasma with a single plasma model (Mazzotta et al. 2004). It can be reduced using spectroscopic temperature maps (Fig. B.2). The use of the broad-band fit induced in fact many other visible biases, notably between abundances (O, Si, and Fe) and temperature (Fig. B.3, upper panel). The use of multi-band fitting (detailed in Sect. 4.6) significantly reduces these biases (Fig. B.3, lower panel) within the statistical variations of the parameters. Despite this improvement, small correlations are visible between abundances and temperature (~1%) and between the normalisation and all other parameters (~5%). Efforts to reduce this bias were conducted by fixing or releasing various fitting parameters, without significant success. A clear rejection of the null hypothesis (H_1) cannot be performed, although results suggest that small residual biases and correlations remain in the current fitting procedure, mainly on normalisation.

Influence of the projection scheme

Depending on the weighting schemes (mass-weighted or emission-measure-weighted) different error distributions of the same quantities can be obtained. These discrepancies add a further complexity to evaluate any potential bias in the pipeline (Fig. B.2).

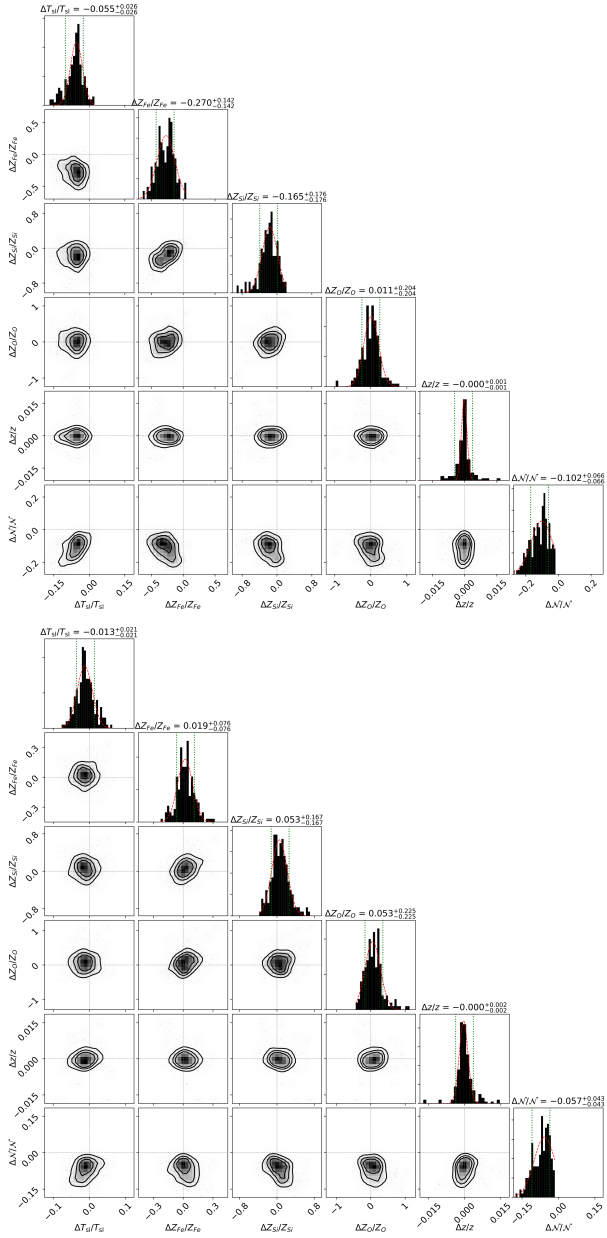

Fig. B.3. Corner plots of the relative error on parameter $(X_{\rm fit} - X_{\rm inp})/X_{\rm inp} = \Delta X/X_{\rm inp}$ as function of parameter $(Y_{\rm fit} - Y_{\rm inp})/Y_{\rm inp} = \Delta Y/Y_{\rm inp}$, for the spectroscopic temperature, $T_{\rm sl}$, oxygen, silicon, and iron abundance, redshift, z, and the normalisation $\mathcal{N}$, for cluster C4. The diagonal panels are the corresponding relative error distribution, where the red solid line indicates the Gaussian best fit of the distribution (parameters $\mu_{\Delta P}$, $\sigma_{\Delta P}$ given above) and the dotted line is the value of $\mu_{\Delta P} \pm \mu_{\rm fit}$. *Top* Broad-band fit. *Bottom*: multi-band fit (Sect. 4.6) considering a spectroscopic temperature.

Accuracy in terms of probability distributions

The previous test is only valid whenever the error distributions are assumed to be Gaussian, which is unfortunately not always the case (slight deviations from Gaussian behaviour are observed). If so, the accuracy of our method needs to be tested in the sense of the statistical distributions performing for instance a Kolmogorov–Smirnov (KS) test over the output and input distributions. The KS test compares the probability $p_{\rm KS}$ for two random variables to be drawn from the same data set (i.e., same

probability density). Using both the input and output distribution, this test showed high $p_{\rm KS}$ values (between 0.6 and 1 for the different parameters), which gives strong hints that the output distribution indeed matches the input. Statistically speaking, no real conclusion can be achieved with a single realization of the observation. To fully validate the pipeline, a large number of observations of the same cluster (either along the same line-of-sight or by taking multiple lines-of-sight) would be needed to perform a meaningful comparison using a KS method. Unfortunately, the duration of one full simulation of a cluster is of the order of a day, making it computationally cumbersome to carry out this test. For simplicity, a very high exposure time simulation of these extended sources were carried out instead. Although beyond the scope of this paper, such observations ($\geq$1 Ms) with the same binning regions decrease most of the biases below 2% and create distributions which much more alike ($p_{\rm KS} \approx 0.8/1$), suggesting that the residual errors are in part related to statistics and to the fitting scheme.

Appendix C: Further results of the simulations

We show here the reconstructed maps of the main physical parameters (spectroscopic temperature, oxygen, silicon and iron) for clusters C1, C3, and C4.

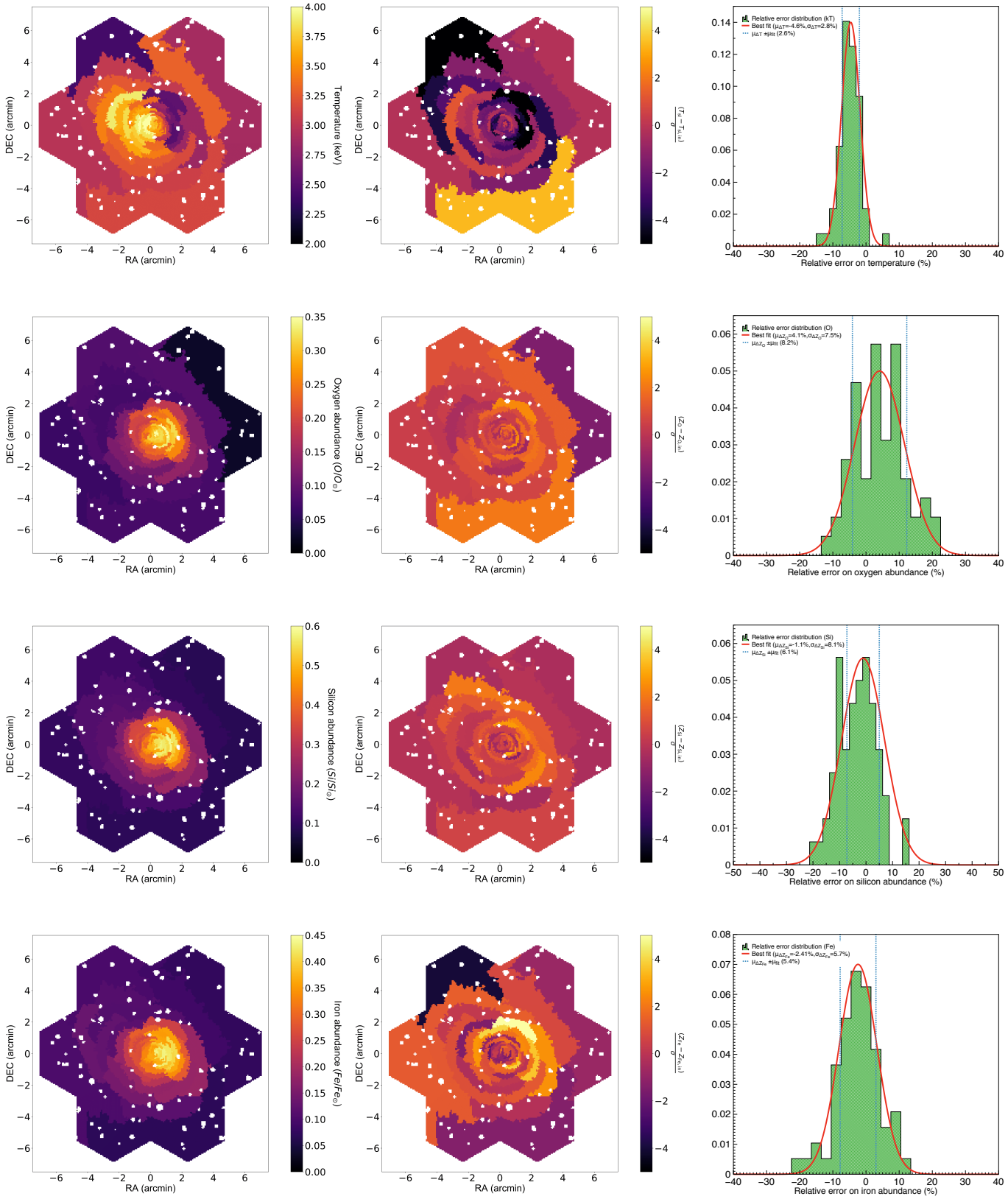

Fig. C.1. As Fig. 3 for cluster C1.

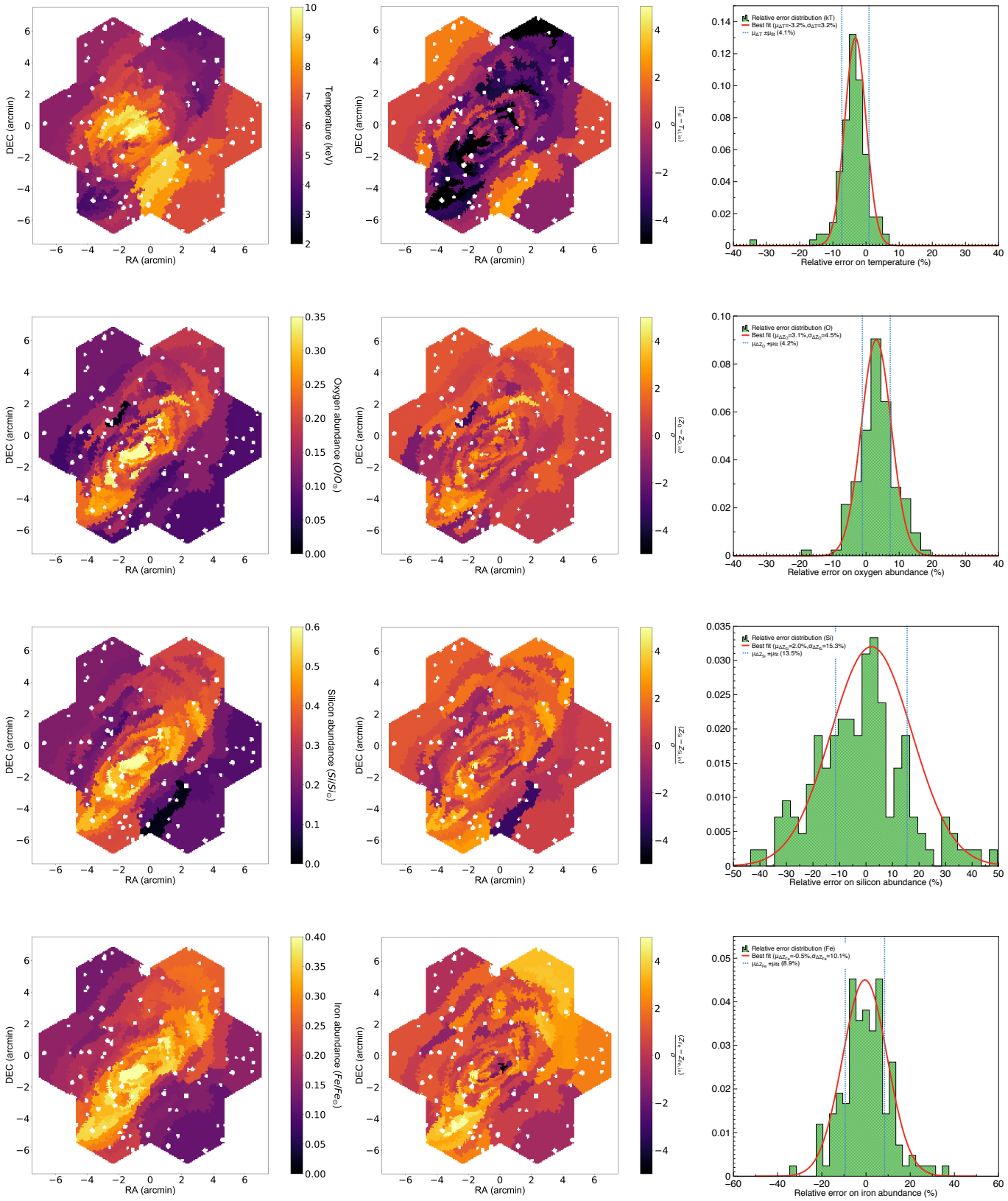

Fig. C.2. As Fig. 3 for cluster C3.

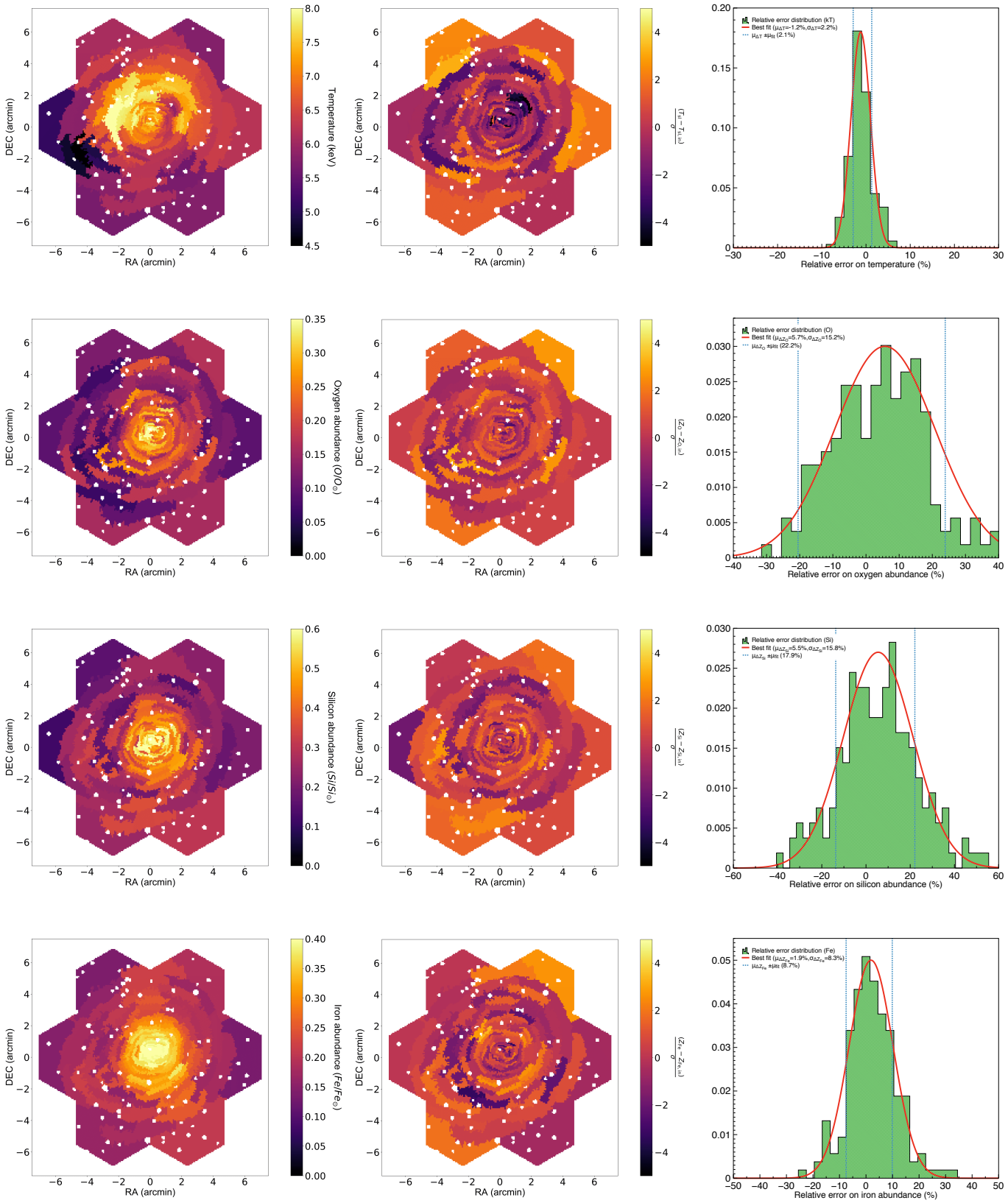

Fig. C.3. As Fig. 3 for cluster C4.