The impact of active galactic nuclei on cooling flows

Vecchia, Claudio Dalla

UNIVERSITY OF DURHAM

THE IMPACT OF ACTIVE GALACTIC NUCLEI ON COOLING FLOWS

CLAUDIO DALLA VECCHIA

INSTITUTE FOR COMPUTATIONAL COSMOLOGY
&
USTINOV COLLEGE

THIS DISSERTATION IS SUBMITTED FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

July 2005

DECLARATION

I hereby declare that my thesis entitled *The impact of Active Galactic Nuclei on cooling flows* is not substantially the same of any that I have submitted for a degree or diploma or other qualification at any other University. I further state that no part of my thesis has already been or is being concurrently submitted for any such degree, diploma or other qualification.

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. Some figures and tables are reproduced from work of other authors, and are always credited in the associated caption. Those parts of this thesis that have been published are as follows.

Chapter 3 was published as:

C. Dalla Vecchia, R. G. Bower, T. Theuns, M. L. Balogh, P. Mazzotta, and C. S. Frenk, 2004, MNRAS, 335:995–1004

Claudio Dalla Vecchia

Durham, 1st June 2005

ABSTRACT

This thesis explores the role of active galactic nuclei (AGN) in the heating of the intracluster medium (ICM). In the centre of many clusters the radiative cooling time of the ICM is much shorter than the Hubble time. Unless cooling is balanced by some form of heating, gas will flow into the cluster centre at rates up to $\sim 1000\ M_{\odot}\ yr^{-1}$. Recent *Chandra* and *XMM-Newton* X-ray observations present almost no evidence that this is happening in cluster cores. Moreover, they show that the ICM has a rather complex structure. Some of the features in the X-ray images can be explained as the interaction of the central AGN with the ICM. The most prominent feature are bubbles of hot and underdense gas inflated by jets coming from massive black holes residing in the centre of giant elliptical galaxies. These bubbles are thought to rise buoyantly through the ICM and heat the gas by depositing their energy.

I start by introducing the cooling-flow problem and by summarising the current understanding of the ICM heating processes. I then introduce the adaptive mesh refinement (AMR) code FLASH that has been used for the simulations in this thesis and the development of new routines and modules. A model of AGN heating is then applied to model clusters to investigate three issues: (1) the quenching of the cooling-flow by injection of bubbles of energy; (2) the determination of the AGN duty cycle by using measurements of sound wave positions; (3) the presence of a mass threshold below which the heating process is no longer effective. I show that cooling can be effectively balanced by AGN heating in a cluster of mass $3 \times 10^{14}\ M_{\odot}$. Then, I argue that by using measurements of sound wave positions it is possible to determine the duty cycle of the AGN with good accuracy. Finally, I show that there is a threshold mass for which the heating process is ineffective. In the light of this, I discuss the importance of the process in shaping the luminosity function of galaxies. I also apply the heating model to a cluster that has formed in a cosmological environment and discuss how to improve the code performance.

ACKNOWLEDGEMENTS

For the work done in this thesis I am indebted to: Carlos Frenk, my supervisor, for his optimistic support and guidance during all three years of PhD; Richard Bower for his amusing approach to science; Tom Theuns for his strict approach to science; Pasquale Mazzotta, Mike Balogh and Elena Rasia for useful conversations; Adrian and Dorothy Jenkins for adopting this Italian; Lydia Heck, the system manager, for solving all the technical problems (she made this thesis possible); Pat, the tea lady, for giving me big smiles; Richard and Noam, my office mates, for crazy discussions about everything but astronomy; Craig, my office neighbour, for bearing my silly questions and crazy mood; Jon and Maria, Chang and Hyun-Ah, Eva and Mary for being so kind to me; Angelo and Daniela, Rodrigo, Raul, Julian, Juan, Daniel, Anna, Valeria and Cinthia, my South American friends; Nadia, who came to England bringing warm Mediterranean wind; Luca, far from but always close to me; my father, my mother and my sister; Franco, my cousin; all the people from any part of the world that I met for a day, an hour or a minute, and that left something in my heart.

Contents

List of Figures

List of Tables

"Nel mezzo del cammin di nostra vita
mi ritrovai per una selva oscura,
ché la diritta via era smarrita."

Dante Alighieri, Inferno, Canto I

Introduction

The hierarchical cosmological model predicts that structures that populate the Universe are the result of the gravitational collapse of primordial density fluctuations. These seed fluctuations have been imprinted in the initial density and energy distribution during a phase called inflation. In the hierarchical scenario, large objects form by merger of smaller objects. Clusters of galaxies are the largest virialised structures in the Universe, containing hundreds to thousands galaxies. Cluster galaxies make up a few percent of the total cluster mass, that ranges from less then 10^{14} *to several* 10^{15} $M_{\odot}$. *More baryonic mass is in the gaseous intra-cluster medium (ICM), while the remaining, dominant fraction is dark*

matter. Being formed through gravitational collapse, where gravity acts on all forms of matter equally, galaxy clusters contain to a good approximation a composition of matter that is representative of the Universe. In this introductory chapter I describe some of the properties of the ICM focusing on the cooling of the gas in the centres of clusters. I introduce the physics of gas cooling and the cooling-flow problem (Section 1.3), and give an overview of possible heating processes that could reduce the cooling rate of the gas in the core of clusters (Section 1.4.1, 1.4.2 and 1.4.3). I also report on recent numerical results about the heating of the ICM.

1.1. The intra-cluster medium

The ICM is a hot plasma with temperature of several tens of million of degrees which has its thermal radiation maximum in the soft X-rays. Therefore, we learn most about the ICM through X-ray astronomy. Until recently, the ICM was thought to settle in hydrostatic equilibrium in the potential wells of relaxed structures. First generation X-ray telescopes seemed to confirm this idea, and showed that the temperature profiles of clusters were almost isothermal. These temperatures are consistent with measurements of galaxy velocities, such that

$$k_{\rm B}T \simeq m_{\rm p}\sigma_v^2$$

where T is the gas temperature, σ_v^2 is the galaxy velocity dispersion and $m_{\rm p}$ is the proton mass. The result supports the picture of clusters as relaxed systems in which galaxies and gas feel the same gravitational potential. However, in recent years, two satellite based X-ray telescopes have changed our vision of the ICM: ESA's *XMM-Newton* and NASA's *Chandra*. They provided a deep insight into the structure of the ICM thanks to their unprecedented spectral and imaging resolution.

The X-ray spectra reveal lines of highly ionised iron, indicating that the primordial gas, a mixture of hydrogen and helium, has been polluted. Elements heavier that helium, com-

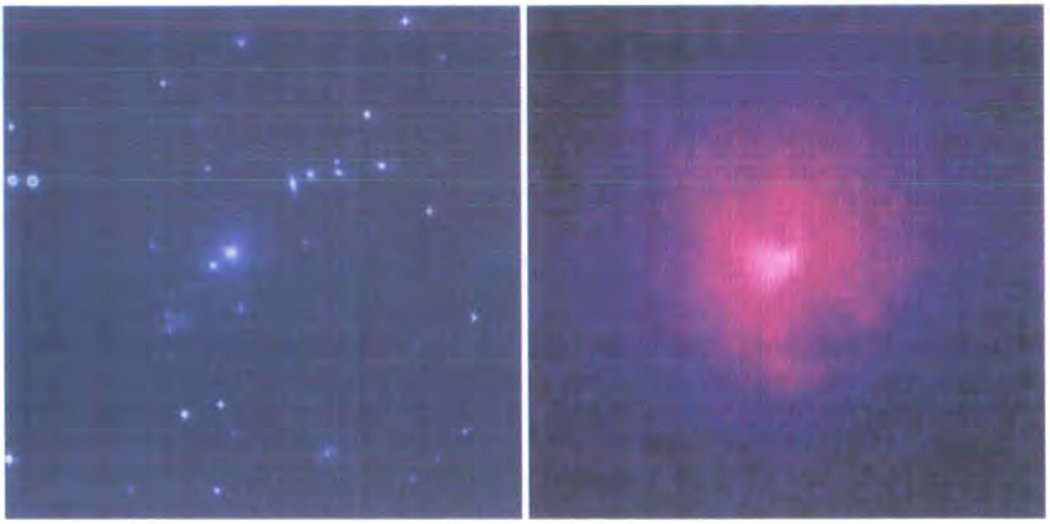

FIGURE 1.1.: The cluster of galaxies *Hydra A*. On the left, the optical image of the cluster; on the right, the X-ray image (2.7 arcmin on a side). The diffuse gas fills the intergalactic space, and presents some characteristic structures like the two X-ray bubbles seen as depression of the X-ray surface brightness (just above and below the centre of the image). From http://chandra.harvard.edu.

monly called *metals* in astronomy, are produced in the core of stars. Stars that end their life as supernovae inject metals in the host galaxy that eventually enrich the surrounding ICM by ram-pressure stripping. The metallicty of the gas is about one-third solar.

Some evidence of hierarchical formation of clusters has been found in X-ray observations. Bullet-like features in projected X-ray maps have been interpreted as colder subhalos falling into clusters. Markevitch et al. (2000) observed the first *cold front* in the cluster Abell 2142. Cold fronts are sharp density discontinuities with higher temperature in front of the edge where the density is low. They could be the result of a merger causing the gas to slosh as first claimed by Mazzotta et al. (2002), and recently investigated with numerical simulations by Mathis et al. (2005).

Depressions of the X-ray surface brightness have been observed in many clusters. They are *bubbles* of hot, under-dense plasma inflated by radio jets generated by *active galactic nuclei* (AGN). They are seen at different stages of their life, and associated with recurrent

activity of the central AGN (e.g. McNamara et al., 2000, 2001; Fabian et al., 2003a). The expansion of bubbles could eventually create weak shocks and sound waves that propagate through the ICM (e.g. Fabian et al., 2003a), while their buoyant rising through the cluster atmosphere could displace metal enriched gas (Sanders et al., 2005). Some of these these structures are shown in Figures 1.1 and 1.2. In Figure 1.1 the optical and X-ray images of the cluster of galaxies *Hydra A* are shown. The hot, under-dense gas fills the intergalactic space. The X-ray emission peak corresponds to the position of the giant elliptical galaxy in the left image. Two bubbles, inflated by the central radio source, are visible, in the right image, as depressions of the X-ray surface brightness. In Figure 1.2 two images of the cluster of galaxies *Perseus* are shown. The left image is a composition of three exposures taken in different X-ray bands. The richness of detail is striking. Two bubbles are close to the centre of the cluster, and two, previous generation bubbles are visible at the bottom and in the top-right quadrant. Sharp edges are also visible. In the right image, the unsharp-mask elaboration enhances the contrast of sharp edges. They are thought to be sound waves created by the bubbles expansion.

The ICM has a very long cooling time—of the order of the age of the Universe (see Section 1.5)—except in denser areas, which can produce a *cooling-flow* (see Section 1.3) onto individual galaxies or in cluster cores. This happens because of the dependence of the cooling rate for the ICM on temperature and density. Since *bremsstrahlung* cooling (see Section 1.2.1) is important at the temperature of the ICM, its ρ^2 dependence can pull the rug out from under a pressure-supported gas. The gas collapses into the centre of clusters increasing its density and, therefore, the cooling rate. In this run-away process a strong inflow of matter should be established. Cooling-flows have been suggested as a way to build central cluster galaxies over cosmic time, but (with a few notable exceptions such as Abell 1795) there is no trace of the massive star formation expected if the gas is being dumped into the centre. Moreover, even in the presence of strong star formation, a large amount of cold gas should be observed in cluster centres. But, there is no observational

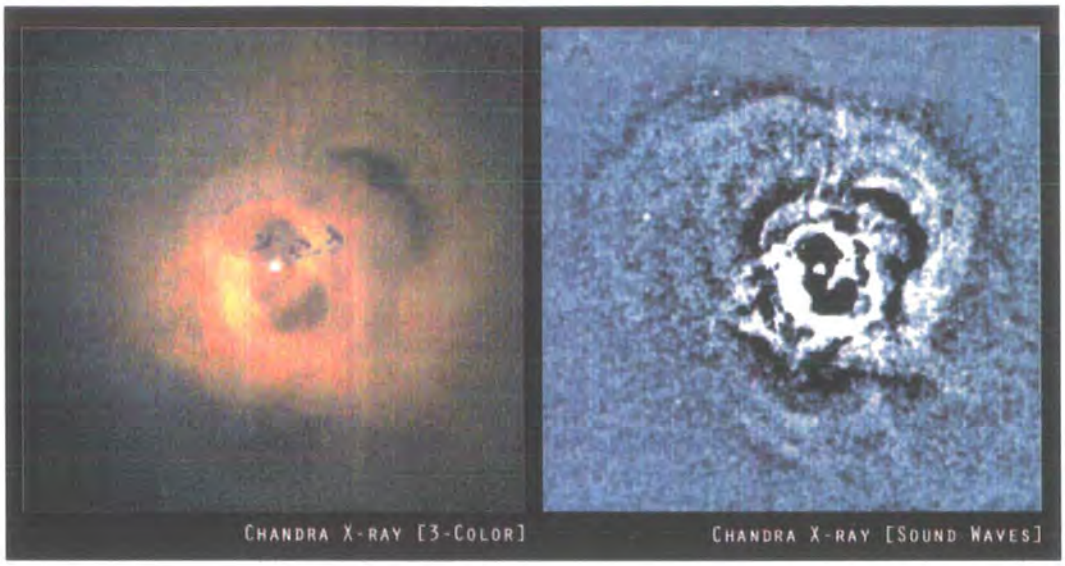

FIGURE 1.2.: *Perseus* cluster. Being the brightest X-ray cluster in the sky, it has been subject of many studies. The X-ray image on the left shows the richness of structure of this cluster. Two generation of bubbles are observed, as well as weak shocks (sound waves) propagating outward from the centre. On the right, an unsharp-mask elaboration of the X-ray observation enhances the contrast and reveal the sound waves. From http://chandra.harvard.edu.

evidence yet that the predicted cold matter mass resides in cluster cores.

In summary, the ICM was thought to be a smooth, almost isothermal medium in hydrostatic equilibrium. Recent X-ray observations showed a different picture: the hierarchical building of clusters perturbs the ICM; AGN activity is associated with many of the X-ray structures; the cooling-flow model predictions have not been confirmed.

1.2. Cooling mechanisms

In this section I introduce different emission mechanisms by which energy of a gas is radiated. Namely, these are: *free-free* emission; *free-bound* emission; *bound-bound* emission. In the first case, photons are emitted by electrons accelerated in the electrostatic field of ions. Free-bound emission occurs when an electron recombines with an ion, emitting a

photon. In the last case, photons are emitted by bound electrons that decay from an excited state to a less excited state. Free-free emission is dominant in a gas that is highly ionised, such as the ICM. In the case of the ICM, free-bound emission contributes to the spectrum continuum much less than free-free emission, and therefore does not warrant a more detailed description. Bound-bound emission becomes important at lower temperatures or when the gas is rich in heavy elements.

1.2.1. Bremsstrahlung emission

At the typical temperatures of the ICM ($\sim$ 1 keV), the gas is fully ionised. Therefore, the predominant emission mechanism is *bremsstrahlung* (*free-free* emission). The bremsstrahlung equation is derived by integrating the radiation emitted in a single encounter between an electron and a stationary ion over a Maxwellian distribution of electron velocities. The spectral emissivity of a charge Z_i in a plasma of temperature T is

$$\varepsilon_{\nu,i}^{\mathrm{ff}} = \frac{1}{3\pi^2}\left(\frac{\pi}{6}\right)^{\frac{1}{2}} \frac{Z_i^2 e^6}{\varepsilon_0^3 c^3 m_{\mathrm{e}}^2}\left(\frac{m_{\mathrm{e}}}{k_{\mathrm{B}}T}\right)^{\frac{1}{2}} n_{\mathrm{e}} n_i g(\nu,T)\ \exp\left(-\frac{h\nu}{k_{\mathrm{B}}T}\right)$$

where $g(\nu,T)$ is a slowly varying function of frequency and temperature known as the Gaunt factor. The factor corrects for quantum mechanical effects, as well as the effect of distant collisions, and its frequency average $\bar{g}(T)$ lies between 1.1 and 1.5. Commonly, the value $\bar{g}(T) = 1.2$ is adopted.

Integrating over all frequencies, and summing over all charges, one obtains the power loss of the plasma per unit volume

$$\begin{aligned}\varepsilon^{\mathrm{ff}} &= \frac{e^6}{3\pi^2\varepsilon_0^3 c^3 m_{\mathrm{e}}^2 h}\ \bar{g}(T)\left(\frac{\pi m_{\mathrm{e}} k_{\mathrm{B}} T}{6}\right)^{\frac{1}{2}} n_{\mathrm{e}} \sum_i Z_i^2 n_i \\ &\simeq 3.0\times 10^{-29}\ \mathrm{erg\ cm^{-3}\ s^{-1}}\left(\frac{T}{10^8\ \mathrm{K}}\right)^{\frac{1}{2}}\left(\frac{n_{\mathrm{H}}}{10^{-3}\ \mathrm{cm^{-3}}}\right)^2 \end{aligned} \tag{1.1}$$

For temperatures above 3×10^7 K the total emissivity of the plasma is $\varepsilon \sim \varepsilon^{\mathrm{ff}}$. One should notice that the emitted energy strongly depends on the density of the gas as n_{H}^2; the temperature dependence is less strong. In a cluster of galaxies the gas density quickly

increases from large radii to the centre. The bremsstrahlung cooling is hence much stronger in cluster cores.

1.2.2. Bound-bound emission

Bound-bound emission occurs when electrons decay radiatively from one excited level to one at lower energy. Since the energy difference between two electron orbits is fixed by quantum mechanics, photons are emitted with specific energies, and appear in spectra as emission lines. At temperatures around 10^8 K, most of the atoms are fully ionised, but, as soon as the temperature drops below $\sim 3 \times 10^7$ K, heavier nuclei recombine with electrons and the bound-bound emission of energy starts to be relevant. Sarazin (1988) gives an approximation for the total emissivity at temperatures below $\sim 3 \times 10^7$ K

$$\varepsilon^{\rm bb} \simeq 9.8 \times 10^{-30} \text{ erg cm}^{-3} \text{ s}^{-1} \left(\frac{T}{10^8 \text{ K}}\right)^{-0.6} \left(\frac{n_{\rm H}}{10^{-3} \text{ cm}^{-3}}\right)^2 \tag{1.2}$$

1.2.3. The cooling function

It is common practise, in numerical computation, to tabulate the total emission per unit mass at given temperatures, for a mixture of hydrogen and helium. This quantity is given by the cooling function $\Lambda(T)$. The cooling function also depends on the metallicity of the gas, and could include time-dependent terms. The total emissivity can be written as

$$\varepsilon = n_{\rm e} n_{\rm H} \Lambda(T)$$

where Λ now includes all the emission terms. At temperatures above $\sim 2 \times 10^7$ K, where bremsstrahlung emission dominates, the cooling function may be approximated by a power-law, with

$$\Lambda \simeq 8.5 \times 10^{-24} \left(\frac{k_{\rm B} T}{\text{keV}}\right)^{\frac{1}{2}} \text{ erg cm}^{-3} \text{ s}^{-1}$$

I show in Figure 1.3 (left panel) the value of the cooling function for a given density as function of the temperature of the gas. The two lines are the cooling function for gas

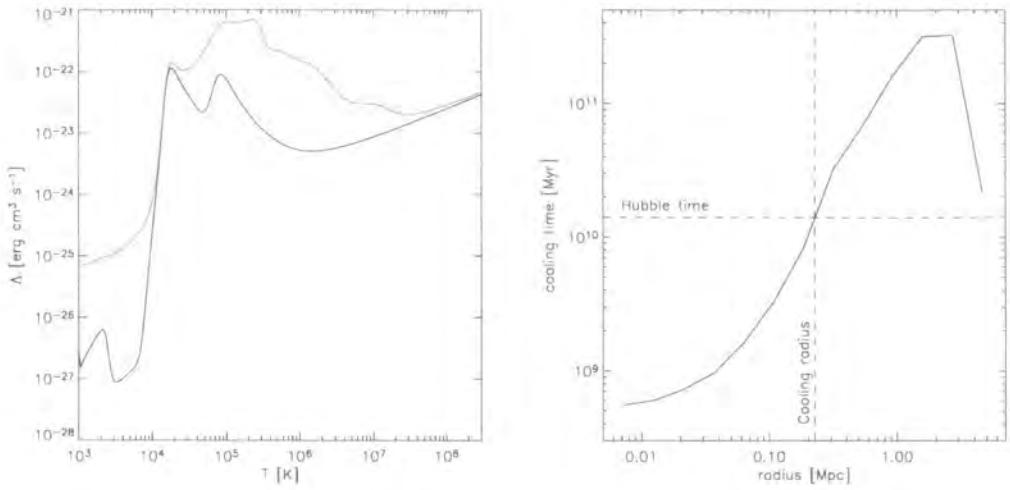

FIGURE 1.3.: **Left panel**: The cooling function Λ at a given density plotted as function of the temperature. Black and red curves are for a gas with metallicity $Z = 0$ and $Z = Z_{\odot}$, respectively. **Right panel**: The average cooling time in a simulated cluster as function of the radius. The cooling time is well below the Hubble time (dashed line) in the inner 100 kpc.

metallicity $Z = 0$ and $Z = Z_{\odot}$, where $Z_{\odot}$ is the solar metallicity.[1]

1.2.4. Definition of cooling time and cooling radius

The cooling time scale can be defined as

$$t_{\rm cool} = \left(\frac{{\rm d}\ln T}{{\rm d}t} \right)^{-1}$$

The power emitted per unit volume of gas cooling at constant pressure is given by

$$\varepsilon = \rho \frac{{\rm d}q}{{\rm d}t} = \frac{5}{2} \frac{k_{\rm B}}{\mu m_{\rm H}} \rho \frac{{\rm d}T}{{\rm d}t}$$

where ${\rm d}q$ is the heat gain per unit mass, $\rho = \mu m_{\rm H} n$ is the gas density, μ is the mean molecular weight of the gas and n is the number of gas particles per unit volume. The

[1]The *metallicity* of a gas is commonly expressed as fraction of the solar metallicity $Z_{\odot}$. This value corresponds to a mass fraction of metals equal to ~ 0.017.

constant pressure cooling time is then

$$t_{\rm cool} = T\left(\frac{{\rm d}T}{{\rm d}t}\right)^{-1} = \frac{5}{2}\frac{nk_{\rm B}T}{\varepsilon}$$

Using Equations 1.1 and 1.2 we can derive an approximation for the cooling time for different temperature intervals

$$t_{\rm cool} = \begin{cases} 0.8 \times 10^{10}\ {\rm yr}\left(\frac{n_{\rm e}}{10^{-3}\ {\rm cm}^{-3}}\right)^{-1}\left(\frac{T}{10^{7}\ {\rm K}}\right)^{1.6} & \text{for } T < 3 \times 10^{7}\ {\rm K} \\ 3.2 \times 10^{10}\ {\rm yr}\left(\frac{n_{\rm e}}{10^{-3}\ {\rm cm}^{-3}}\right)^{-1}\left(\frac{T}{10^{7}\ {\rm K}}\right)^{0.5} & \text{for } T > 3 \times 10^{7}\ {\rm K} \end{cases}$$

The *cooling radius* is defined as the radius at which the cooling time equals the Hubble time (see Figure 1.3, right panel, for an explanatory example taken from a simulated cluster).

1.3. Cooling in clusters of galaxies

X-ray observations show that several clusters of galaxies have a sharp central surface brightness peak. The central (inside 100 kpc) density of these objects is generally of the order of 10^{-2} to 10^{-1} cm^{-3}, while the temperature usually drops by a factor of three in the centre (De Grandi and Molendi, 2002). The high densities of cluster cores imply that the cooling time is shorter than the Hubble time, and less than 300 Myr in the very centre (e.g. Fabian and Nulsen, 1977). Radiative losses become important inside the cooling radius. In the absence of any heat source to balance the lost energy, the gas cools. As the temperature drops, a subsonic inflow of mass develops to maintain the central pressure required to support the weight of the overlying gas. This process is called *cooling-flow*. A review of the cooling-flow theory can be found in Fabian (1994).

Gas cooling at the centre of a cluster halo is an inherently unstable process: cooling increases the gas density, which in turn enhances the cooling rate, as shown in Equations 1.1 and 1.2. Unless cooling is balanced by some form of heating, gas will flow into the cluster centre at rates up to $\sim$ 1000 M$_\odot$ yr^{-1} (e.g. Kaastra et al., 2001). What happens to the cooled gas is unclear; it could be consumed by star formation, or lead to a reservoir of low

Name	No. of O5 stars	B5/O5	A5/O5	F5/O5	Visible SFR ($M_\odot$ yr^{-1})
A115	$9.3 \pm 1.2 \times 10^3$	—	2257 ± 864	—	0.20
A291	$9.1 \pm 0.3 \times 10^4$	—	—	—	0.99
RXJ0751.3+5012b	—	—	—	—	0.09
A646	$1.7 \pm 0.2 \times 10^4$	—	—	19254 ± 7158	0.40
RXJ1000.5+4409	—	—	—	—	0.52
A795	$1.3 \pm 0.1 \times 10^4$	—	—	—	0.14
Z3146	$1.51 \pm 0.03 \times 10^6$	—	—	—	16.43
A1068	$5.0 \pm 1.2 \times 10^5$	86 ± 61	1744 ± 875	29840 ± 10093	30.59
A1361	$1.1 \pm 0.2 \times 10^4$	—	—	—	0.12
RXJ1206.5+2810	—	—	—	—	0.06
RXJ1223.0+1037	—	—	—	—	0.03
RXJ1230.7+1220	$2.3 \pm 0.2 \times 10^2$	—	—	—	0.02
A1664	$1.17 \pm 0.03 \times 10^6$	—	1927 ± 173	—	23.06
A1795	—	—	—	—	1.94
A1835	$1.35 \pm 0.16 \times 10^6$	86 ± 31	2599 ± 505	17149 ± 4249	76.98
	$1.79 \pm 0.28 \times 10^6$	144 ± 43	2382 ± 636	13852 ± 2814	124.37
	$9.2 \pm 3.1 \times 10^5$	360 ± 145	2427 ± 1369	25877 ± 13472	125.13
RXJ1442.2+2218	$2.1 \pm 1.4 \times 10^2$	—	—	—	0.02
A1991	—	—	—	—	0.09
Z7160	$4.2 \pm 1.4 \times 10^4$	—	7565 ± 4562	232815 ± 94533	8.55
A2052	—	—	—	—	0.96
A2072	$3.5 \pm 1.6 \times 10^3$	—	—	—	0.04
RXJ1532.9+3021	$4.4 \pm 0.9 \times 10^5$	—	3724 ± 154	—	12.26
A2146	$1.4 \pm 0.2 \times 10^5$	—	—	41956 ± 11194	5.62
A2199	—	—	—	—	0.21
A2204	$1.2 \pm 0.1 \times 10^5$	—	—	—	1.29
RXJ1715.3+5725	$1.2 \pm 0.1 \times 10^4$	—	—	—	0.13
Z8197	—	—	—	—	7.86
Z8276	$2.8 \pm 0.2 \times 10^4$	—	—	—	0.30
A2390	$1.3 \pm 0.1 \times 10^5$	—	1218 ± 676	36993 ± 10350	5.40
A2495	$5.7 \pm 0.9 \times 10^3$	—	—	—	0.06

TABLE 1.1.: Table from Crawford et al. (1999). The rate of visible star formation is given in the last column (refer to the paper for more details).

cluster	redshift	optical line luminosity (erg s^{-1})	mass flow rate ($M_\odot$ yr^{-1})	dust mass ($M_\odot$)	molecular gas mass estimate ($M_\odot$)	gas-to-dust ratio
A11	0.1503	1.0×10^{42}	–	1.4×10^{7}	$2.6\pm0.7\times10^{10}$	1820
A262	0.0171	6.0×10^{39}	27	4.4×10^{5}	$9.0\pm1.3\times10^{8}$	2040
A291	0.196	4.6×10^{41}	–	2.9×10^{7}	$<2.3\times10^{10}$	<795
RXJ0338+09	0.0338	1.0×10^{41}	325	7.3×10^{5}	$3.9\pm0.4\times10^{9}$	5350
RXJ0352+19	0.109	5.8×10^{41}	–	$<6.9\times10^{6}$	$1.2\pm0.3\times10^{10}$	>1700
A478	0.0882	1.1×10^{41}	616	1.0×10^{7}	$4.5\pm2.6\times10^{9}$	450
RXJ0439+05	0.208	1.1×10^{42}	–	$<3.1\times10^{7}$	$<3.3\times10^{10}$	–
RXJ0747−19	0.1028	1.4×10^{42}	1038	$<1.8\times10^{7}$	$<7.6\times10^{9}$	–
A646	0.1268	1.6×10^{41}	–	$<1.3\times10^{7}$	$<1.3\times10^{10}$	–
RXJ0821+07	0.110	3.0×10^{41}	–	2.2×10^{7}	$3.9\pm0.4\times10^{10}$	1750
4C+55.19	0.242	$\approx10^{42}$	–	$<4.6\times10^{7}$	$<4.5\times10^{10}$	–
Zw2089	0.235	$\approx10^{42}$	–	$<4.2\times10^{7}$	$<4.6\times10^{10}$	–
Hydra-A	0.052	1.6×10^{41}	264	1.3×10^{6}	$<2.0\times10^{9}$	<1540
Zw3146	0.2906	7.0×10^{42}	1358	2.2×10^{8}	$1.6\pm0.3\times10^{11}$	740
A1068	0.1386	1.7×10^{42}	937	7.7×10^{7}	$8.5\pm0.6\times10^{10}$	1110
Zw3916	0.204	3.0×10^{41}	–	$<3.5\times10^{8}$	$<1.6\times10^{10}$	–
A1664	0.1276	1.1×10^{42}	260	$<1.6\times10^{7}$	$4.4\pm0.7\times10^{10}$	>2720
RXJ1347−11	0.4503	3.0×10^{42}	1790	1.8×10^{8}	$<6.8\times10^{10}$	<375
A1795	0.0620	1.1×10^{41}	381	$<3.1\times10^{6}$	$<2.7\times10^{9}$	—
A1835	0.2523	1.4×10^{42}	2111	1.0×10^{8}	$1.8\pm0.2\times10^{11}$	1760
Zw7160	0.2578	5.0×10^{41}	1227	1.5×10^{8}	$6.1\pm2.4\times10^{10}$	410
RXJ1532+30	0.3615	4.2×10^{42}	–	$<1.2\times10^{8}$	$2.5\pm0.4\times10^{11}$	>2110
A2146	0.2343	1.4×10^{42}	–	5.6×10^{7}	$<3.5\times10^{10}$	<620
A2204	0.1514	1.6×10^{42}	1660	$<4.3\times10^{7}$	$2.3\pm0.6\times10^{10}$	>540
Zw8193	0.1825	1.5×10^{42}	–	$<2.2\times10^{7}$	$<4.3\times10^{10}$	–
Zw8197	0.1140	1.6×10^{41}	–	$<6.7\times10^{6}$	$1.1\pm0.3\times10^{10}$	>1640
Zw8276	0.0757	1.3×10^{41}	–	2.6×10^{6}	$8.2\pm1.2\times10^{9}$	3140
A2390	0.2328	6.2×10^{41}	600	9.0×10^{7}	$<4.9\times10^{10}$	<540
A2597	0.0852	5.2×10^{41}	271	$<4.1\times10^{6}$	$8.1\pm3.3\times10^{9}$	>1990
NGC1275	0.0184	4.7×10^{42}	556	5.3×10^{7}	$1.7\pm0.2\times10^{10}$	323
I09104+41	0.4420	$>1\times10^{42}$	1060	1.6×10^{8}	$<5.1\times10^{10}$	< 319
3C48	0.3695	$>1\times10^{42}$	300	2.0×10^{8}	$1.6\pm0.6\times10^{10}$	80
R0107+32	0.0175	6.0×10^{39}	–	5.8×10^{5}	$2.2\pm0.3\times10^{9}$	3793
A1367	0.0218	5.0×10^{39}	0	4.1×10^{5}	$5.2\pm1.0\times10^{8}$	1268

TABLE 1.2.: Data from Table 9 of Edge (2001).

temperature (< 1 keV) material in the core. Although this may be the fate of a fraction of the gas, there are indications that most of the gas in fact does not follow either route. First, the star formation rates in the central galaxies of clusters are much lower than the inferred mass inflow rates (Johnstone et al., 1987; O'Connell and McNamara, 1989), rarely approaching ~ 100 $M_\odot$ yr^{-1}. To estimate the star formation rate in a typical cluster, we can average over the large sample of clusters studied by Crawford et al. (1999). Table 1.1 has been taken from Crawford et al. (1999); the inferred star formation rates are in the last column. This suggests that the typical star formation rate is less than 10 $M_\odot$ yr^{-1}. Secondly, we can compare the mass deposition rates with observations of the molecular gas content of clusters. Edge (2001) find molecular gas masses ranging from 5×10^8 to 2.5×10^{11} $M_\odot$, with an average of 2.6×10^{10} $M_\odot$. The results are in Table 1.2, where the estimate of molecular gas mass in the core of observed clusters is in the penultimate column. One could calculate an average estimate of the mass deposition rate assuming a gas consumption timescale of 10^9 yr. This is roughly the mean value of the cooling time in the proximity of the centre of clusters (e.g. Allen et al., 2001). The deposition rate may be as high as 200 $M_\odot$ yr^{-1} in a few clusters, but is $\simeq 30$ $M_\odot$ yr^{-1} on average. Similar limits are obtained using the data in Salomé and Combes (2003). Finally, recent spectroscopic X-ray observations show no evidence for significant gas cooling below 1 keV (Kaastra et al., 2001; Peterson et al., 2001; Tamura et al., 2001), and observations of molecular and neutral material reveal that the amount of cold gas in clusters of galaxies is also much less than expected from the integrated cooling-flow rate (typically less than 30 $M_\odot$ yr^{-1} – Edge et al., 2002; Edge and Frayer, 2003; Salomé and Combes, 2003).

1.4. Cooling-flow vs heating

The cooling-flow paradox has led many authors to investigate mechanisms to quench gas cooling. Observations of merging clusters show little evidence for cooling-flows, which suggests that the merger process might be implicated in disrupting cooling. Recent sim-

ulations have shown that sub-halo merging can indeed heat up gas (Burns et al., 2004); however, the amount of cold gas produced in these simulations is still too large compared to that observed, leading to the conclusion that additional heating processes must be involved. The evidence for non-gravitational heating in clusters has also been observed in a statistical sense in cluster scaling relations. These relations show departures from the self-similar scalings and reveal systematic excess of entropy (Ponman et al., 2003).

Several alternatives have been proposed, including energy injection from radio sources or active galactic nuclei (AGN; Binney and Tabor, 1995; Quilis et al., 2001; Churazov et al., 2002), viscous dissipation of sound waves (Fabian et al., 2003a; Ruszkowski et al., 2004a) and thermal conduction (Voigt and Fabian, 2004; Dolag et al., 2004). Each of these has advantages and disadvantages. In particular, it is difficult to balance cooling, with its ρ^2 density dependence, with heating processes that typically scale as ρ. Since cooling and heating can then balance only at one particular density, some of these feedback mechanisms may lead to unstable solutions with some regions of the cluster being efficiently heated while others continue to cool catastrophically. I report some of the results in the following sections.

1.4.1. Thermal conduction

Thermal conduction is a diffusive process in which heat is transported from hotter to cooler regions via collisions between particles. In astrophysical plasmas like the ICM, electrons can diffuse more effectively than ions because of their smaller mass. So, heat is transported mainly by electrons. The mean electron free path in an unmagnetised plasma is given by

$$\lambda_{\rm e} = \frac{3^{\frac{3}{2}}(k_{\rm B}T_{\rm e})^2}{4\pi^{\frac{1}{2}} n_{\rm e} e^4 \ln\Lambda}$$

where $n_{\rm e}$ and $T_{\rm e}$ are the electron number density and temperature, respectively, and $\ln\Lambda$ is the Coulomb logarithm. The Coulomb logarithm is a slowly varying function of tempera-

ture and gas density in the case of ICM plasma, and is given by

$$\ln\Lambda = 37.8 + \ln\left[\frac{T_e}{10^8\ \mathrm{K}}\left(\frac{n_e}{10^{-3}\ \mathrm{cm}^{-3}}\right)^{-1}\right]$$

The mean free path can be approximated as

$$\lambda_e \simeq 23\ \mathrm{kpc}\left(\frac{T_e}{10^8\ \mathrm{K}}\right)^2\left(\frac{n_e}{10^{-3}\ \mathrm{cm}^{-3}}\right)^{-1}$$

If the scale length of the temperature gradient, $\ell_T = T/|\nabla T|$, is much greater than the electron free path, then the heat flux is proportional to the temperature gradient, and can be written as

$$q = -\kappa\nabla T$$

where κ is the thermal conductivity coefficient. This is the case of the ICM, where the electron free path is of the order of 1 kpc and temperature changes on scale of the order of 10 kpc.

The conductivity coefficient for a plasma of ionised gas, in absence of any magnetic field, has been formulated by Spitzer (1962). The Spitzer conductivity is given by

$$\kappa_S \simeq 5.0 \times 10^{13}\ \mathrm{erg\ s^{-1}\ cm^{-1}\ K^{-1}}\left(\frac{T}{10^8\ \mathrm{K}}\right)^{\frac{5}{2}}\left(\frac{\ln\Lambda}{37}\right)^{-1}$$

and, with a further approximation, it could be written as

$$\kappa_S \simeq \kappa_0 T^{\frac{5}{2}}$$

Last two equations underline the strong dependence of the thermal conductivity on the temperature of the gas. It is effective in high mass clusters—in which the virial temperature is higher—, but almost ineffective in small clusters and galaxy groups. The thermal diffusivity of the plasma is

$$\chi = \kappa/(c_p\rho) \tag{1.3}$$

where c_p is the specific heat at constant pressure.

In the first hydrodynamical simulations of a sample of clusters including star formation and feedback, cooling, heating, and thermal conduction, Dolag et al. (2004) find that, for high mass clusters, the temperature profiles agree with observations. As expected from the strong temperature dependence, small clusters maintain the same temperature profile in simulations with and without conduction. This implies that thermal conduction cannot easily account for the observed self-similarity of the temperature profiles of fairly relaxed clusters.

Dolag et al. (2004) find that conduction can not prevent a central cooling-flow. It apparently only transports the energy gained by compressional heating of inflowing gas to outer regions of the cluster. They observe that the inclusion of conduction may even make the central cooling-flow stronger, depending on the resulting density and temperature structure of the inner parts in dynamical equilibrium. Moreover, the presence of a tangled magnetic field could reduce the diffusive motion of the electrons and, consequently, reduce or even suppress conduction.

There is also observational evidence that conduction at the Spitzer rate could not be able to balance the radiative losses of cluster cores. Voigt and Fabian (2004) studied the structure of the clusters Abell 1835 and 2199 concluding that in both clusters the required conductivity values in the innermost regions give rates equal or even higher that the full Spitzer rate.

1.4.2. Viscous dissipation

Shocks and sound waves in the ICM have been recently observed with *Chandra* in the *Perseus* cluster (Fabian et al., 2003a) and in the *Virgo* cluster (Forman et al., 2003). Other example of observed shock heating are *Hydra A* (Nulsen et al., 2004), MS0735.6+7421 (McNamara et al., 2005), Abell 478 (Sanderson et al., 2004). Based on observations of the *Perseus* cluster, it has been suggested by Fabian et al. (2003a) that viscosity may play an important role in dissipating energy injected by the central AGN. They observed a loss of

the sound waves power larger than predicted and argued that viscosity could be responsible for this dissipation. On the other hand, Churazov et al. (2002) argue that weak shocks do not dissipate energy efficiently and may be the major source of energy leakage from the cooling flow region. Strong shocks are more efficient at energy dissipation and there would be direct heating of the gas as it passed through any strong shock. They could leave behind relatively compact regions of high-entropy gas, which would contain significant fractions of energy.

The dissipation of mechanical energy due to viscosity, per unit mass of the fluid, is given by (e.g. see Brüggen et al., 2005, for more details)

$$\dot{\varepsilon}_{\rm visc} = \frac{2\mu}{\rho}\left(e_{ij}e_{ji} - \frac{1}{3}\Delta^2\right)$$

where μ is the dynamical coefficient of viscosity, $\Delta = e_{ii}$ and

$$e_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right)$$

The viscosity coefficient for an unmagnetised plasma, is given by Spitzer (1962)

$$\mu \simeq 6.0 \times 10^{-17}\ {\rm g\ cm^{-1}\ s^{-1}} \left(\frac{T}{10^8\ {\rm K}}\right)^{\frac{5}{2}} \left(\frac{\ln\Lambda}{37}\right)^{-1}$$

The kinematic viscosity is defined as

$$\nu = \mu/\rho \tag{1.4}$$

Ruszkowski et al. (2004a) (see also Ruszkowski et al., 2004b, for a three-dimensional description of the same results) simulated the viscous energy dissipation of sound waves which are generated by periodic AGN activity. Their model makes use of the full Spitzer viscosity and, for a given AGN power and duty cycle, they are able to balance the energy radiated in shells at different radii with the heating induced by sound waves dissipating their energy. Even though the model is quite simplistic, it looks promising. Heating is efficient in the central 100 kpc of their model cluster, and taking into account Spitzer conduction (not included in the simulation) could make the dissipation of energy even

quicker. The major uncertainty is the effect of magnetic fields on these processes. In the next paragraph I briefly discuss the relative importance of viscosity and thermal conduction and the role of magnetic fields.

The relative importance of viscosity and thermal conduction for the transport of energy in the gas is described by the dimensionless Prandtl number (see Kaiser et al., 2005, for more details)

$$\mathrm{Pr} = \nu/\chi$$

Using Equations 1.3 and 1.4 one finds that for a plasma the Prandtl number is

$$\mathrm{Pr} \sim 0.4 \left(\frac{m_\mathrm{e}}{m_\mathrm{p}}\right)^{\frac{1}{2}} \sim 0.01$$

This means that if viscous effects are important in the cluster gas, then thermal conduction of heat will also play a significant role in the energy budget of the cluster gas. The reverse argument can not be applied to viscosity. The gas in cluster atmospheres, in which thermal conduction is important, is not necessarily viscous. In the case of a magnetic field in the plasma, thermal conduction and therefore viscous transport of momentum is suppressed perpendicularly to the field lines. If the magnetic field is tangled on a scale much smaller than the electron free path the average thermal conductivity is significantly reduced. It is not clear which description of thermal conductivity or viscous transport of momentum is applicable to the magnetised plasma in galaxy clusters (e.g. Voigt and Fabian, 2004). In the simulations presented in this thesis I do not model viscosity, thermal conduction or magnetic fields. This is motivated by the fact that magnetic fields are likely to suppress conduction and viscosity well below the Spitzer values. Also, simulations including magnetic fields require an implementation of magneto-hydrodynamics in the code which was not available at the time this project started.

1.4.3. Energy injection from AGN

AGN are particularly promising candidates for balancing cooling, given their potentially large energy reservoir (Tabor and Binney, 1993; Binney and Tabor, 1995). Croston et al. (2004) separated a sample of groups of galaxies into radio quiet and radio loud objects, and showed that radio loud groups deviate from self-similarity. This result demonstrates that AGN could play a crucial role in heating the ICM. According to radio observations (e.g. Burns, 1990; Ball et al., 1993), 70% of cD galaxies at the centres of cooling-flows show evidence for radio activity. This fraction is much higher than in non-cooling-flow clusters ($\sim$ 23%), suggesting a link between the presence of a cooling-flow and the activity of a central super-massive black hole. Black holes of mass $10^8 - 10^{10}$ $M_\odot$ are likely to reside in these elliptical galaxies. Even accreting mass at a rate well below the Eddington limit, these super-massive black holes could provide enough energy to offset cooling. In fact, the accretion of gas onto super-massive black holes in galactic cores could give an output of order $2 \times 10^{62}(M/M_\odot)$ erg where M is the accreted mass and a standard mass-energy conversion efficiency of 0.1 is assumed (Wu et al., 2000; Bower et al., 2001).

In most numerical simulations of this process to date, energy injection produces bubbles at high temperature and low density that, after a short expansion phase, gain momentum by buoyancy (Brüggen and Kaiser, 2001; Quilis et al., 2001; Basson and Alexander, 2003). This mimics the effect of a jet which rapidly loses its collimation, as often observed in local clusters (Eilek, 2004). In other simulations, gas is injected at high velocity to mimic jets which retain their large-scale coherence (Reynolds et al., 2002; Omma et al., 2004).

Jets of high energy particles produced by massive black holes interact with the ICM through shock heating. The radius at which kinetic energy is deposited depends on the characteristics of the jets and the ICM. This radius ranges between several kpc and a few tens of kpc. Kinetic energy is quickly thermalised and bubbles of hot, under-dense plasma are formed. Inflated bubbles expand sub-sonically until pressure equilibrium with the surrounding gas is reached. A lower limit on the energy blown into a bubble of radius r_b is

the work it has done on the surrounding medium (Soker et al., 2002)

$$E_{\min} = \frac{\gamma}{\gamma - 1} p_0 V$$

where γ is the adiabatic index, $V = (4\pi/3) r_b^3$ is the volume of the bubble and p_0 is the pressure of the medium. Taking into account the internal energy of the bubble will give roughly two times as much energy.

The bubble internal density is much lower than the density of the medium. Therefore, the surrounding gas will exert a buoyancy force, and the bubble will move in the direction of decreasing pressure. The buoyancy force can be written as

$$F = \rho V \frac{\mathrm{d}\phi}{\mathrm{d}r}$$

where ρ is the density of the medium, ϕ is the gravitational potential, and r is the distance of the bubble from the centre of the cluster. After a short accelerated motion, equilibrium between buoyancy force and drag force is reached, and the bubble will move at low, almost constant velocity. The integral of the drag force over distance gives the mechanical energy exchange between the bubble and the medium. Assuming that the bubble expands adiabatically during sub-sonic motion through the cluster atmosphere in hydrostatic equilibrium, one could calculate explicitly how much energy was extracted from the bubble (Churazov et al., 2002):

$$\Delta E = -\int \rho V \frac{\mathrm{d}\phi}{\mathrm{d}r} \mathrm{d}r = E_i - \frac{\gamma}{\gamma - 1} pV = E_i \left[1 - \left(\frac{p}{p_i} \right)^{1 - 1/\gamma} \right]$$

where E_i, p_i and p are the initial energy, the initial pressure and the final pressure of the bubble, respectively. A large fraction of energy goes into the kinetic and potential energy of the entrained gas and turbulence in the wake of the bubble. It is likely, given the steep entropy profile in typical cooling-flow clusters, that the entrained gas can not travel a long distance with the rising bubble so that the kinetic energy can not efficiently leak outside the cooling-flow region (Churazov et al., 2001, 2002).

The model of *effervescent heating* (Begelman, 2001) has been investigated by many authors as mechanism to avoid the *cooling catastrophe* in the central region of clusters.

Ruszkowski and Begelman (2002) show that the joint heating by AGN energy injection and thermal conduction could lead to stable configurations of cluster temperature and entropy profiles. Moreover, in a recent work, Roychowdhury et al. (2004) show that the observed cluster entropy floor (Ponman et al., 2003) could be explained by the effect at large radii of AGN energy injection. From their calculations, they conclude that such a heating mechanism is able to solve the entropy problem, provided that the total energy injected by the AGN is roughly proportional to the cluster mass. This correlation is consistent with a linear relation between the mass of the central black hole and the mass of the cluster, which is reminiscent of the Magorrian relation (Magorrian et al., 1998) between the black hole and bulge mass.

In the work presented in this thesis I focus on the effect of recurrent AGN activity in cluster of galaxies, mainly on the role of AGN bubbles in reducing the cooling-flow of clusters. Recurrent energy injection is applied to a model cluster to infer the energy required to balance the cooling (Chapter 3). This model is then extended to halos of decreasing mass to investigate whether there is a lower mass limit at which the energy injection is no longer efficient (Chapter 5). Finally, I apply the energy injection model to a cluster formed in a cosmological environment (Chapter 6). In Chapter 4, I also show that sound waves can be detected by the *Chandra* X-ray telescope using synthetic images of a simulated cluster.

1.5. Cosmological parameters

I conclude the introduction of the thesis reporting on the latest estimate of the parameters describing the ΛCDM cosmological model of the Universe. The detailed discussion of these results is beyond the aim of this thesis. I simply stress that all the cosmological parameters assumed in our simulations are consistent with those predicted by the standard cosmological model and confirmed by the latest experiments.

Observations of the *cosmic microwave background* (CMB) have put tight constraints on the values of the cosmological parameters. The first full map of the sky was observed by

DESCRIPTION	SYMBOL	VALUE
Total density	$\Omega_{\rm tot}$	$1.02^{+0.02}_{-0.02}$
Dark energy density	Ω_Λ	$0.73^{+0.04}_{-0.04}$
Baryon density	$\Omega_{\rm b}h^2$	$0.0224^{+0.0009}_{-0.0009}$
Baryon density	$\Omega_{\rm b}$	$0.044^{+0.004}_{-0.004}$
Matter density	$\Omega_{\rm m}h^2$	$0.135^{+0.008}_{-0.008}$
Matter density	$\Omega_{\rm m}$	$0.27^{+0.04}_{-0.04}$
CMB temperature	$T_{\rm CMB}$	$2.725^{+0.002}_{-0.002}$ K
Hubble constant	h	$0.71^{+0.04}_{-0.03}$
Age of the Universe	t_0	$13.7^{+0.2}_{-0.2}$ Gyr

TABLE 1.3.: WMAP results from Spergel et al. (2003).

Cobe (see Bennett et al., 1996, for a summary of the results). In the next decade, many other experiments have measured the anisotropy in the fluctuations of the CMB temperature. Together, they have detected the first acoustic peak of oscillations in the CMB power spectrum. This peak comes from oscillations in the coupled photon-baryon fluid before the epoch of recombination[2] and gives support to the standard recombination model.

Recently, a higher resolution map of the entire sky has been produced by the Wilkinson Microwave Anisotropy Probe (WMAP) team (see Bennett et al., 2003a,b). The results of this experiment are in agreement with the standard cosmological model—a model consistent with observations from galactic scales to the largest observable scales. In the standard model, the Universe is spatially flat, and comprises radiation, baryonic matter, non-baryonic dark matter and dark energy.

[2]Before recombination, the baryonic matter is a plasma of electrons and ions due to the high temperature of the Universe. Photons interact strongly with charged particles and their mean free path is very short. Thus, the Universe if effectively opaque. As the Universe expands, the temperature drops, and, after 380000 yr after the Big Bang, baryons start to recombine. The Universe becomes transparent to radiation, and photons decouple from matter. The distribution of temperature fluctuations of the CMB are those effectively imprinted at the epoch of recombination.

The determination of some parameters with WMAP data suffers from some degree of degeneracy. In order to break this degeneracy, the inclusion of independent data sets is necessary. For example, combined results from the Two Degree Field Redshift Galaxy Survey (2dFRGS) and WMAP break degeneracies such as the geometric degeneracy and lead to precise estimates of the content and geometry of the Universe (Spergel et al., 2003; Cole et al., 2005).

Some useful values are tabulated in Table 1.3. I conclude this section with the definition of the Hubble time, which is often used in this work to refer to the age of the Universe. This is defined as

$$t_{\rm H} = H_0^{-1} = 9.8h^{-1}\ {\rm Gyr}$$

Adaptive mesh refinement with FLASH

The code used for the simulations in this thesis is FLASH *(Fryxell et al., 2000).* FLASH *was in part developed by the DOE-supported ASC/Alliance Center for Astrophysical Thermonuclear Flashes at the University of Chicago.* FLASH *is a block-structured Eulerian code that makes use of adaptive mesh refinement (AMR) for solving the hydrodynamics equations. In this chapter I describe the code introducing the Eulerian approach to hydrodynamics and the definition of the AMR grid. Then, I introduce the equations of hydrodynamics, gravity and cosmology the code solves. I also describe in detail the new modules I wrote for performing specific tasks in my simulations. I show some scaling tests that has been ran on the computing facilities of the Institute for Computational Cosmology of Durham, and I report on performance problems and their possible solutions.*

2.1. Introduction

FLASH is a *block-structured* Eulerian code that makes use of AMR for solving the hydrodynamics equations, as well as the N-body gravity. The strategy adopted by AMR codes is to increase the resolution of the computational grid where and when higher accuracy is needed; the grid "adapts" to the time evolution of gas and collisionless particles. The equations solved by FLASH are presented in Sections 2.3.1 and 2.3.2. There, some concepts regarding the AMR grid structure are assumed to be known by the reader. Therefore, I introduce the AMR approach starting from the description of the grid structure in which the physical equations are computed. In the following section I briefly describe the difference between *Lagrangian* and *Eulerian* approaches to gas dynamics, *block-structured* and *mesh-structured* AMR, and the AMR implementation of FLASH.

2.2. Adaptive mesh refinement

2.2.1. Lagrangian and Eulerian approaches to gas dynamics

Codes used for solving the equations of gas dynamics can be grouped into two main categories: *Lagrangian* and *Eulerian*. Lagrangian codes which make use of the *smoothed particle hydrodynamics* (SPH) formalism (see Monaghan, 1992, for a review of the method) represent the continuous distribution of the gas properties (e.g. density) by smoothing discrete properties of particles (e.g. mass) over a volume defined by a function called *kernel*. These codes follow the time evolution of the position of particles which carry the discrete values. A different approach is adopted in Eulerian codes which follow the evolution of continuous properties of the gas in discrete spatial positions. That is, the simulated volume is subdivided in cells in which gas physical quantities change with time. The space discretisation can be uniform (a single mesh covers the whole volume), or adaptive (different meshes with different size and resolution are placed to cover the entire volume). The adaptive mesh has the advantage of changing shape increasing or decreasing its resolution when

and where needed. This limits the amount of memory used and the number of calculations performed by the code. Two approaches to adaptive refinement are the *block-structured* and *mesh-structured* described in the following section.

2.2.2. Block-structured vs mesh-structured AMR

Structured Cartesian grids are often used in AMR applications. As the name suggests, structured Cartesian grids use a regular underlying Cartesian grid. The obvious first choice, suggested by the nested hierarchical nature of the grid itself, is to use an *octree* in 3D or *quadtree* in 2D. The connectivity of the tree also provides the information needed in a multigrid method. Although local refinement is easy to implement with this data structure, drawbacks to tree approaches include the difficulties of vectorizing (on vector architectures) and minimizing bandwidth to preserve locality (on cache-based machines). To avoid the tree traversal overhead, a mapping of leaf nodes to some other data structure is often used.

Two approaches are commonly used in AMR: *block-structured* and *mesh-structured*[1] grids. The definition of them is based on the shape of the meshes used in the AMR grid. Block-structured grid refinement makes use of "identical" elements, usually called blocks, whose dimension (number of cells) is the same. The resolution of the mesh changes by varying the spatial size of the blocks. Usually, the resolution is increased by reducing the size by a factor of two. Thus, the grid is formed by a hierarchical adaptive structure of blocks of different size. An example of a block-structured grid is shown in Figure 2.1 (left panel). On the other hand, mesh-structured grid refinement makes use of rectangular meshes of different dimension. Regions that need to be refined are covered with one or more rectangular meshes that adapt their dimension to the dimension of the region. An example of mesh-structured mesh refinement is shown in Figure 2.1 (right panel).

Both approaches have advantages and disadvantages, and it is difficult to say which is

[1] Mesh-structured grids are simply called in literature structured grids. I prefer to call them mesh-structured for sake of clearness.

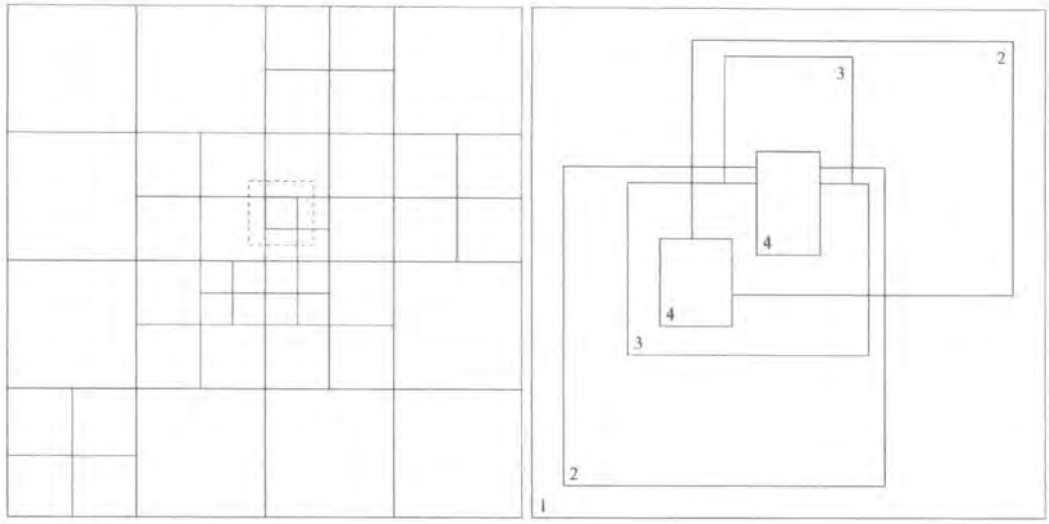

FIGURE 2.1.: **Left panel**: Example of block-structured AMR grid. A block is highlighted to show its boundary. Each block contains nxb×nyb×nzb computational zones. **Right panel**: Example of mesh-structured AMR grid. Numbers design different levels of refinement.

better. For example, block-structured meshes provide easier handling of the grid structure, easier programming, and eventually better computational performance when the dimension of blocks fits the cache memory of processors. However, each block needs to carry and to keep updated its boundary cells, and this can result in either a communication over-head on parallel computers and a larger use of memory. Mesh-structured grids are memory conservative, and can better adapt to the shape of refined regions. I do not give any further detail on this argument. In the next section I introduce the structured approach adopted by the developers of FLASH.

2.2.3. AMR with FLASH

FLASH makes use of the PARAMESH routines for handling the block-structured adaptive grid, and the message-passing interface library MPI to run on massive parallel computers. The basic constituents of the AMR grid are blocks. Each block consists of a grid of nxb×nyb×nzb computational *zones* (nxb, nyb and nzb cells along the x, y and z

directions, respectively; see Figure 2.2). The boundaries of each block, in each direction, consist of `nguard` cells called *guard cells*. Guard cells store physical quantities of neighbour blocks that are used to compute fluxes and perform interpolations at the boundaries. The grid at the coarsest level of refinement contains `nblockx`×`nblocky`×`nblockz` top blocks. Blocks are organised in a tree structure that can be walked up and down going from *parent blocks* to their *children* and *vice versa*. The blocks at the end of any branch of the tree are called *leaves*. These are the blocks containing the finer solution on the AMR grid.

A refinement criterion is necessary to adapt the mesh resolution when and where higher accuracy has to be achieved for solving the hydrodynamics equations. Moreover, most refinement criteria provide also a way to de-refine regions in which the current numerical accuracy is too high to avoid unnecessary computations. In pure hydrodynamical simulations most of the effort is concentrated in resolving shocks and discontinuities in the fluid. This is the purpose of the standard refinement criterion implemented in FLASH. The refinement criterion used by PARAMESH is based on the Löhner (1987) estimator, defined as a modified second derivative normalised by the average of the gradient over a computational cell. In one dimension it is

$$\mathcal{D} = \frac{|u_{i+1} - 2u_i + u_{i-1}|}{|u_{i+1} - u_i| + |u_i - u_{i-1}| + \varepsilon(|u_{i+1}| - 2|u_i| + |u_{i-1}|)}$$

where u_i is the refinement variable value of the i-th cell. The last term in the denominator is a filter that prevents refinement of small ripples, and can be adjusted by changing the value of ε (0.01 by default) at run time. The extension of the estimator to three dimensions is

$$\mathcal{D} = \left(\frac{\sum_{pq} \left(\dfrac{\partial^2 u}{\partial x_p \partial x_q} \Delta x_p \Delta x_q \right)^2}{\sum_{pq} \left[\left(\left| \dfrac{\partial u}{\partial x_p} \right|_{i_p+1/2} + \left| \dfrac{\partial u}{\partial x_p} \right|_{i_p-1/2} \right) \Delta x_p + \varepsilon \dfrac{\partial^2 |u|}{\partial x_p \partial x_q} \Delta x_p \Delta x_q \right]^2} \right)^{1/2} \tag{2.1}$$

where the sums are over coordinate directions, $p, q = 1, 2, 3$.

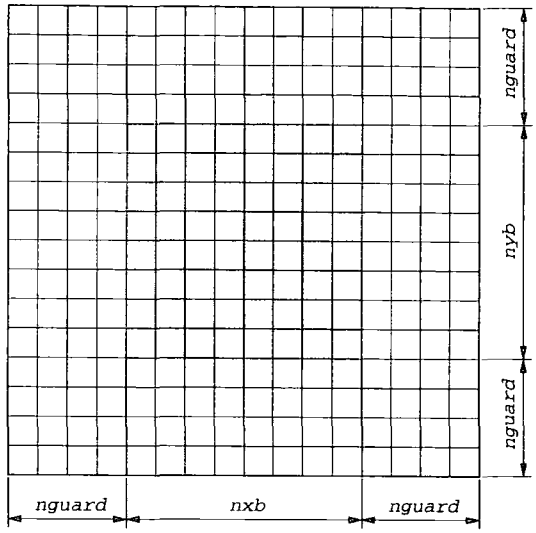

FIGURE 2.2.: Structure of a grid block in FLASH. The grey cells are the computational zones surrounded by guard cells. In this plot, nxb = nyb = nzb = 8 and nguard = 4.

The resolution is adaptively increased whenever $\mathcal{D} > c_2$ and decreased when $\mathcal{D} < c_1$, where $c_{1,2}$ are tolerance parameters. When a region requires refining ($\mathcal{D} > c_2$), child blocks with size half that of the parent blocks are placed over the offending region, and the coarse solution is interpolated. When a region requires de-refining leaf blocks are removed from the tree structure and the fine solution is restricted to their parent blocks. Blocks of different size are allowed to be next to each other only when there is just one refinement step between them. Other user-defined criteria can be specified (see Section 2.5.1). An example of a refined AMR grid is shown in Figure 2.1 (left panel).

2.3. FLASH *solvers*

FLASH is capable of simulating general compressible flow problems found in many astrophysical environments, as well as N-body gravity. It is a modular code, meaning that the routines performing different tasks are grouped in different directories and come along with `Config` and `Makefile` files. In the next sections I explain how the *hydro* and

gravity modules work showing the equations that are solved, their discretisation and the algorithms used for their integration. A more detailed description is beyond the scope of this thesis. More details can be found in the FLASH manual (http://flash.uchicago.edu/).

2.3.1. Hydrodynamics

The `hydro` module solves Euler's equations for compressible fluids in one, two, or three spatial dimensions. The conservative form of these equations is

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) &= 0 \\ \frac{\partial \rho \mathbf{v}}{\partial t} + \nabla \cdot (\rho \mathbf{v}\mathbf{v}) + \nabla p &= \rho \mathbf{g} \\ \frac{\partial \rho E}{\partial t} + \nabla \cdot [(\rho E + p)\mathbf{v}] &= \rho \mathbf{v} \cdot \mathbf{g} \end{aligned} \tag{2.2}$$

where ρ is the density of the fluid, $\mathbf{v}$ is the velocity, p is the pressure, E is the sum of the internal energy ϵ and kinetic energy per unit mass

$$E = \epsilon + \frac{1}{2}|\mathbf{v}|^2$$

$\mathbf{g} = -\nabla\phi$ is the acceleration due to gravity, and t is the time coordinate. The equation of state for an ideal gas

$$p = (\gamma - 1)\rho\epsilon$$

where γ is the ratio of specific heats, is used to compute the pressure. Along with these equations, Poisson's equation is needed to compute the acceleration $\mathbf{g}$ if a gravitational field ϕ is specified.

$$\nabla^2\phi = 4\pi G\rho$$

Other definitions of the acceleration are possible. For example, constant and point mass gravity can be chosen in the setup of the simulation, and specific user-defined gravity routines can be easily implemented.

The system of equations 2.2 is *hyperbolic*. Thus mass, momentum and energy are conserved—in the absence of source terms—in the finite difference formulation of the conservation laws. The system can be written as (see Appendix A for the mathematical details)

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}^{\alpha}(\mathbf{u})}{\partial x^{\alpha}} = 0$$

or, in integral form, as

$$\frac{\partial}{\partial t}\int_{\Omega} \mathbf{u}\,\mathrm{d}\Omega + \int_{\partial V} \mathbf{F}(\mathbf{u})\cdot\mathbf{n}\,\mathrm{d}S = 0$$

where $\mathbf{u}$ is the vector of unknowns (in our case $\mathbf{u} = (\rho, \rho v_x, \rho v_y, \rho v_z, E)$), and $\mathbf{F}^{\alpha} = (\mathbf{f}, \mathbf{g}, \mathbf{h})$ are functions called fluxes.

One could define a computational cell as the discrete domain $\Omega = [t, t+\Delta t] \times \Delta V$, where $\Delta V = \Delta x \Delta y \Delta z$ is now the volume of the cell. Then, Gauss's theorem allows rewriting the system above as

$$\begin{aligned}(<\mathbf{u}_{t+\Delta t}> - <\mathbf{u}_t>)\Delta V = &-\left[\int_{\partial V_{x+\Delta x}} \mathbf{f}\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}z - \int_{\partial V_x} \mathbf{f}\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}z\right]\\ &-\left[\int_{\partial V_{y+\Delta y}} \mathbf{g}\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}z - \int_{\partial V_y} \mathbf{g}\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}z\right]\\ &-\left[\int_{\partial V_{z+\Delta z}} \mathbf{h}\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y - \int_{\partial V_z} \mathbf{h}\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y\right]\end{aligned}$$

where $\partial V_{x(y,z)}$ and $\partial V_{x(y,z)+\Delta x(y,z)}$ are the cell faces, and

$$<\mathbf{u}> = \frac{1}{\Delta V}\int_{\Delta V} \mathbf{u}\,\mathrm{d}V$$

is the cell average of $\mathbf{u}$ at a given time. That is, the variation of $\mathbf{u}$ with time is the flux of $\mathbf{u}$ through the faces of the cell. The fluxes are computed at the cell boundaries by solving the local Riemann problem (i.e., the initial value problem with discontinuous data; see Appendix A for the mathematical definition of Riemann problem). Algorithms solving the Riemann problem are caller *Riemann solvers*. In the next paragraphs I give an overview of the solver implemented in FLASH.

The exact iterative Riemann solver typically needs the computation of an initial guess solution for the Newton iterative method, the identification of the pattern of the solution and finally some Newton iterations. Without explaining more the details of the algorithm, the reader should be aware that it is a clearly computationally expensive procedure. Moreover, most of the structure of the resulting solution is lost in the Godunov's method, due to the final cell-average operation on each grid cell. This suggests that good numerical results could be obtained by calculating an appropriate approximate solution instead of the exact one.

The Godunov-type scheme and all its higher-order generalisations require the solution of a Riemann problem for every cell boundary at each time-step. Godunov's method uses a finite-volume spatial discretisation of the Euler equations together with an explicit forward time difference. Time-advanced fluxes at cell boundaries are computed using the numerical solution to the Riemann's shock tube problem at each boundary. Initial conditions for each Riemann problem are determined by assuming the non-advanced solution to be piecewise-constant in each cell. Using the Riemann solution has the effect of introducing explicit non-linearity into the difference equations and permits the calculation of sharp shock fronts and contact discontinuities without introducing significant non-physical oscillations in the flow. Since the value of each variable in each cell is assumed to be constant, Godunov's method is limited to first-order accuracy in both space and time.

FLASH includes a directionally split *piecewise-parabolic method* (PPM) solver, which is a higher-order accurate Godunov method developed by Woodward and Colella (1984). PPM improves the method by representing the flow variables with piecewise-parabolic functions. It also uses a monotonicity constraint rather than artificial viscosity to control oscillations near discontinuities. Although this could lead to a method which is accurate to third order, PPM is formally accurate only to second order in both space and time. Nevertheless, PPM is considerably more accurate and efficient than most formally second-order algorithms. PPM is particularly well-suited to flows involving discontinuities, such

as shocks and contact discontinuities. The method also performs very well for smooth flows. The high accuracy and resolution of PPM are obtained by the explicit non-linearity of the scheme and through the use of intelligent dissipation algorithms. These algorithms are described in detail by Colella and Woodward (1984).

2.3.2. Gravity

The `gravity` module makes use of a multigrid Poisson solver. Multigrid techniques are widely used in AMR codes to solve *elliptic* equations like the Poisson's equation

$$\nabla^2\phi(\mathbf{x}) = c\rho(\mathbf{x})$$

where ϕ and ρ are the potential and the density of a scalar field, respectively, and c is a proportionality constant. The advantage of using adaptively refined grids is to accelerate the convergence of *relaxation* methods. For these methods (the one used in FLASH is the Gauss-Seidel; see below) the convergence rate decreases when the size of the mesh increases. Modal analysis shows that the longest wavelength components of the error requires many more iterations to decrease below a given tolerance. In multigrid algorithms all wavelengths converge at the same rate, because long wavelengths on a fine mesh appear to be short wavelengths on a coarse mesh.

The relaxation operator used by FLASH is the Gauss-Seidel with *red-black ordering*. The Gauss-Seidel operator is first applied to those cells for which $i + j + k$ is even (the red checks on a chess board). Then, these updated values are used to compute the *residuals* on the cells with odd $i+j+k$ (the black checks). Once the residual is known, the *correction* to the solution is calculated. The operation is repeated until a convergence criteria is satisfied, or a maximum number of iterations has been reached. The residual is defined as

$$R = \nabla^2\phi(\mathbf{x}) - \nabla^2\tilde{\phi}(\mathbf{x}) = c\rho(\mathbf{x}) - \nabla^2\tilde{\phi}(\mathbf{x})$$

where $c\rho$ is the *source* (in the case of gravity it is $4\pi G\rho$, where G is the gravitational constant), ϕ is the solution, and $\tilde{\phi}$ is the guess of the solution. Since the Poisson's equation is

linear, the residual satisfies the equation $\nabla^2 C(\mathbf{x}) = R(\mathbf{x})$ whose solution is the correction $C(\mathbf{x})$. Once the value of the correction has been calculated, the solution is updated to the value

$$\phi(\mathbf{x}) = \tilde{\phi}(\mathbf{x}) + C(\mathbf{x})$$

The discretisation on the grid of $\nabla^2\phi$ can be written as

$$\nabla^2\phi^{\ell}_{i,j,k} = \frac{\phi^{\ell}_{i+1,j,k} + \phi^{\ell}_{i-1,j,k} - 2\phi^{\ell}_{i,j,k}}{h_\ell^2} + \frac{\phi^{\ell}_{i,j+1,k} + \phi^{\ell}_{i,j-1,k} - 2\phi^{\ell}_{i,j,k}}{h_\ell^2} + \frac{\phi^{\ell}_{i,j,k+1} + \phi^{\ell}_{i,j,k-1} - 2\phi^{\ell}_{i,j,k}}{h_\ell^2}$$

where i, j, k are the indexes of the computational zone of interest, ℓ is the level of refinement, and h_ℓ is the linear size of the computational zones at level ℓ. For the sake of simplicity, $\Delta x = \Delta y = \Delta z = h$. Let us define for each cell in a block at level ℓ the residual $R^{\ell}_{i,j,k}$, the source $S^{\ell}_{i,j,k}$ and the correction $C^{\ell}_{i,j,k}$ in their discrete form. With these definitions, the multigrid algorithm used by FLASH can be summarised as follow:

`if (first call of the solver)`

initialise solution, correction and residual to zero

`else`

use the solution at the previous step as initial guess

`end if`

compute the residual $R^{\ell}_{i,j,k} = S^{\ell}_{i,j,k} - \nabla^2\phi^{\ell}_{i,j,k}$ on all leaf blocks

compute the discrete L2 norm of the residual and the source

`do while (the ratio of residual and source norm > tolerance)`

zero the correction C on the highest level of refinement, $\ell_{\mathbf{max}}$

`do` $\ell = \ell_{\mathbf{max}}, 2, -1$

copy the solution $\phi^{\ell}_{i,j,k}$ to a temporary variable $\tau^{\ell}_{i,j,k}$

zero the correction $C^{\ell-1}_{i,j,k}$

apply the relaxation operator several times to the correction equation on level ℓ, $\nabla^2 C^{\ell}_{i,j,k} =$

$R^{\ell}_{i,j,k}$

add the correction $C^{\ell}_{i,j,k}$ to $\phi^{\ell}_{i,j,k}$

compute the residual of the correction equation on all blocks on level ℓ, and restrict it to $R^{\ell-1}_{i,j,k}$

compute the residual of the source equation on all leaf blocks of level $\ell - 1$ and leave the result in $R^{\ell-1}_{i,j,k}$

`end do`

solve the correction equation on the coarsest level, and correct the solution.

`do` $\ell = 2, \ell_{\max}$

prolongate the correction from level $\ell - 1$ and add the result to $C^{\ell}_{i,j,k}$

replace $R^{\ell}_{i,j,k}$ with the residual of the correction equation

zero a second temporary variable $\mu^{\ell}_{i,j,k}$

apply the relaxation operator several time to $\nabla^2 \mu^{\ell}_{i,j,k} = R^{\ell}_{i,j,k}$

add $\mu^{\ell}_{i,j,k}$ to the correction $C^{\ell}_{i,j,k}$

copy $\tau^{\ell}_{i,j,k}$ back into $\phi^{\ell}_{i,j,k}$ on all leaf blocks

add the correction $C^{\ell}_{i,j,k}$ to the solution $\phi^{\ell}_{i,j,k}$ on all leaf blocks

`end do`

compute the residual $R^{\ell}_{i,j,k} = S^{\ell}_{i,j,k} - \nabla^2 \phi^{\ell}_{i,j,k}$ on all leaf blocks

compute the discrete L2 norm of the residual

`end do while`

As one can notice, the relaxation is done going from the maximum level of refinement $\ell_{\max}$ down to the coarsest level, and back to $\ell_{\max}$. This algorithm is called V-cycle.

2.3.3. Cosmology

The `cosmology` module solves the Friedmann's equation. It includes also library functions for various calculations needed by FLASH for initialising, performing, and analysing cosmological simulations. All these calculations take place in *co-moving* coordinates

$$\mathbf{x} = \mathbf{r}/a$$

where $\mathbf{r}$ is the proper position vector and $a(t)$ is the time-dependent cosmological scale factor. The gas velocity $\mathbf{v}$ is then the co-moving peculiar velocity $\dot{\mathbf{x}}$, and the co-moving hydrodynamical quantities are defined as

$$\rho = a^3\tilde{\rho}$$

$$p = a\tilde{p}$$

$$T = \tilde{T}/a^2$$

$$\rho\epsilon = a\tilde{\rho}\tilde{\epsilon}$$

With these definitions, the Euler's equations 2.2 can be written as follow:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0$$

$$\frac{\partial\rho\mathbf{v}}{\partial t} + \nabla\cdot(\rho\mathbf{v}\mathbf{v}) + \nabla P + 2\frac{\dot{a}}{a}\rho\mathbf{v} + \rho\nabla\phi = 0$$

$$\frac{\partial\rho E}{\partial t} + \nabla\cdot[(\rho E + p)\mathbf{v}] + \frac{\dot{a}}{a}[(3\gamma - 1)\rho\epsilon + 2\rho v^2] + \rho\mathbf{v}\cdot\nabla\phi = 0$$

2.4. Running FLASH

For each numerical problem to be simulated, the user must provide a directory containing initialisation routines and a `Config` file specifying which modules are to be included for the compilation of the code. The `Config` file of each module or problem setup defines the dependencies on other modules and allows the definition of parameters and variables used during the simulation by the module itself and eventually by the modules which might depend on it. Parameters that are specified in the `Config` file can be changed for any run of the code using the file `flash.par`. A `flash.par` file must be provided for any problem setup in order to perform the simulation. The most important initialisation routine is `init_block.F90`. This routine is called at the beginning of a new run to set up the initial condistions on the AMR grid. Depending on the specific problem, it must be provided by the user.

The compilation of the code consists of two steps. First, a setup script is executed.

The name of the setup directory must be specified, and other parameters can also be set (for example, the maximum number of blocks to be allocated for each processor). The script links into the `object` directory all the source files of the modules specified in the `Config` of the setup. That is, only those source files are compiled and linked into the FLASH executable. The second step is of course the compilation. An executable file called `flash2` is created in the `object` directory.

2.5. Implementation of new routines and modules

In this section I describe the routines and modules that I have been adding to FLASH to perform specific tasks. Most of these routines have been used in the simulations presented in this thesis, while some of them were attempts to find a better solution to the problem of interest.

2.5.1. Refinement routines

The default refinement criterion of FLASH is the one in Equation 2.1. It works very well when the quantities used to check the refinement (density, temperature, etc.) are smoothly distributed. This criterion has been used for the simulations of Chapter 3 in which a refinement on temperature was enough to follow the evolution of bubbles of energy; only an adjustment of the refinement thresholds $c_{1,2}$ was necessary. The case of cosmological simulations in quite different. Once initial perturbations in the density start to collapse, density and temperature gradients become steep everywhere. In this regime, the standard refinement criterion causes the code to refine the entire box, resulting in an undesired over-refinement of the AMR grid. This is not good not only because an over-refined grid would require a larger computational effort, but also because collisionless particles in under-dense regions would act as gravitational point-sources adding unrealistic small scale fluctuations. One would prefer to refine only regions in which structures are collapsing, adding more resolution where the density is higher and the gravitational potential is steeper. For cosmo-

logical simulations, I wrote two routines to perform the refinement of the grid based on the gas density and the collisionless particle number density. The second one has been used in the cosmological simulations in this thesis.

Gas overdensity threshold criterion

As said above, one would prefer to resolve overdense regions of the simulated box. This routine implements a criterion based on the overdensity of the gas. Blocks are marked to be refined up to level ℓ whenever they contain cells with gas overdensity above the logarithmic threshold

$$\delta > c^{\ell-1}$$

where δ is the gas overdensity with respect to the critical density $\rho_{\rm crit}$ scaled by the baryon fraction $\Omega_{\rm b}$

$$\delta = \frac{\rho - \Omega_{\rm b}\rho_{\rm crit}}{\Omega_{\rm b}\rho_{\rm crit}}$$

and c is the runtime parameter that defines the overdensity threshold. For example, if $c = 8$ the overdensity thresholds for level $\ell = 1, 2, 3\ldots$ are $\delta = 8, 64, 512, \ldots$, respectively. At the same time, the routine checks that blocks are not refined at a level above the one required by the logarithmic threshold. If this is the case, blocks are marked to be de-refined.

Particle number density criterion

A criterion similar to the one described in the previous section can be applied to the collisionless particle number density. This approach has the advantage that the particle density is kept almost constant on the refined grid. Thus, one can be sure that the dominant term of the gravitational potential (the one given by dark matter) is properly resolved. The routine marks for refinement blocks which contain computational zones with more than a given number of particles. Once a block has been refined, the particle number density of the cells of its children is roughly one eighth of their parent. Then, newly created child blocks have

particle density below the refinement threshold. These are not marked for de-refinement unless their parent block no longer obeys the refinement criterion.

Temperature threshold criterion

In the cosmological simulations of Chapter 6, I found useful to add an extra refinement criterion on top of the one described in the previous section to resolve bubbles of energy. This criterion requires the code to refine, up to a given maximum level, all the blocks containing cells with temperature above a given threshold. In the case of offending regions (the ones with temperature larger that the threshold) satisfying other de-refinement criteria, the refinement is forced. The temperature threshold is given in physical units, and the routine takes care of its scaling with the expansion factor a.

2.5.2. Importing GADGET *outputs*

The simulations presented in 6 were performed restarting with FLASH a simulation that was ran with GADGET (Springel et al., 2001; Springel, 2005). GADGET is an N-body hydrodynamical code. It uses the formalism of SPH to represent the gas. Therefore, particles are used to simulate both dark and baryonic matter, and restarting the simulation with FLASH requires the mapping of the gas physical quantities on the grid. I describe here the routines I wrote for this task.

Since I was going to use the particle number density refinement criteria of Section 2.5.1, the shape of the initial grid must be defined by the dark matter particle distribution. In FLASH any collisionless particle belongs to the processor computing the leaf block in which the particle resides. Therefore, one needs an initial grid to distribute particles among processors. This was achieved by first refining uniformly the volume up to the minimum level of refinement; the blocks are now distributed among processors. Then, the GADGET output (dark matter and SPH particles) was read-in, and the dark matter particles added to the particle array structure of FLASH. At this point, each processor has all the particles in

its leaf blocks, and further refinement can be initiated. Cycling from the minimum to the maximum refinement level allowed, the refinement criteria is applied to the whole volume. New blocks are added where needed, and redistributed among processors as well as their particles. At the end of this process, the AMR grid on which the SPH gas particles are going to be mapped, is set.

The mapping of the gas particles is done by computing density, internal energy and velocity at the centre of each cell. Particles properties are spread over a sphere of radius h (smoothing length) using the same kernel function as GADGET:

$$\begin{cases} \frac{8}{\pi}(1-6q^2(1-q)) & \text{for } 0 \leq q \leq 1/2 \\ \frac{4}{\pi}(1-q)^3 & \text{for } 1/2 < q \leq 1 \end{cases}$$

where $q = r/h$ and r is the distance. Each particle contributes to the physical quantities in the cells its kernel sphere intersects. The resolution is set by the minimum value between the smoothing length of the particle, h, and the resolution of the AMR grid, $2\Delta x$. I imposed that the used smoothing length, $h_{\rm used}$, to be

$$h_{\rm used} = \min(h, 2\Delta x)$$

where Δx is the size of the cell. The total mass on the grid is then calculated and eventually rescaled to match the total mass of the SPH particles.

2.5.3. Cooling module

For the simulations I performed for this thesis, I needed a more robust and accurate cooling routine than the one in FLASH. I adopted the cooling functions of Theuns et al. (2002), which include cooling by H, He and metals, in the presence of an ionising UV-background. This cooling function uses interpolation tables for the cooling rate due to a solar admixture of metals, obtained from CLOUDY (version 94, Ferland et al., 1998). I assume a metallicity of one third solar, and the UV-background from Haardt and Madau (1996), as updated by Haardt and Madau (2001). The cooling routine calculates the internal energy and element

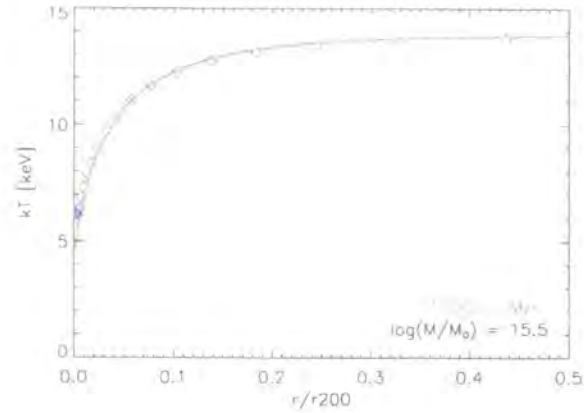

FIGURE 2.3.: Comparison between the analytic model of McCarthy et al. (2004), red line, and the results of the same simulation setup run with FLASH, blue diamonds. The data are from the simulation output at the time of 12 Gyr. Data for the plot have been provided by Ian McCarthy.

abundances implicitly. That is, at each time-step, the quantities are iteratively computed integrating over the time-step with increasing accuracy, so that no time-step limiter is necessary. The cooling routine was tested against the MEKAL (Mewe et al., 1986) cooling routines in the X-ray analysis package XSPEC (Kaastra, 1992) and against the cooling rates used by McCarthy et al. (2004). I show in Figure 2.3 the comparison with McCarthy's analytical model. We set up the same model cluster with initially isothermal gas in hydrostatic equilibrium. The simulation I ran is fully three-dimensional, while McCarthy's is uni-dimensional. The agreement is very good, and the only significant difference is at the centre of the cluster. This is only a resolution issue, since the grid I used does not approach the infinite resolution of the analytic model. Moreover, in the analytic model mass is removed from the very centre of the cluster, while in my simulation mass accumulates in the central cells.

2.5.4. Material composition module

The definition of the composition of elements and their attributes for a given problem is possible through the module `material`. FLASH evolves the mass fractions of any defined element advecting them with the density. The `material` module provides also a routine for computing the mean molecular weight using the equation

$$\frac{1}{\mu} = \sum_{i=1}^{N} \frac{1}{A_i} \tag{2.3}$$

where μ is the mean molecular weight, N is the number of elements, and A_i is the atomic weight of each element. The numerical solution of the conservation laws could lead to overshoot of the mass fraction values in some cells. To prevent non-physical situations, the mass fractions are forced to lie in the interval $[0, 1]$ and renormalised any time they do not sum to unity.

In the case of simulations with cooling, one has to take into account the evolution of hydrogen and helium ion abundances, and the metallicity of the gas. I wrote a new module which defines seven different species: H, H^+, He, He^+, He^{++}, e^- and Z, where Z is the metallicity in units of the solar metallicity $Z_\odot$. Notice that in order to have the right value of the mean molecular weight, the electron mass fraction is needed, so the atomic weight of the electron is defined as $A_e = m_e/m_p$, where m_e and m_p are the masses of the electron and proton, respectively. For the metals, I assume the average value $A_Z = 16.0$, that is a good approximation for cosmological simulations. The electron mass fraction is derived from the hydrogen and helium mass fractions without accounting for the metal contribution that is very small.

The module presented above is the most general solution to the problem of defining the composition of the gas in my simulations. There is no need to modify any existing routine but the one initialising the composition data base, and it provides all the quantities required by the cooling routine. This module is used in the cosmological simulations in this thesis. For the simulations in Chapter 3, in which the metallicity of the gas is assumed

to be constant, I did not define the metal mass fraction. The metallicity value was set using a run-time parameter.

A module that avoids carrying all the seven variables, saving memory and disk space, has been written, but not intensively used. In this case only H^+, He^+, He^{++}, e^- and Z mass fractions are used. H and He mass fractions are calculated knowing the total abundances of the two elements. Then, Equation 2.3 is no longer valid, and the check that the mass fractions sum to the unity is no longer applied. This module has not been used for the simulations presented in this work, but will be adopted in future.

2.5.5. Heating module

The injection of bubbles of energy has been implemented adapting an existing routine in the `source_terms` module. A single bubble is defined by its position, initial radius $r_{\rm b}$, injection starting time $t_{\rm b}$, injection duration time $t_{\rm b}^{\rm inj}$ and energy $E_{\rm b}$. These values are in a parameter file called `mheat.dat`. The routine `tstep_heat` has been modified to have the first (last) time-step of energy injection matching $t_{\rm b}$ ($t_{\rm b} + t_{\rm b}^{\rm inj}$). Energy is injected over the time $t_{\rm b}^{\rm inj}$ at constant rate $\dot{E} = E/t_{\rm b}^{\rm inj}$. The energy is injected with a three-dimensional Gaussian distribution proportional to $e^{(x^2+y^2+z^2)/2\sigma^2}$ where $\sigma = r_{\rm b}/3$. The bubble is truncated at a maximum radius of 5σ. Thus, over several time steps the total energy of the bubble is injected in computational cells whitin a sphere of radius 5σ. The resolution of the AMR grid must be chosen to have a smooth representation of the volume filled by the bubble. The bubble can then expand until it reaches pressure equilibrium with the surrounding gas. The final size of the bubble is several times σ–the sphere of radius σ is the region in which most of the energy is dumped. Injecting uniformly the energy in cells whitin a sphere of given initial radius (top-hat) would have been another valid way of making the bubble. But the sharp edge of the sphere would be poorly resolved by the grid adopted. This would result in a too large error in energy conservation (of the order of several percent). Resolution is not an issue when smoothing the energy with a Gaussian

kernel because in the external shell of the sphere ($r > 3\sigma$) less than 1% of the energy is injected.

A co-moving frame is used in cosmological simulations and the scaling of the bubble quantities must be taken into account. Let $\mathbf{x}_\mathrm{b}$, r_b, E_b be the physical values of bubble position, radius and total energy, $a(t) = 1/(1 + z(t))$ the expansion factor at redshift z, t and $\mathrm{d}t$ the current time and time-step. The co-moving position and radius are then $\tilde{\mathbf{x}}_\mathrm{b} = \mathbf{x}_\mathrm{b}/a(t)$, and $\tilde{r}_\mathrm{b} = r_\mathrm{b}/a(t)$. At each time-step, the injection rate must be recomputed. The co-moving total energy is $\tilde{E}(t) = E/a(t)^2$ so that the injection rate is $\dot{\tilde{E}}(t) = \tilde{E}(t)/t_\mathrm{b}^\mathrm{inj}$. The amount of injected energy is then $\dot{\tilde{E}}(t)\mathrm{d}t$. In the cosmological simulations presented in Chapter 6, even though the initial conditions have been spatially shifted to have the cluster forming in the centre of the box, the cluster bulk velocity is not null. The centre of the cluster is then moving and bubble positions defined as above are no longer valid (unless one knows *a priori* the path of the cluster). Instead of bubble positions, I defined the offsets with respect to the centre of the cluster. The centre of the cluster is computed (in parallel) at each time-step, just before injecting the energy, using a shrinking sphere about the centre of the cluster. The centre of the cluster is saved to have a better initial guess at the next step.

2.6. Scaling tests

Performance tests on the computer facilities of the Institute of Computational Cosmology are presented in this section. I ran the scaling test that comes with the code, and a test based on a cosmological ΛCDM simulation.

2.6.1. Computer facilities

The facilities available at the Institute of Computational Cosmology are the three super-computers called TITANIA, CENTAUR and QUINTOR. Here is a brief description of each one.

TITANIA

TITANIA is a SUN shared-memory machine. It consists of 24 UltraSPARC-III 750 MHz processors sharing 48 GByte of memory. TITANIA is mainly used for developing codes, and for running serial codes that require huge amounts of memory (e.g. initial condition generator, friend-of-friend codes). It also manages the input/output to the tape storage unit and the disk space of 4.4 TByte shared by the three supercomputers.

CENTAUR

CENTAUR is a cluster of 64 SUNBlade 1000s connected with a 64-Port Mirynet 2000 switch. Each node consists of two 750 MHz UltraSPARC-III processors with 8 MByte of cache each, 1 GByte of RAM memory, 20 GByte of sequential data space, and Myrinet 2000 card (150-200 MByte/s).

QUINTOR

QUINTOR is a cluster of 259 SUNFire V210s nodes with a 2 Gbit interconnection. Each node consists of two 1002 Mhz UltraSPARC-IIIi processors with 1 MByte of cache each, 2 GByte of RAM memory, 40 GByte of sequential data space. 32 of the 259 nodes have 4 GByte of RAM. The 256 nodes devoted to computation are grouped in racks of 32 nodes. In any rack, groups of 16 nodes are connected with a fast switch. Another switch is used to connect the two partitions of 16 processors. A code running on processors in the same rack will perform at the maximum speed allowed by the network interface. Different racks are connected by a Gbit network that suffers of long latency time. This can cause a loss of 50% in performance for FLASH when ran on processors distributed among different racks (see discussion below).

FIGURE 2.4.: Results of the hydrodynamics scaling test on TITANIA, CENTAUR and QUINTOR machines. The scaling is almost linear for all the machines. The y axes are normalised to 8977, 11225 and 16255 cells s^{-1} speeds, respectively.

2.6.2. Built-in hydrodynamics scaling test

I first tested the behaviour of the hydrodynamical part of the code by running the built-in scaling test (`setups/scaling_test`) on the three supercomputers. The simulation does not include any gravity computation. The setup scales the size of the problem proportionally to the number of processors ensuring that the number of operations performed by each processor is constant in any run. Thus, the overhead introduced varying the number of processors should be caused only by communications *via* the MPI libraries. Communications involve the redistribution of blocks after each refinement through the Morton ordering, and the exchange of guard cell information at any time-step. One should expect that the code scales very well on each computer, because the number of communications is quite limited. The results of the tests are shown in Figure 2.4. The scaling of the code in very good in all the three cases. On QUINTOR the test runs were performed in a single rack to prevent the use of the slow Gbit channel. This is the best configuration available on the machine.

FIGURE 2.5.: Results of the cosmological scaling test on TITANIA, CENTAUR and QUINTOR machines. The scaling is super-linear for TITANIA. On CENTAUR the code scales reasonably well up to 16 processors, while on QUINTOR the performances are poorer because of the slow TCP/IP interconnection. The y axes are normalised to 807, 1009 and 1210 cells per second speeds, respectively.

2.6.3. Cosmological simulation scaling test

I also tested the performances of the `gravity` module, running the cosmological simulation presented in Chapter 6 for three time-steps. I started from the initial conditions at redshift $z = 63$, restricting them to two levels of refinement. The initial conditions contains gas densities on a uniform grid covering the whole simulation box, and one extra level of refinement covering one eighth of the volume. Dark matter particles on the fine grid have mass one eighth of those on the coarse grid. It may not be the ideal case since the distribution is very smooth, and one could expect the code to perform better than with a more complex AMR grid. The problem size is not scaled with the number of processors. However, the number of operations should be large enough to obtain a sensible comparison at least up to 16 processors. The `multigrid` Poisson solver is $\mathcal{O}(N)$. That is, it scales linearly with the number of cells N, and the only decrease in performance should derive from the parallel overhead of the communications. So, one should not expect the scaling to deviate from linear.

The results of the tests are shown in Figure 2.5. The test on TITANIA was performed on 1, 2, 4, 8, and 16 processors; the one on CENTAUR, due to memory limitation, was

performed on 8, 16 and 32 processors and normalised to the 1 processor run on TITANIA scaled by the speed ratio between the two machines showed in the hydrodynamics test; the test on QUINTOR was performed on 2, 4, 8, 16 and 32 processors. TITANIA does not perform any communication: the calls to MPI routines are converted in call to shared memory routines, and any processor can access directly any portion of the memory. The only overhead comes from the synchronisation of memory reading and writing. As result of this, the code scales very well up to 16 processors. Moreover, the scaling appears to be super-linear. That is, doubling the number of processors makes the code run at more than twice the speed. It is a well known behaviour on shared-memory machines, usually due to better cache fitting when varying the number of processors. When running on the distributed-memory computers, the parallel overhead becomes significant, and runs with a number of processors larger than 16 do not provide any significant performance improvement. However, increasing the size of the problem, the code should perform better and approach the linear regime. Taking into account that, for each processor domain, only the surface blocks need to exchange boundaries, one could argue that the time to perform a simulation step is

$$t = t_{\mathrm{p}} \frac{N}{n_{\mathrm{p}}} + t_{\mathrm{c}} \left(\frac{N}{n_{\mathrm{p}}} \right)^{\frac{2}{3}}$$

where t_{p} and t_{c} are the average computation and communication time for each cell, respectively, N is the total number of cells and n_{p} is the number of processors. The "speed" of the code is then

$$n_{\mathrm{p}} \frac{N/n_{\mathrm{p}}}{t_{\mathrm{p}}(N/n_{\mathrm{p}}) + t_{\mathrm{c}}(N/n_{\mathrm{p}})^{2/3}}$$

This leads to three different values depending on the ratio $t_{\mathrm{p}}/t_{\mathrm{c}}$:

$$\begin{cases} \dfrac{n_{\mathrm{p}}^{2/3}}{t_{\mathrm{c}}} N^{1/3} & \text{for } t_{\mathrm{p}}/t_{\mathrm{c}} \ll 1 \\ \dfrac{n_{\mathrm{p}}^{2/3}/t_{\mathrm{p}}}{1 + (n_{\mathrm{p}}/N)^{1/3}} & \text{for } t_{\mathrm{p}}/t_{\mathrm{c}} = 1 \\ \dfrac{n_{\mathrm{p}}}{t_{\mathrm{p}}} & \text{for } t_{\mathrm{p}}/t_{\mathrm{c}} \gg 1 \end{cases}$$

I tried another test with a cosmological simulation setup in which initial conditions were created with different resolution. Several simulations were timed on the same number of processors (16 and 32). Changing the resolution of the initial conditions, a sensible improvement on the speed of the code is achieved, but still the code does not perform better than the worst case in which $t_{\rm p}/t_{\rm c} \ll 1$. A large simulation ran on a small number of processors could be the ideal configuration for achieving the best performance on our distributed memory machines. However, one would like to take advantage of the large number of processors of these machines to perform the same simulations in shorter time and, ideally, to use all the available processors at the same time.

One possible explanation of this behaviour is that the `multigrid` module requires many calls per time-step to the routine that updates the guard cells of blocks, `mg_bndry`. This involves a large number of simultaneous communications that could saturate the slow Gbit network interconnection, and appears as a *latency* problem. Latency is the time interval between a communication request and the start of data transfer. In the ideal case of communication between only two processors, latency time is usually small, depending only on the network physical architecture. When many processors are communicating at the same time through a single network channel, a processor has to wait until the channel is free before starting the data transfer. Therefore, latency appears to increase dramatically. In Table 2.1, I show the timing of the cosmological simulation tests. Looking at the test on TITANIA, one can easily recognize that the hydro and gravity computational time scales linearly with the number of processors (the 16 processors run deviates from linear scaling for the reasons discussed above). The ratio between gravity and hydro times (fifth column) decreases with the number of processors, but it is still far from what was claimed by the FLASH developers; in their perfromance tests, they find this ratio to be close to one. I observed an even larger value in the simulations of Chapter 6, in which the ratio is about twenty. So, this could be just related to the choice of the simulated problem. The most interesting result is that the boundary cells updating time is approximately the same

N.proc	hydro	gravity	mg_bndry	g/h	b/g
		TITANIA			
1	333.189	2801.71	835.4	8.4	29.8%
2	169.354	1358.56	422.6	8.0	31.1%
4	91.168	556.376	152.7	6.1	27.3%
8	47.138	238.285	68.6	5.1	28.8%
16	30.184	172.525	65.0	5.7	37.8%
		QUINTOR			
2	125.155	800.844	269.9	6.4	33.6%
4	83.226	514.049	257.3	6.2	50.0%
8	52.926	336.242	194.7	6.3	57.8%
16	37.876	270.472	183.2	7.1	67.8%
32	39.084	356.062	225.6	9.1	63.4%

TABLE 2.1.: The timing of the cosmological simulation tests for TITANIA and QUINTOR. Times are in seconds.

fraction of the total time spent computing the gravity ($\simeq$ 30%); this is shown in the sixth column. This time scales linearly with the number of processor (or with the number of blocks per processor, since the total number of blocks is constant). Everything looks just as one should expect because in these runs the parallel overhead is negligible. What happens on a distributed memory machine is quite different. In the second part of Table 2.1 I show the results of the tests on QUINTOR. Hydro and gravity do not scale linearly, and the fraction of time spent updating blocks boundaries increases with the number of processors. In the next section I describe some possible performance improvements.

2.7. Performance improvements

When the performance issue arose, I tried to investigate some possible solutions. The first attempt was to redesign the routines that exchange the guard cells. These routines use a common communication scheme. All the receive calls are non-blocking and done all together in a loop. Since non-blocking calls have been issued, processors can move to the next step and start to send data to other processors. Each send call is blocking, thus each processor has to wait until its completion. Once all the data are sent, a check is necessary to make sure all the data has been received. This scheme has the disadvantage that a processor could await long time before sending its data if the receiving processor is busy communicating with another processor. I tried to avoid this by making the routines more general and introducing non-blocking send calls. In this way, any processor can send and receive data at any time, as soon as another processor is free to communicate. The performance improvement was insignificant, suggesting that the problem is intimately connected to the number of communications. A better solution could be to reduce the number of communications by exchanging more data every single time. This would partially avoid the latency problem. On the other hand, the code would require a larger amount of memory. However, given the structure of the data that are exchanged, it is not a trivial task to find a way of packing them into a single array and the attempts I made did not work.

Some authors have suggested another way of approaching the problem. The red-black ordering scheme of the Gauss-Seidel iterative solver requires the values in the $(i-1)$-th and $(i+1)$-th cells to compute the value in the i-th cell. If all the guard cell are filled, it could be possible to perform four iterations before updating their values. So, the communication required should be reduced by a factor of four. However, in their test, they do not show any performance improvement. I will investigate this in the future, but I am more interested in trying the approach discussed next.

Another possible solution to the performance problems of the gravity solver is to change the algorithm partially. I have been working on a new implementation of it, but I have

not yet completed it. The idea is as follow. Many codes compute the gravitational potential on a uniform grid with *fast Fourier transformation* (FFT) techniques. Using the FFTW libraries, the FFT can be also performed in parallel with a small amount of extra coding required. So, at each time-step the density could be mapped on a uniform grid, and the gravitational potential computed. The grid could be placed at the maximum level of refinement covering the entire computational volume. The solution at higher levels of refinement could be calculated in two ways: (1) the solution could be computed on higher levels of refinement using Gauss-Seidel iterations; (2) the solution could be propagated to a higher level of refinement by a simple interpolation. In the first case the accuracy would be improved. However, the performance improvement introduced by the FFT could be reduced by the relaxation phase. Since an accurate solution of the gravitational potential is achieved with a resolution usually lower than the one required by the hydrodynamics solver, the Gauss-Seidel iteration, and all the communications it requires, could be avoided by just propagating the solution. Moreover, in many cases, one is more interested on resolving the hydrodynamics than on achieving high resolution for the gravity. For example, this is the case of the simulations in Chapter 6. Lower accuracy could be obtained by limiting the FLASH solver to computing the gravity to a level of refinement lower than the maximum, instead of using the FFT. But, again, the performance improvement would be reduced by the relaxation algorithm.

2.8. Conclusions

In summary, I have introduced in this chapter the concept of adaptive mesh refinement and the algorithms used by FLASH for solving the hydrodynamics and gravity equations on the AMR grid. I also presented the routines and modules I have addied to the code, giving more emphasis to the refinement criteria and cooling routines. Several performance tests have been described. These tests were ran on the computer facilities of the Institute of Computational Cosmology. They present puzzling preformance degradation issues when

FLASH is ran on the distributed memory machines. Even though it seems clear this problem is related to the network configuration of the super-computer, many doubts on the implementations of the gravity solver are still unresolved. I have suggested some ways of solving the speed problem of the gravity solver and, in conclusion, the use of FFT techniques with and without extra relaxation is what I will explore in the future.

(...) numerous theoretical models in which the gas is reheated in some fashion: by powerful radio jets, by rising bubbles, by shocks, by magic, etc.

Mathews, W., Brighenti F., and Buote, D. (2004)

I like to think I am doing magic.

Quenching cluster cooling flows with recurrent AGN bubbles

The contents of this chapter are as in the work published in Dalla Vecchia et al. (2004). Some parts have been repeated and developed in other chapters of the thesis. An extra section has been added at the end of the chapter (Section 3.6), and the appendix to the paper has been extended and moved to the end of the thesis (Appendix B).

3.1. Introduction

Gas cooling at the centre of a cluster halo is an inherently unstable process: cooling increases the gas density, which in turn enhances the cooling rate. X-ray observations show that the cooling time of gas in most cluster cores is less than the Hubble time (e.g. Fabian and Nulsen, 1977). Unless cooling is balanced by some form of heating, gas will flow into the cluster centre at rates up to $\sim$ 1000 $M_\odot$ yr^{-1} (e.g. Kaastra et al., 2001). What happens to the cooled gas is unclear; it could be consumed by star formation, or lead to a reservoir of low temperature ($<$ 1 keV) material in the core. Although this may be the fate of a fraction of the gas, there are indications that most gas in fact does not follow either route. First, the star formation rates in the central galaxies of clusters are much lower than the inferred mass inflow rates (O'Connell and McNamara, 1989; Johnstone et al., 1987), rarely approaching $\sim$ 100 $M_\odot$ yr^{-1}. To estimate the star formation rate in a typical cluster, one can average over the large sample of clusters studied by Crawford et al. (1999). This suggests that the typical star formation rate is less than 10 $M_\odot$ yr^{-1}. Secondly, we can compare the mass deposition rates with observations of the molecular gas content of clusters. Edge (2001) finds molecular gas masses ranging from 5×10^8 to 2.5×10^{11} $M_\odot$, with an average of 2.6×10^{10} $M_\odot$. Assuming a gas consumption timescale of 10^9 years (see Edge, 2001), this implies a deposition rate that may be as high as 200 $M_\odot$ yr^{-1} in a few clusters, but is $\sim$ 30 $M_\odot$ yr^{-1} on average. Similar limits are obtained using the data from Salomé and Combes (2003). Finally, recent spectroscopic X-ray observations show no evidence for significant gas cooling below 1 keV (Kaastra et al., 2001; Peterson et al., 2001; Tamura et al., 2001), and observations of molecular and neutral material reveal that the amount of cold gas in clusters of galaxies is also much less than expected from the integrated cooling flow rate (typically less than 30 $M_\odot$ yr^{-1} – Edge et al., 2002; Edge and Frayer, 2003; Salomé and Combes, 2003).

This cooling-flow paradox has led many authors to investigate mechanisms to quench gas cooling. Observations of merging clusters show little evidence for cooling flows, which

suggests that the merger process might be implicated in disrupting cooling. Recent simulations have shown that sub-halo merging can indeed heat up gas (Burns et al., 2004); however, the amount of cold gas produced in these simulations is still too large compared to that observed, leading to the conclusion that additional heating processes must be involved. Several alternatives have been proposed, including energy injection from radio sources or active galactic nuclei (AGN; Binney and Tabor, 1995; Quilis et al., 2001; Churazov et al., 2002), viscous dissipation of sound waves (Fabian et al., 2003a; Ruszkowski et al., 2004a) and thermal conduction (Voigt and Fabian, 2004; Dolag et al., 2004). Each of these has advantages and disadvantages. In particular, it is difficult to balance cooling, with its ρ^2 density dependence, with heating processes that typically scale as ρ. Since cooling and heating can then balance only at one particular density, some of these feedback mechanisms may lead to unstable solutions with some regions of the cluster being efficiently heated while others continue to cool catastrophically.

AGN are particularly promising candidates for balancing cooling, given their potentially large energy reservoir (Tabor and Binney, 1993; Binney and Tabor, 1995; Bower et al., 2001). In most numerical simulations of this process to date, energy injection produces bubbles at high temperature and low density that, after a short expansion phase, gain momentum by buoyancy (Brüggen and Kaiser, 2001; Quilis et al., 2001; Basson and Alexander, 2003). This mimics the effect of a jet which rapidly loses its collimation, as often observed in local clusters (Eilek, 2004). In other simulations, gas is injected at high velocity to mimic jets which retain their large-scale coherence (Reynolds et al., 2002; Omma et al., 2004). In this chapter, I consider the first mechanism and show that such bursts of localised energy can induce convection of the intra-cluster medium (ICM), which leads to a quasi-stable cluster configuration and reduces the average mass deposition rate to within observational limits. I discuss whether the required energy and duty cycle of AGN activity are compatible with observational limits. This chapter is organised as follows. In Section 3.2 I describe the setup of my simulations and discuss the parameters explored.

Results are presented in Section 3.3 and, in Section 3.4, I compare the required energy and duty cycle with observational constraints. Finally, my conclusions are summarised in Section 3.5.

3.2. Simulations

In order to concentrate the numerical resolution on the interactions between the cooling material and the buoyant hot bubbles, my simulations use a fixed external gravitational potential and neglect the self-gravity of the gas. This allows us to run the simulation over cosmologically significant timescales. I model the input of energy as a cycle of energy injection and quiescent phases. In the active phase a small region of the gas is heated so that it expands to form a bubble that rises out of the confining potential.

3.2.1. The code

My simulations were performed with FLASH 2.2, an Adaptive Mesh Refinement (AMR) hydrodynamical code developed and made public by the ASC/Alliance Center at the University of Chicago (Fryxell et al., 2000). FLASH is a modular block structured AMR code, parallelised using the Message Passing Interface (MPI) library.

FLASH solves the Riemann problem on a Cartesian grid using the Piecewise-Parabolic Method (PPM; Woodward and Colella, 1984). It uses a criterion based on the dimensionless second derivative $\mathcal{D}^2 \equiv |(Fd^2F/dx^2)/(dF/dx)^2|$ of a fluid variable F to increase the resolution adaptively whenever $\mathcal{D}^2 > c_2$ and de-refine the grid when $\mathcal{D}^2 < c_1$, where $c_{1,2}$ are tolerance parameters. When a region requires refining ($\mathcal{D}^2 > c_2$), child grids with cell size half that of the parent grid are placed over the offending region, and the coarse solution is interpolated. In my simulations, I have used density and temperature as the refinement fluid variable F (see Fryxell et al., 2000, for more details). I have performed a series of simulations with a different maximum level of refinement to investigate the effect of numerical resolution, as described in Section 3.2.4

FLASH interpolates the imposed analytic gravitational potential (see Section 3.2.2) to the grid cells, and computes the corresponding gravitational acceleration, neglecting the self-gravity of the gas. To interpolate the initial density field from the analytical solution to the FLASH mesh, I impose an initial grid with increasing resolution from large radii to the centre of the cluster. This is the minimum refinement level allowed during the simulation.

For the boundary conditions, I impose that the values of density, temperature and velocity on the boundary cells (guard cells) remain the ones computed at the initial time and satisfy the hydrostatic equilibrium conditions. With the cluster located at the centre of the box, the box size of 6×10^{24} cm (1.9 Mpc) is large enough to ensure that cooling does not affect the boundaries. This guarantees the hydrostatic equilibrium condition is not broken, and no evident inflow or outflow occurs at the border of the grid. Moreover, the temperature difference between the last computational cell and the guard cell next to it, at the end of the simulation, is always less that 2%. Likewise, sound waves propagating across the volume do not generate any unphysical artifacts at the boundary.

The time step is derived by the Courant condition $\mathrm{d}t = C\Delta x/c_s$, where Δx is the dimension of a cell, c_s is the sound speed in that cell and C is a coefficient, usually less than one. The time step is uniform; thus, all the cells evolve at each time step.

3.2.2. Initial conditions

The time-independent gravitational potential of the dark matter is that corresponding to a Navarro et al. (1996, 1997) density profile

$$\rho_{\mathrm{DM}}/\rho_{\mathrm{crit}} = \frac{\delta_{\mathrm{c}}}{(r/r_{\mathrm{s}})(1+r/r_{\mathrm{s}})^2},$$

where ρ_{crit} is the critical density, $r_{\mathrm{s}} = r_{200}/c$ and $r_{200} = 1.38$ Mpc. I adopted a halo mass of $M_{200} = 3 \times 10^{14}$ $\mathrm{M}_{\odot}$ and set the concentration parameter to $c = 6$ (Eke et al., 2001; Wechsler et al., 2002; Dolag et al., 2004).

I use the prescription of Wu et al. (2000) to set up an isothermal gas distribution in hydrostatic equilibrium within this dark matter potential. I adopted this prescription because

of its simplicity. Fits to numerical simulations (Komatsu and Seljak, 2001), suggest that the temperature should decline with increasing radius. However, my isothermal profiles do match the observed temperature structure of clusters within $0.2r_{200}$, where the observational data show an isothermal temperature profile (eg. Allen et al., 2001; De Grandi and Molendi, 2001; Pointecouteau et al., 2004) sometimes with a central decrement. A central decrement is quickly generated as the model cluster cools.

I choose an initial gas temperature of $T = 3.1$ keV, consistent with the observed correlation between temperature and M_{200} (Reiprich and Böhringer, 2002). The appropriate gas density is then

$$\rho_{\rm GAS}/\rho_0 = (1 + r/r_{\rm s})^{\eta/(r/r_{\rm s})},$$

where ρ_0 was chosen to satisfy the relation

$$M^{200}_{\rm GAS}/M^{200}_{\rm DM} = \Omega_{\rm b}/(\Omega_{\rm m} - \Omega_{\rm b}),$$

Here, $M^{200}_{\rm DM}$ denotes dark matter mass contained within a sphere whose mean interior mass density is 200 times the critical density. I used $\eta = 10.25$, $\Omega_{\rm m} = 0.3$ and $\Omega_{\rm b} = 0.04$ for the matter and baryon densities in units of $\rho_{\rm crit}$, and $h = 0.7$ for the Hubble parameter.

I adopt the cooling functions of Theuns et al. (2002), which include cooling by H, He and metals, in the presence of an ionising UV-background. This cooling function uses interpolation tables for the cooling rate due to a solar admixture of metals, obtained from CLOUDY (version 94, Ferland et al., 1998). I assume a metallicity of one third solar, and the Haardt and Madau (1996) UV-background at redshift $z = 0$, as updated by Haardt and Madau (2001). The cooling routine was tested against the MEKAL (Mewe et al., 1986) cooling routines in the X-ray analysis package XSPEC (Kaastra, 1992) and against the cooling rates used by McCarthy et al. (2004).

Heating by an AGN is modelled by injecting energy at predefined points in space and time. This allows us to simulate multiple episodes of AGN activity as I now describe.

3.2.3. Energy injection

To simulate the effect of a central AGN, bubbles of energy are distributed randomly inside a radius of 50 kpc from the centre. I do not attempt to model the processes leading to the formation of the bubbles, but I consider it likely that the bubbles will be injected at different positions because of the precessional motion of the jet axis of the central source.

Energy is injected with a three-dimensional Gaussian distribution proportional to $e^{(x^2+y^2+z^2)/2\sigma^2}$ and I take the initial size of the bubble to be characterised by $\sigma = 10.3$ kpc. The bubble is truncated at a maximum radius of $r_{\rm b} = 5\sigma = 51.5$ kpc, which is seven cells in my standard six level resolution runs. The final radius (after the expansion of the bubble) depends on the amount of energy injected.

Starting from $t = 0$, energy is injected every 10^8 yr for a time of 10^7 yr in my standard simulations. I assume energy values from 1×10^{59} to 3×10^{60} erg (Table 3.1). My choice of parameters deliberately lies at the upper end of observed AGN duty cycles and jet powers (Willott et al., 1999; Owen et al., 2000; Fabian et al., 2003a; Eilek, 2004).

3.2.4. The simulations

My main results are derived from a series of nine simulations (Table 3.1):

1. an adiabatic test simulation (A) without cooling or AGN bubbles;
2. a simulation with cooling but no AGN bubbles (S0.0);
3. seven simulations with cooling and AGN bubbles of varying strength (S0.1–S3.0), but the same duty cycle. In these simulations only the energy of each bubble varies. The position and timing of the heating events are the same.

The results of these simulations are presented in Section 3.3.

I also performed simulations with the following changes in order to test the sensitivity of my results to the details of the heating cycle:

4. the bubbles were randomly distributed inside a sphere of 25 or 100 kpc, instead of 50 kpc;

5. the same total number of bubbles (15) was generated within 1.5 Gyr, but at random time intervals;

6. the energy was injected as bubble pairs, with each bubble containing half the energy per event.

Finally, I also reran simulation S2.0 with increased resolution to check for numerical convergence. The results of this test, described in Appendix B, led us to adopt a maximum refinement level of six. I also continued the simulation S2.0 to 5 Gyr to test the long term stability of this simulated cluster.

In the adiabatic simulation, the gas remains almost static as expected for my equilibrium set-up, apart from small re-adjustments induced by discretisation. Energy and mass are conserved to $\simeq 0.4\%$ and 1% respectively. With cooling allowed (simulation S0.0), the loss of pressure support causes a cooling flow to be established.

The initial conditions have a relatively high X-ray luminosity ($\log L_{\rm X,bol} = 45.0$) compared to the observed X-ray luminosities of clusters of this mass and temperature (Markevitch, 1998; McCarthy et al., 2004). This is a well-known problem that results from the low entropy gas in the centre of a cuspy dark matter potential. Starting with such a luminous cluster ensures that my simulations conservatively explore the maximum energy necessary to stabilise cooling. In order to be fully successful, however, my energy injection must not only balance the cooling, but also raise the entropy of the cluster gas sufficiently to reduce the total luminosity by a factor ~ 10.

By default, I ran each simulation for 1.5 Gyr. Without a cosmological context, a longer simulation is not well justified since cluster-cluster mergers provide an important additional mechanism for re-organising the gas distribution and erasing the cooling flow structure. Nevertheless, I ran the S2.0 simulation for 5 Gyr to check the evolution of the solution over very long time scales.

		INJECTED/RADIATED POWER	
name	E	$\dot{E}_{\rm i}$	$< \dot{E}_{\rm r} >_{100}$
	10^{60} erg	10^{44} erg s^{-1}	10^{44} erg s^{-1}
A	0	0	no cooling
S0.0	0	0	28
S0.1	0.1	0.32	18
S0.3	0.3	0.95	18
S0.6	0.6	1.9	13
S1.0	1.0	3.2	13
S1.5	1.5	4.7	8.7
S2.0	2.0	6.3	7.9
S3.0	3.0	9.5	6.8

TABLE 3.1.: Summary of the nine simulations performed. The different simulations are referred to as Sn, where the energy E of a single bubble is $n \times 10^{60}$ erg. $\dot{E}_{\rm i}$ is the mean energy injection rate over a duty cycle, and $< \dot{E}_{\rm r} >_{100}$ is the mean emitted energy rate, averaged over the last 10^8 yr of the simulation.

3.3. Results

3.3.1. The effect of energy injection

In Figure 3.1 I show, for each simulation, the variation in the total energy within the simulation volume. The pulsed injection of the energy and subsequent cooling gives rise to a saw-tooth pattern.

In simulation S0.0, which has no heating, the cooling rate increases with time as the cluster collapses. Including energy injection (simulations S0.1–S3.0) reduces the overall amount of energy that is radiated. This is not a trivial result since the energy injection events can actually promote cooling in the compressed regions around the rising bubbles. In practice, however, the dominant effect is the mixing of high and low entropy material

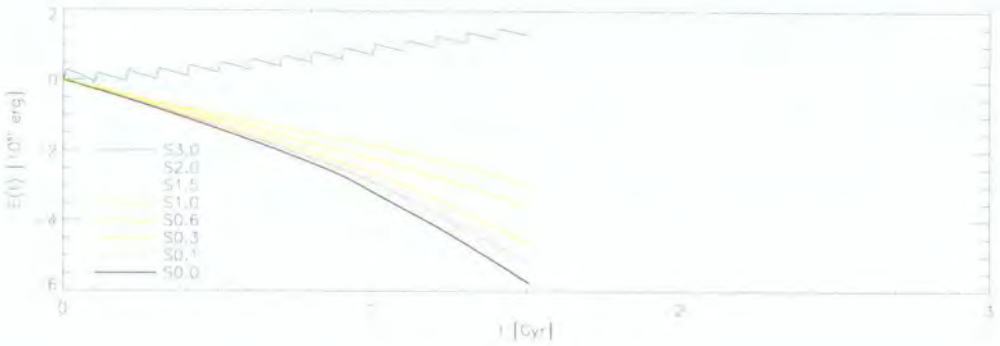

FIGURE 3.1.: The evolution of the total energy of each simulation is shown by plotting $\Delta E_{\rm T}(t) = E_{\rm T}(t) - E_{\rm T}(0)$, where $E_{\rm T}(t)$ is the sum of internal, kinetic and potential energy at time t and $E_{\rm T}(0)$ is its initial value. The saw-tooth shape of the curves results from the discrete AGN events and subsequent cooling. At the mean injection rate of simulation S2.0, the energy keeps an almost constant value within the simulation time. I tested that this behaviour is maintained up to 5 Gyr, though I show the evolution only to 3 Gyr for clarity.

in the large-scale gas motions induced by the bubbles. Therefore, as the injected energy is increased, the amount of radiated energy drops. At a mean energy injection rate of 6.3×10^{44} erg s^{-1} (simulation S2.0), the radiated energy and the injected energy are balanced. This balance does not guarantee that the structure of the cluster is stable, as continuing collapse of the central region might be balanced by expansion of the outer part of the cluster. However, as Figure 3.2 shows, the whole entropy profile of the cluster changes little. I investigated the long term behaviour of this model by continuing this simulation for 5 Gyr (Figure 3.1) and found that the cluster maintains this behaviour even over this longer timescale. Of course, over such long timescales, the effects of cluster mergers can not be neglected.

The emissivity-weighted temperature and entropy profiles of each simulation are shown, for 50 output times, in Figure 3.2. I define the entropy as $K \equiv kT_{\rm e}n_{\rm e}^{-2/3}$ keV cm^2, where $T_{\rm e}$ and $n_{\rm e}$ are the electron temperature and density, respectively. This is related to the thermodynamic entropy by a logarithm and additive constant. If cooling is not balanced by the energy injection, the entropy profile drops dramatically in the centre (top-

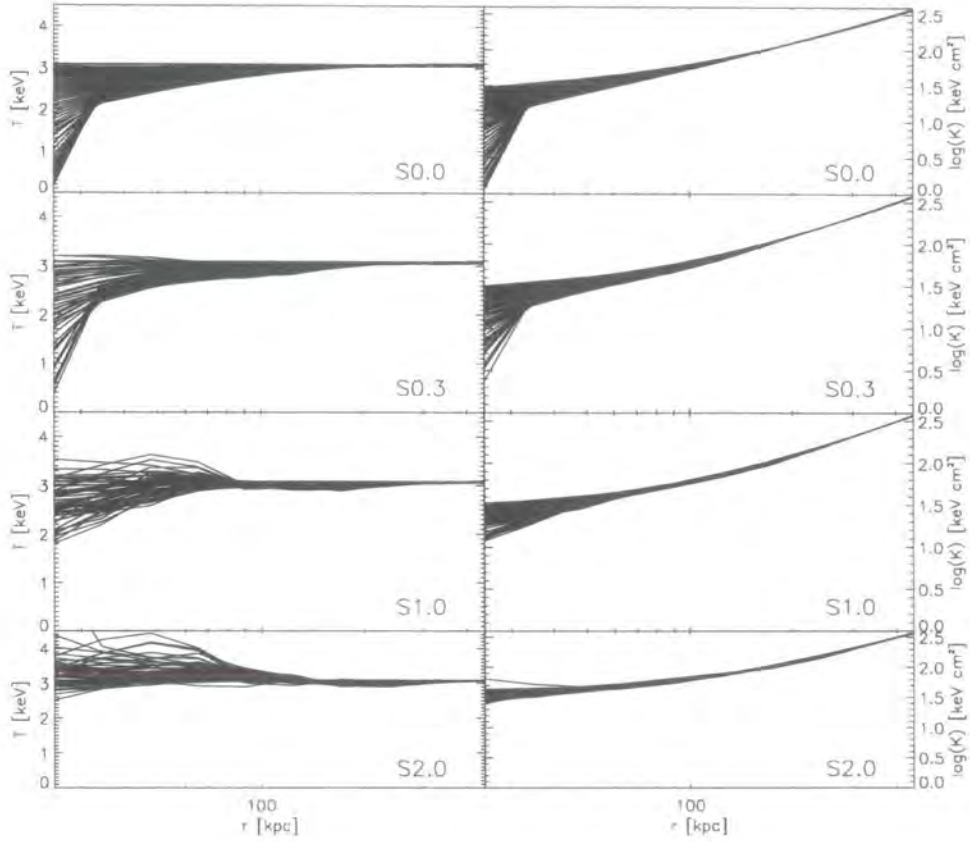

FIGURE 3.2.: The emissivity–weighted temperature (left panels) and entropy (right panels) profiles of each simulated cluster are shown for 50 simulation times. From top to bottom, simulations S0.0 (without any AGN bubble), S0.3, S1.0 and S2.0 are shown. The effect of energy injection is to maintain the initial, relatively shallow entropy profile and isothermal temperature in the cluster core.

right plot). This is reflected as a marked dip in the temperature profile of the cluster (top-left plot). Energy injection reduces this trend, maintaining the relatively flat entropy profile and isothermal temperature of the initial cluster.

The cooling of material in the central cluster regions leads to a net inflow of gas. In Figure 3.3, I plot the mass accumulation rate within the central regions of my simulated clusters. The rate is determined by averaging over the last 0.3 Gyr of the simulation, and is computed within the central 100 kpc and 50 kpc. The results for both radii are

similar, as expected if the cluster is in a quasi-steady state. The figure also shows the mass accumulation rate of material with temperatures less than 1 keV. In the absence of energy injection, the cooling cluster deposits material in the cluster centre at a rate that exceeds 1000 $M_\odot$ yr^{-1}, a factor $\sim$ 10–100 higher than the observational limits (Crawford et al., 1999; Edge, 2001; Salomé and Combes, 2003). Similarly, this model has a large mass flux of material cooling below 1 keV (1/3 of the ambient temperature).

Energy input at an average rate of 4.7×10^{44} erg s^{-1} (S1.5) reduces the deposition rate below 30 $M_\odot$ yr^{-1} at the end of the simulation, as required by observations (Peterson et al., 2001; Tamura et al., 2001; Kaastra et al., 2001; Peterson et al., 2003). Even with a lower energy injection rate of 3.2×10^{44} erg s^{-1} (S1.0), the rate at which material cools below 1 keV is still consistent with observed X-ray spectra. I conclude that energy injection through the creation of buoyant bubbles can successfully prevent catastrophic cooling.

3.3.2. Two-dimensional morphologies

In Figure 3.4, I show the time evolution of various quantities in two-dimensional sections of simulation S2.0. The temperature and entropy distributions are shown for a slice in the x-y plane. They show that cold, low entropy gas is pushed to large radii by the bubble's buoyancy and mixed with gas at those radii. Thus, the cooling rate is reduced not by direct heating of the cooling gas, but by convective transport of this material to regions of lower pressure. Similar results have been seen in simulations of single bubbles (e.g. Quilis et al., 2001; Churazov et al., 2002).

The temperature distribution also reveals the presence of sound waves propagating through the ICM. The sound waves are almost concentric and regular, a consequence of the periodic energy injection events near the cluster centre.

In order to illustrate what might be observable with an idealised X-ray satellite, I also calculated the projection of the quantity $\rho^2 T^{1/2}$, which is approximately proportional to the bolometric gas emissivity. Despite the large amounts of energy being dumped into the cen-

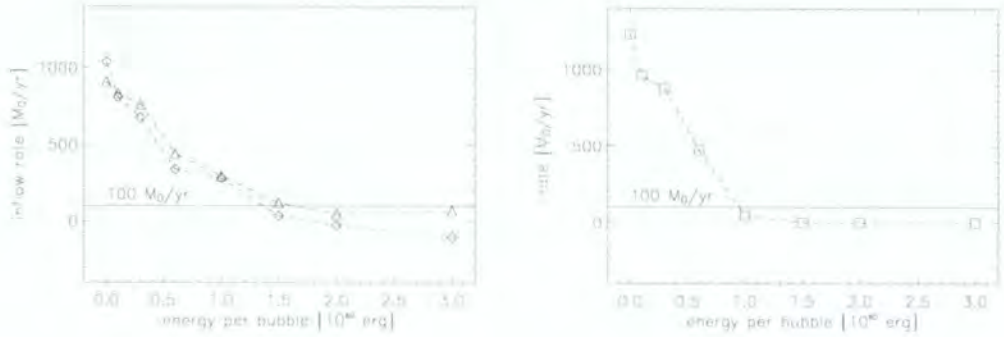

FIGURE 3.3.: **Left panel**: The mass deposition rate into a sphere of radius 50 kpc (dashed line) and 100 kpc (dash dotted line), averaged over the time interval 1.2 to 1.5 Gyr, as a function of the energy injected per bubble. For bubbles of energy above 10^{60} erg the deposition rate is $\lesssim 100$ $M_\odot$ yr^{-1}. **Right panel**: The mean growth rate of the total mass at temperature < 1 keV for each simulation. The amount of gas cooling below 1 keV drops down to almost zero when the injected energy per bubble is greater than 10^{60} erg.

tral regions, the large-scale gas distribution appears smooth and undisturbed in projection. To reveal the presence of perturbations on this smooth profile, I created an unsharp masked image through the transformation $(\mathrm{img}_0 - \mathrm{img}_s)/\mathrm{img}_s$, where img_0 and img_s are the original image and its smoothed version, respectively. The initial image was taken from a high resolution simulation ($\sim$ 2 kpc) and the final unsharped image is shown in Figure 3.5. The contrast in this image is high enough to show the sound waves distinctly as *ripples*. As suggested by Fabian et al. (2003a), these ripples in X-ray images of observed clusters could be the fingerprint of multiple outbursts and could be used to measure their periodicity. Furthermore, if the cluster gas is sufficiently viscous, these sound waves might help to offset the cooling flow by directly heating the gas (Fabian et al., 2003a; Ruszkowski et al., 2004b). By comparing the properties of successive ripples it may be possible to measure the viscosity of the ICM. Although the ripples in my simulation extend to very large radii, as result of the small viscosity in the simulations, it is unlikely that realistic observations would have sufficient signal–to–noise ratios to see these weak features at such distances. I will investigate this further in Chapter 4, where I synthesise artificial X-ray images, with

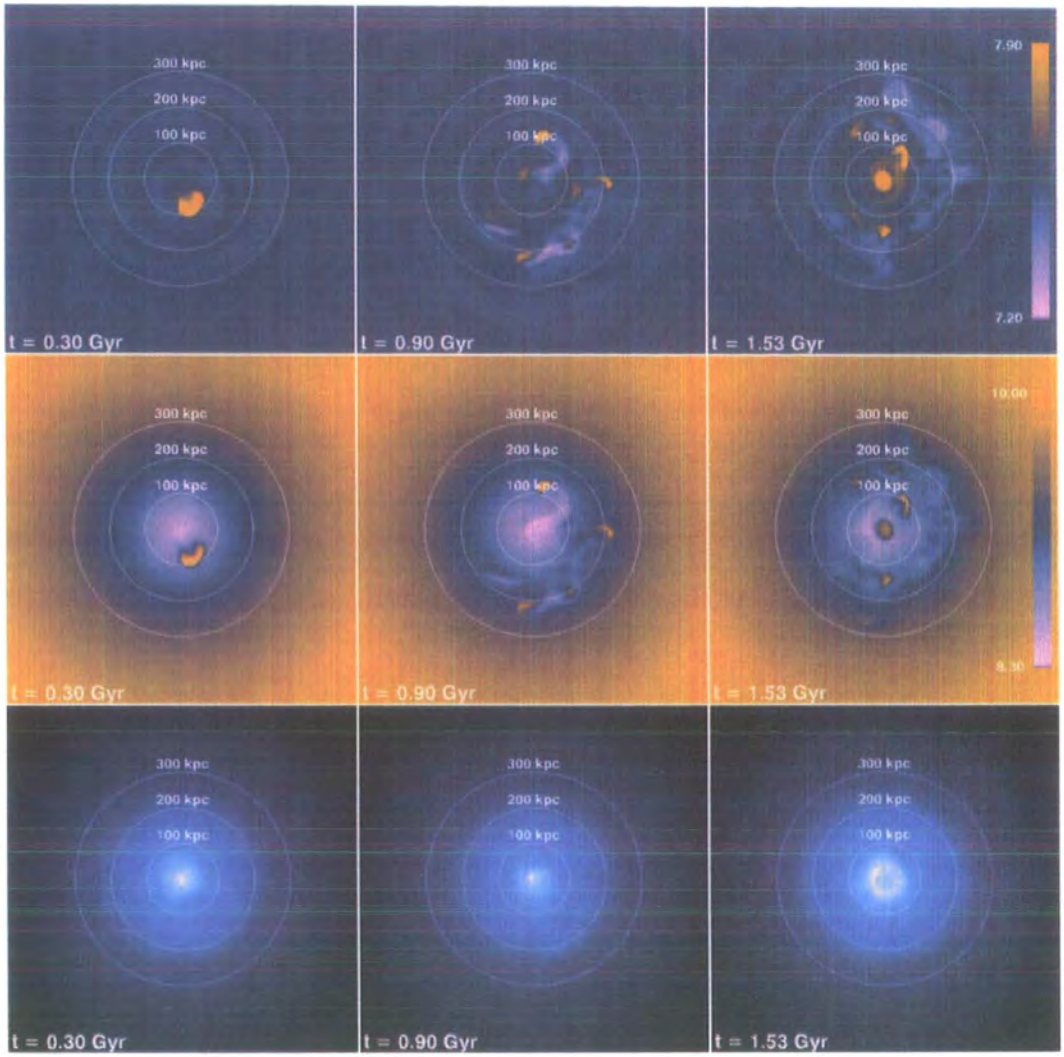

FIGURE 3.4.: The time evolution (from left to right) of simulation S2.0. From top to bottom, the quantities shown are the temperature [K] and entropy [ergs $g^{-2/3}$ cm^2] on the x-y plane crossing the centre. In the bottom row I show the approximate bolometric emissivity $\rho^2 T^{1/2}$ projected through the simulation volume. The temperature distribution reveals the presence of sound waves propagating through the ICM. The sound waves are almost concentric and regular, a consequence of the periodic energy injection events near the cluster centre.

the appropriate transmission functions, from my simulations.

FIGURE 3.5.: An unsharp masked projection of the approximated X-ray emissivity, $\rho^2 T^{1/2}$, for the central (0.9 Mpc diameter) part of the cluster simulated at a maximum resolution of $\sim$ 2 kpc. The "granular" structure of the image is due to the resolution of the simulation grid. The ripples from successive energy injection events can be clearly seen.

3.3.3. Sensitivity to duty cycle parameters

Simulations S0.1–S3.0 were based on a regular injection of energy. This is somewhat artificial, so I investigated whether randomly–timed energy injection events would have the same effect. I also investigated what happens if the energy injection occurs in less frequent bursts of greater intensity, or if bubbles are distributed inside a sphere of larger (100 kpc instead of 50 kpc) or smaller (25 kpc) radius. All my tests were performed choosing the parameters of simulation S2.0.

The resulting energy evolution of each simulation is illustrated in Figure 3.6. First, I double the time between bursts, while keeping the total average energy rate constant; thus, each bubble has twice as much energy injection as in S2.0. In this case, the variation in total energy during each cycle is much larger than before. However, the long term evolution is very similar. Randomising the injection time does not have a large effect on the average energy evolution over the duration of the simulation either. Where a particularly long time

interval elapses between bubbles, the energy drops further as cooling progresses; however this is offset by later events where bubbles appear in quick succession. My results are also insensitive to whether the bubbles are injected singly, or in pairs with the total energy divided between the two bubbles, possibly a better model for the AGN activity. The main factor that does make a significant difference is the location of the bubbles. Placing the bubbles initially within a larger radius (100 kpc) is much less effective at disrupting the cooling flow, since these bubbles tend to rise buoyantly without disturbing the central, cooling gas. By the end of this simulation, the total cluster energy has dropped below that of simulation S1.5 and is falling rapidly. Conversely, placing all the bubbles within 25 kpc is much more effective at preventing cooling, and the total energy of the system actually increases with time. To summarise, I see that the energy balance is most sensitive to the total amount of injected energy, and to the initial location of the bubbles.

3.4. Discussion

3.4.1. Energy requirements

Most central galaxies in clusters are radio sources (Ledlow and Owen, 1996); I can therefore expect that events like those I have simulated will play an important role in cluster evolution. I have shown that the cluster cooling rate can be reduced so that it agrees with observations if the AGN activity is sufficiently energetic. I have modelled the energy injection events as short outbursts separated by relatively long quiescent periods in order to make an extreme test of the long term effect of the energy injection mechanism. To stabilise the long term evolution of the model cluster in this way, I require an energy of $\sim 2 \times 10^{60}$ erg for each of the injection events, if the bubbles are distributed within a ~ 50 kpc radius. I can associate these events with outbursts of powerful AGN activity with a duration of $\sim 10^7$ yr and a power of $\sim 6 \times 10^{45}$ erg s^{-1}.

The observed radio power of central galaxies varies by factors of ~ 100 for clusters with

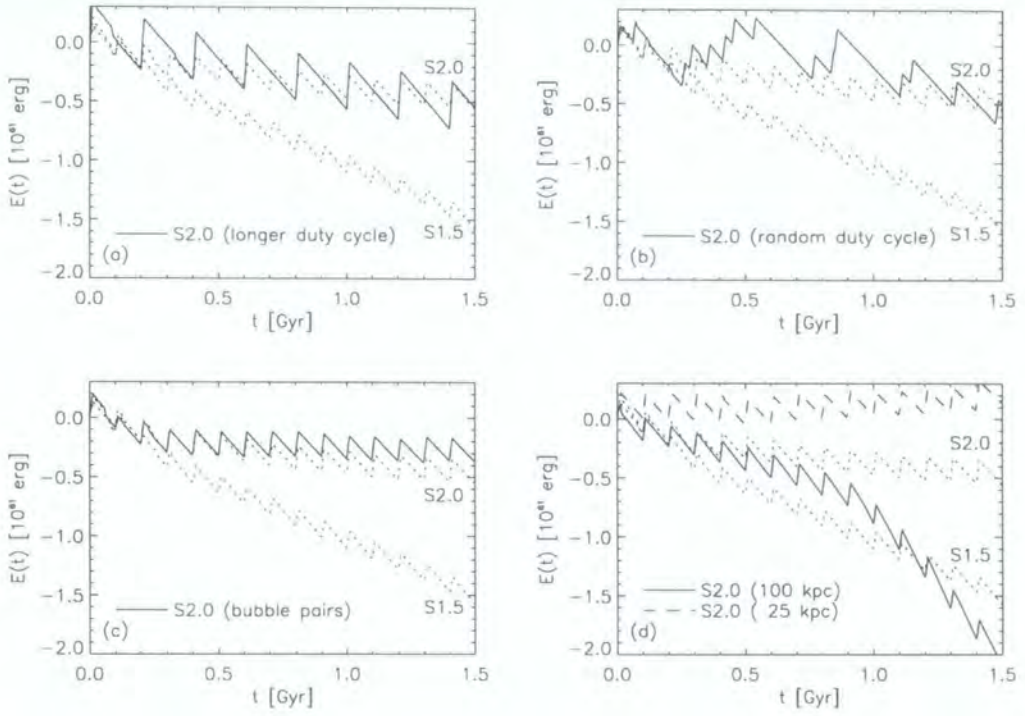

FIGURE 3.6.: In each panel I show the evolution of the total energy in simulations S2.0 and S1.5 (dotted lines), as in Figure 3.1. The other lines represent models with the same total energy as simulation S2.0, but with different parameters for the energy injection mechanism. In panel (a) I show a simulation where the duty cycle is regular, but twice as long as in S2.0. In panel (b) I show the effect of randomising the time interval between injection events. In panel (c) I show the effect of injecting the energy within pairs of bubbles. Finally, in panel (d) I show, as the dashed and continuous curves, the effect of placing the bubbles randomly within a 25 kpc or 100 kpc area, respectively.

similar core X-ray luminosities. The most powerful radio sources can have monochromatic radio power of $\nu \cdot S_\nu \sim 10^{41.5}$ erg s^{-1} (where S_ν is the radio luminosity density at $\nu = 1.4$ GHz); around 10% of clusters host such powerful sources. Eilek's (2003) analysis of central cluster radio source morphologies and ages suggests that the ratio of the total jet power ($P_{\rm jet}$) to $\nu \cdot S_\nu$ is more than 10^4. Under minimal assumptions, this can be confirmed in the case of M87, where the most detailed radio observations are available (Owen et al., 2000); such powerful jets are also inferred for classical FR II sources (Willott et al., 1999).

This power is close to the Eddington limit for accretion by a 10^9 $M_\odot$ black hole. Thus, the total jet power of the most powerful radio galaxies is comparable to the energy required to counter-balance cooling in my simulated cluster. The frequency with which these sources occur in clusters also seems consistent with the duty cycle that I have assumed in my simulations.

It is worthwhile to emphasise that the energy injection can vary significantly (by a $\sim$ 50%) while still reducing the mass deposition rate to an acceptable level after 1.5 Gyr. These variations in the injected energy lead to an overall cooling or heating of the cluster: however subsequent mergers with massive substructures will disrupt the evolution presented in my simulations (which assume a static gravitational potential). The shocks and turbulence generated in the mergers will lead to a re-distribution of energy and entropy that may erase the differences that have built up during the quiescent phase that I simulate (Gómez et al., 2002; Burns et al., 2004). My model does not contain any mechanism to regulate the energy injection – this would require detailed understanding of the physics powering the central source from the surrounding gas reservoir. However, as Binney and Tabor (1995) (see also Binney, 2004) have argued, it seems natural that the cooling rate and the energy injected by the central energy source should be linked. In this case, a balance between heating and cooling is naturally expected when these quantities are integrated over sufficiently long timescales: if insufficient energy is injected into the ICM, the net cooling leads to a high mass deposition rate, and thus an increase in the fuelling of the central engine. Nevertheless, the delays inherent in this feedback loop lead to a short-term imbalance, and thus to the out bursts that I model here.

3.4.2. Global X-ray properties

Although the injected energy is able to counterbalance the cooling flow in the cluster, the effect on the cluster's overall structure is modest, as I showed in Figure 3.2. In the absence of any heating, cooling causes the central entropy and temperature to drop precipitously.

In simulation S1.0, the AGN energy is sufficient to reduce this drop to only $\sim$ 50 per cent. Thus, by preventing runaway cooling, my energy injection mechanism reduces the change in the initial entropy profile of the cluster. The result is that the integrated luminosity and average temperature of the cluster evolve little. Doubling the amount of energy injected (S2.0) results in a small increase in the central temperature ($\sim$ 30%), and a 20% reduction in luminosity.

I simulated an initially isothermal cluster, with temperature observationally consistent with its mass (Reiprich and Böhringer, 2002), in order to see how it would evolve in the luminosity-temperature plane. The experiment raises two issues that are not fully resolved by my model: (1) The cluster is too luminous, for its emission weighted temperature, compared with observed clusters (e.g., Markevitch, 1998), even when uncorrected for cooling flows (McCarthy et al., 2004). (2) The temperature profiles of the simulated clusters do not show the temperature decrement seen in the central regions of real clusters (e.g. Allen et al., 2001).

In order to lower the luminosity of the simulated cluster enough to agree with observations, I can either adjust the gravitational potential (by assuming a lower halo concentration) or increase the entropy of the central gas. In order to get good agreement with the average luminosities of clusters of this temperature, I would need to adopt $c \sim 3$, well below the concentrations inferred from dark matter simulations (Eke et al., 2001; Wechsler et al., 2002). On the other hand, the central entropy could be raised to $\gtrsim 160$ keV cm^2 (Babul et al., 2002; Voit et al., 2002), which is a factor of five greater than the initial entropy at 20 kpc. Even if enough energy were injected to achieve this (for example by running simulation S2.0 for a much longer time), this would make the core temperature even hotter, exacerbating the disagreement with observed temperature profiles. Thus, I conclude that the initial cluster gas distribution must have been different from the isothermal model I have assumed. Since the high cooling rate of the unperturbed cluster is a direct consequence of its low central entropy (leading to high X-ray luminosity), my energy injection

mechanism has removed the symptoms of the over-cooling problem, but not addressed its cause.

Global preheating of the intra-galactic medium prior to cluster formation should lead to cluster gas distributions with integrated luminosities and temperatures that are consistent with observations (Kaiser, 1991; Navarro et al., 1995; Ponman et al., 1999; Balogh et al., 1999; Muanwong et al., 2002; Borgani et al., 2002). An alternative (but related) possibility is that heating events similar to the ones I have simulated here occur in the cluster progenitors and/or surrounding filamentary structure, at earlier times. Raising the entropy in these lower mass structures may result in a higher entropy for the final cluster, as the lower density results in an increase in the accretion shock strength (Voit et al., 2003; Voit and Ponman, 2003). In either scenario, subsequent cooling would naturally lead to a temperature decrement in qualitative agreement with observations. Calculations by McCarthy et al. (2004) suggest that although the cluster would then increase in luminosity, it would decrease in emission weighted temperature so as to remain in acceptable agreement with the observed L-T relation.

3.4.3. Entropy transfer and generation

In Figure 3.1, I have shown that an energy injection rate of 6.3×10^{44} erg s^{-1} is sufficient for the total energy of the cluster to remain constant and for the density and temperature distribution of the ICM to reach a quasi-equilibrium profile in which the radiative heat losses are balanced by the energy deposited in the ICM by the hot bubbles. The slow evolution of the density profile implies that the energy injected is somehow shared with the cooling ICM and does not remain trapped in the hot material. In particular, the energy must be efficiently shared with the dense material near the centre of the cluster, where the cooling time is short, leading to an overall increase in entropy in these regions.

There are three mechanisms by which the energy can be shared:

(i) $P\mathbf{d}V$ **work** – as energy is injected into the bubble, it expands doing work on the sur-

rounding material. I calculate the PdV work by calculating the volume of the bubble at a series of time-steps. 40% of the injected energy is used in this way. The process is largely adiabatic, however, and has little effect on the entropy distribution of the surrounding ICM.

(ii) **shocks** – If the motion of the bubble were supersonic, shocks would dissipate energy in the surrounding ICM leading to an overall increase in the entropy of the system. In practice, however, the bubble does not reach supersonic speeds: as the speed of the bubble increases the energy dissipation in the turbulent wake grows. In my simulations, the maximum Mach number in the flow is 0.8. Thus I do not expect shocks to be a major source of entropy in the surrounding ICM.

(iii) **up-lift and turbulence** – The rising bubble generates a complex flow pattern, with material from the central regions of the cluster being up-lifted in the wake behind the bubble. Vortices in this flow dissipate energy and mix the bubble material, low entropy material from the centre with the surrounding ICM. These irreversible processes generate an increase in the entropy of the cluster.

Even though I assume that the viscosity of the ICM is small, it is certainly not negligible, and viscous dissipation on molecular scales can still be a major source of heating in turbulent regions of the flow. Clearly, my simulations are far from resolving the relevant molecular diffusion scales. However, in turbulent flows, the non-linear couplings cause a cascade of energy from large to small scale motions (Kolmogorov, 1941). The viscosity determines the smallest scales in the flow, while the rate at which energy is dissipated is determined by the largest scale eddies: these control how quickly energy is fed into the turbulent cascade (Tennekes and Lumley, 1972). In simulations, the development of the turbulent cascade is limited by the numerical mixing of fluid elements within a grid cell, rather than by molecular viscosity. However, the large scale properties of the flow are not expected to be dependent on the smallest scales that are resolved. I checked this by comparing the evolution of a single bubble in two simulations: namely the fiducial resolution and another with almost an order of magnitude greater linear resolution. Plotting the two

simulations to the same base resolution I see that the gross structure of the flow is the same. Thus, because of the weak dependence of the large scale structure of the flow on the numerical resolution, it is plausible that the energy dissipation and mixing rates in the regions around the rising bubble will converge over the range of resolution that I can test in Appendix B.

In addition to being a source of dissipative energy, the turbulence in the flow provides a means of increasing the entropy of the surrounding ICM through mixing. The mixing may occur between the hot material within the bubble and the surroundings, or between the surrounding ICM and the low entropy material drawn out of the cluster centre by the up-draft created by the rising bubble. In both cases, the mixing is irreversible and leads to an overall increase in the entropy of the system. In the first case, the mixing transfers energy from the bubble to the surrounding ICM (in addition to the $P\mathrm{d}V$ work discussed above). The second case does not directly transfer energy from the bubble, but by driving an entropy increase in the lowest entropy material (the material with the shortest cooling time), it prevents this material from cooling out of the flow.

3.5. Conclusions

I have presented 3-dimensional gas-dynamical simulations using the FLASH adaptive–mesh refinement code of a cooling flow in an isothermal, X-ray luminous cluster of galaxies, with periodic injections of thermal energy. The initial cluster has mass of 3×10^{14} $\mathrm{M}_{\odot}$ and a temperature of 3.1 keV, consistent with the observed mass–temperature relation (Reiprich and Böhringer, 2002). The injected energy is manifested as hot bubbles of buoyant gas that rise out of the cluster core, convectively mixing the cooling material. I treat these injection events as sporadic outbursts with a typical duty cycle of 100 Myr. The parameters of these injection events match the observed luminosity and frequency with which the most powerful radio galaxies are seen in clusters.

In the absence of any energy injection, mass rapidly flows to the cluster centre, at a rate

that exceeds observational limits by at least an order of magnitude. Based on simulations with a variety of parameters related to the energy injection rate, run for between 1.5 Gyr and 5 Gyr, I draw the following conclusions:

1. For a time averaged energy injection rate of 6×10^{44} erg s^{-1}, the mass inflow rate is less than 30 $M_{\odot}$ yr^{-1}, compatible with available observational limits. With this amount of heat input, the total energy of the cluster remains approximately constant over 5 Gyr.

2. The evolution of the total cluster energy depends primarily on the total amount of energy injected and the spatial distribution of bubbles, but is only weakly sensitive to the duty cycle of heating events or to whether the bubbles are produced singly or in pairs.

3. The bubble activity generates concentric sound waves that are clearly evident in unsharp–masked projections of cluster emissivity.

4. When the injected energy just balances cooling, the entropy and temperature profile of the cluster remain approximately unchanged from their initial configurations.

In summary, periodic energetic events of the kind I have simulated can reduce the mass flow rate and accumulation of cold gas in massive clusters to within observational limits. However, this mechanism operating on a fully formed cluster does not result in a final luminosity consistent with observations. It is likely that the structure of the progenitors from which the cluster formed was affected by heating events prior to the assembly of the final cluster.

3.6. On turbulence, viscosity and metallicity gradient of clusters

After the publication of the paper Dalla Vecchia et al. (2004), I received useful comments on and interesting questions about my results. Here, I would like to offer replies to some

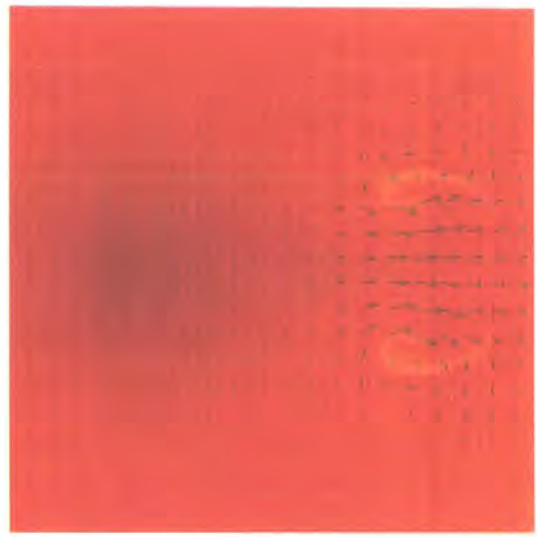

FIGURE 3.7.: Density plot on the $x - y$ plane through the centre of the cluster. Over-plotted is the $x - y$ velocity field of the gas. Turbulent motion is present at the edge of the *mushroom* head. In the wake, the flow is clearly laminar.

of these issues. I will briefly discuss turbulence, viscosity and metallicty gradients.

The heating mechanism I described in this chapter involves turbulent gas mixing that changes the entropy distribution of the gas. This happens when bubbles rise buoyantly in the intracluster medium and create vorteces at their edge. One objection to this picture is that the Hα filaments observed in the wake of rising bubbles are strikingly straight (Fabian et al., 2003b). This point is valid and suggests that the motion of the gas should be laminar, with no turbulence present. This is not inconsistent whit what I observe in my simulations. The bubble motion tends to create a vacuum in its wake that lifts the gas behind. The motion of this gas is laminar in the wake—at least to the resolution scale of the simulation—as shown by the velocity field plotted in Figure 3.7. What I have argued in this chapter is that the turbulent mixing takes place at the edge of the rising bubbles. The up-lifted material at the edge of the wake is entrained in vorteces and mixed with material external to the wake. This is clearly visible in Figure 3.7, at the edge of the *mushroom head*.

The laminarity of the flow has been also interpreted as requiring high viscosity for the

gas. In the light of what I have just discussed, viscosity is not required in order to abtain a laminar flow in the wake of bubbles. However, in my simulations, the bubbles flatten and disrupt quickly, while round depressions in the X-ray surface brightness are observed in *Chandra* images at large radii (Sanders et al., 2005). The numerical simulations by Reynolds et al. (2002) suggest that shear viscosity in the cluster gas leads to bubble morphologies and flow patterns consistent with observations. However, the importance of shear viscosity could be signifcantly reduced if the plasma is magnetised. The measure of cluster magnetic field is still inprecise, and different methods can lead to different results (Carilli and Taylor, 2002). However, a magnetic field of strength of the order of few μG is not unlikely in the vicinity of buoyant bubbles in some cluster (e.g. in *Virgo* cluster Owen et al., 1990). As shown by Kaiser et al. (2005), if the component of the magnetic field[1] parallel to the bubble surface is of the order of 0.1 μG, Rayleigh-Taylor instabilities[2] can be effectively suppressed. The magnetic field acts as a surface tension at the interface between the bubble gas and the ICM. Moreover, Kaiser et al. (2005) point out that the difference between the morphology of the bubbles in the *Perseus* cluster and in the *Virgo* cluster could be due to the strength of their gravitational potential. The stronger gravitational potential in *Virgo* could be more effective at mixing the bubble gas with the surrounding ICM through Rayleigh-Taylor instabilities. This could explain why no X-ray surface brightness depressions are seen in the proximity of the central radio galaxy.

Another objection is that the turbulent mixing would reduce, and eventually flatten, any metallicity gradient in the cluster. This is a very interesting point, and I ran a simulation to investigate it. I added an extra "dye fluid" to my model cluster. The fluid does not contribute to the hydrodynamics of the gas, but just advects with the density. One can thus follow the motion of the gas and see where the fluid is deposited. In order to do this,

[1]The magnetic field of the ICM, not the bubble magnetic field (see Kaiser et al., 2005, for more details).

[2]Rayleigh-Taylor instabilities happen at the boundary between two fluids of different densities when a gravitational field is present. Where the two fluids are moving with repsect to each other, Kelvin-Helmotz instabilities develop. These instabilities are suppressed by shear viscosity, and are not important if the ratio of the fluid densities is large, even in the case of inviscid fluids.

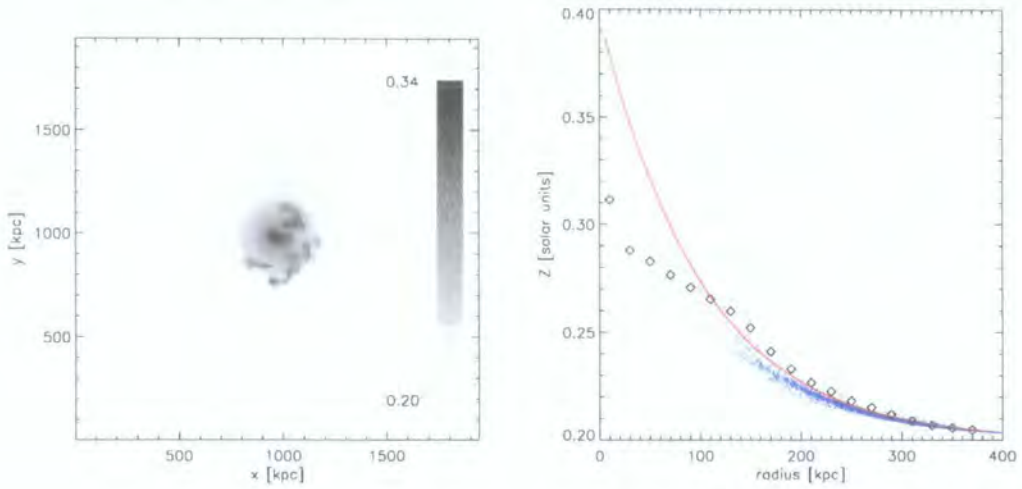

FIGURE 3.8.: Left panel: distribution of the advected fluid on a plane through the centre of the cluster. Material is up-lifted by bubbles in a region of radius $\sim$ 200 kpc. Right panel: shell averaged profiles of the advected fluid. The red line is the initial profile; the diamonds are the profile after 1.5 Gyr. The blue points are a sample of density values in single cells.

initially, I distributed the fluid with a gradient that increases toward the cluster centre—in Figure 3.8, the initial gradient is represented by the red line. Then, I ran the simulation for 1.5 Gyr, injecting the same bubbles as in simulation S2.0. The left panel of Figure 3.8 shows the distribution of the fluid on a plane through the centre of the cluster, at the end on the simulation. The up-lifted gas is clearly visible in the wake of the bubbles and in the bubbles remnants. Most of this low entropy material will eventually fall back towards the centre, except that which is mixed with higher entropy gas. In the right panel I show the final gradient (diamonds) and a sample of fluid density in single cells (blue dots). While some cells have a high value of density, the average gradient is very similar to the original one beyond 150 kpc. Inside 100 kpc it is shallower, but considering the timescale of the simulation and that there are no other processes contributing to maintaining the steepness of the gradient (in the case of a real cluster, supernova explosions could keep injecting metals at the centre), I may conclude that turbulent mixing will not destroy a metallicty

gradient. I will investigate this issue in more detail in the near future. A better model should take into account a simple star formation and metal enreachment recipe. This could be done by adding an analitic prescription for star formation to the model cluster setup used in the simulations of this chapter, or simulating star formation with the cosmological setup of Chapter 6.

I think I used to have a voice

Now I never make a sound

Nine Inch Nails

Detection of sound waves

Part of the work presented in this chapter is a result of using the X-MAS *code in collaboration with Pasquale Mazzotta and Elena Rasia.* X-MAS *is not yet publicly available. Therefore, the X-ray observations presented here were generated by Elena Rasia, who is the main developer of the code.*

4.1. Introduction

The unprecedented resolution of the *Chandra* X-ray space observatory is providing us with very detailed images of the intracluster medium. These images unveil a wealth of structures which are identified as cold fronts, cavities, bubbles, sound waves and weak shocks (e.g. McNamara et al., 2000, 2001; Mazzotta et al., 2002; Fabian et al., 2003a). Sound waves have been studied to prove the dynamics of bubble expansion and the recurrence of AGN activity. The results show that the bubbles expand subsonically or at most transonically producing weak shocks in the surrounding gas. Sound waves propagate through the intracluster medium, possibly dissipating their energy by viscous processes (Ruszkowski et al., 2004a).

Periodic AGN activity is thought to be the dominant mechanism responsible for producing several generations of sound waves. The variability of mass ejection from stellar mass black holes has been observationally studied. It is thought that the periodic activity of these black holes depends on instabilities on the accretion disk due to hydrogen ionisation processes. Instabilities affect the disk thickness, and hence the transport efficiency, reducing the power of black hole jets. There is observational support for the theoretical speculation that this argument could not be applied to massive black holes (Done and Gierlinski, 2005). Moreover, the accretion disk in AGN is much cooler than in stellar mass black holes and is not observable in X-ray bands. An indirect way of testing the accretion theory of massive black holes is therefore required in order to understand the role of variable accretion. One can infer the time elapsed between two active phases of the AGN, given the sound speed in the intracluster medium and the spatial distance between the waves (Fabian et al., 2003a). The measure of the duty cycle of AGN could give more insight on the timescale over which these processes happen.

In this chapter I investigate the detectability of sound waves by the *Chandra* satellite by carrying out simulations adopting the properties (temperature, density, etc.) of the well-studied Perseus cluster. Then, I use a new definition of projected temperature (Mazzotta

et al., 2004) to estimate whether sound waves could be observed. I perform a synthetic *Chandra* observation of the cluster with the code X-MAS (Gardini et al., 2004) and, finally, I calculate the duty cycle of the AGN from projected temperature images. This work is mainly qualitative—I am planning to apply this methodology to cosmological simulations in the near future, performing a proper analysis of the X-ray images.

4.2. Simulation of Perseus cluster

I performed a simulation of a model cluster with a set-up similar to the one presented in Chapter 3. The main difference is that the cluster is not isothermal, but resembles the temperature and gas density profiles of the Perseus cluster (Ettori et al., 1998; Schmidt et al., 2002). Density, pressure, and temperature profiles are shown in Figure 4.1. In order to model the Perseus cluster, I first impose an NFW profile for the dark matter with virial mass $M_{200} = 6 \times 10^{14}$ M$_\odot$ h^{-1} and concentration $c = 5$ (Ettori et al., 1998). The resulting virial radius is $r_{200} = 1.37$ Mpc h^{-1}. The gas is initially settling in the dark matter potential well in hydrostatic equilibrium. To represent the temperature drop in the core of the *Perseus* cluster (Schmidt et al., 2002), the core density must be higher than in the isothermal case. I therefore modified the isothermal ($T = 6.3$ keV) gas profile making it steeper in the inner 100 kpc. The pressure profile is derived by imposing the hydrostatic equilibrium condition, and the temperature profile then recalculated using the equation of state of the gas. The initial luminosity of the cluster is $\sim 3.7 \times 10^{45}$ erg s^{-1}. This is in agreement with data from Markevitch (1998). Adopting observed density and temperature profiles prevent the problem of the high luminosity of the model cluster discussed in 3.2.4.

The size of the box has been reduced along x and y in order to speed up the calculation. I use a rectangular volume with dimension of $972 \times 972 \times 1944$ kpc^3 (projected along z for the analysis). The code ran with a maximum refinement of 8 levels, achieving a resolution of 0.95 kpc. The simulation was performed with cooling switched on, and ran for a time interval of 100 Myr.

Pairs of bubbles of energy were injected every 10^7 yr; the duty cycle is in agreement with the one inferred by Fabian et al. (2003a) from their observations of *Perseus*. I did not try to match the position of the bubbles in Perseus, but I imposed a distance from the centre similar to what is observed. The energy injected in each bubble is 2.5×10^{59} erg, and the duration of this injection is 3×10^5 yr.

The refinement criterion was adjusted in order to follow the propagation of the sound waves in the intracluster medium. To keep the sound waves refined to the maximum resolution, I proceeded in the following way. I first performed a simulation to measure their average propagation speed, and then I used this average value to set up the refinement to the maximum level on spherical shells centred on the bubbles and expanding at this velocity. In this way, the simulation can resolve the temperature fluctuations in more detail throughout their evolution, even when they are at considerably large distances from the centre of the cluster.

4.3. X-ray MAp Simulator (X-MAS)

X-MAS has been developed to provide X-ray images from hydrodynamical N-body simulations. The software package simulates ACIS-S3 *Chandra* observations, and the simulated data can be analysed in the same way and using the same tools as real observations. One of the main features of X-MAS is the ability to generate event files following the same standards used for real observations. The produced synthetic X-ray images have the same high spatial resolution provided by the *Chandra* satellite.

The package can be divided into two main units. The first unit is independent of the specific characteristics of the simulated X-ray telescope. For each energy channel considered, it generates a map of the differential flux obtained by projecting the specific emission of each particle along the line-of-sight. That means that, for each line-of-sight, it calculates and stores the corresponding projected mass-weighted spectrum. The second unit takes each spectrum and simulates the data relevant to an observation with a specific X-ray tele-

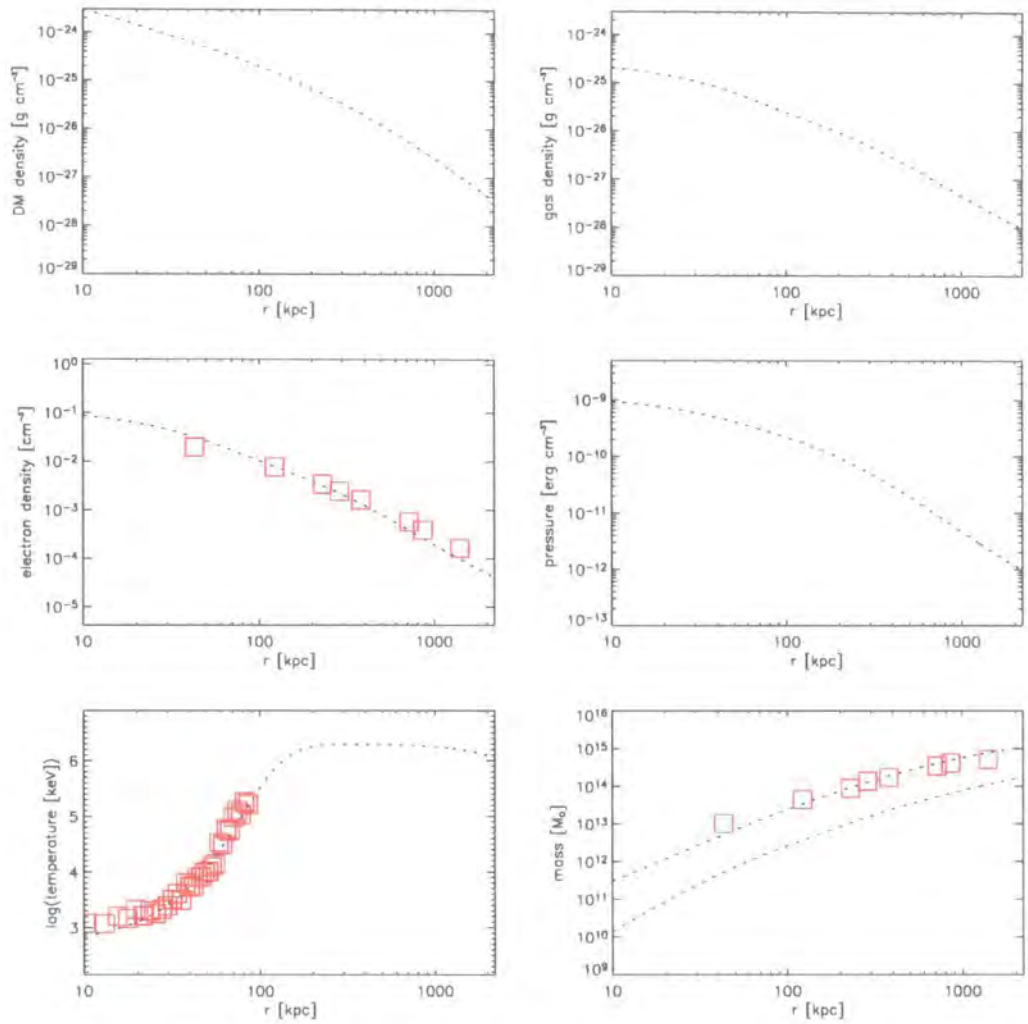

FIGURE 4.1.: The initial setup of the simulation. From the top, dark matter, gas, electron number, pressure and temperature density and gravitational mass profiles. Over-plotted red squares are the temperature profiles from Schmidt et al. (2002) and electron number density and gravitational mass from Ettori et al. (1998). The values have been normalised to the adopted Hubble parameter, $h = 0.7$.

scope and detector for a defined amount of time. Many artifacts introduced by instrumental response, sky background and instrumental noise are taken into account to obtain images as close as possible to those produced in ACIS-S3 *Chandra* observations.

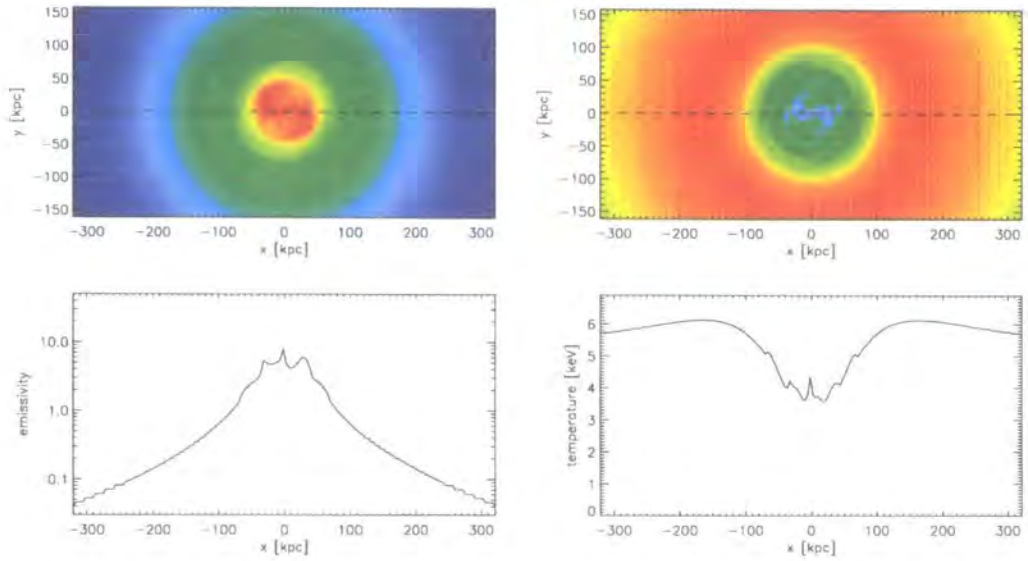

FIGURE 4.2.: Top panels show the projected emissivity (left) and spectroscopic-like temperature (right). Bottom panels show the value of the emissivity (left) and spectroscopic-like temperature (right) along the dashed lines plotted above. Sound waves are visible in the temperature plot as small bumps in the temperature profile.

4.4. Detection of sound waves

Using these techniques, on can ask: could the sound waves observed in our previous simulations (Dalla Vecchia et al., 2004, see Chapter 3) be detected by the *Chandra* observatory? As a first test, I investigated the projected emissivity and temperature resulting from our simulation. The temperature was projected using the weighting method described in Mazzotta et al. (2004). The so-called *spectroscopic-like temperature* gives a better description of observed clusters with temperature above 2 keV. The emissivity-weighted temperature, commonly used in the analysis of simulations, tends to be higher then the temperature inferred with X-ray observations—the emissivity-weighted temperature does not reflect the actual spectral properties of the observed source. This has important implications for the study of thermal structures in clusters, especially when strong temperature gradients and shocks are present. An application of this method can be found in Rasia et al. (2005).

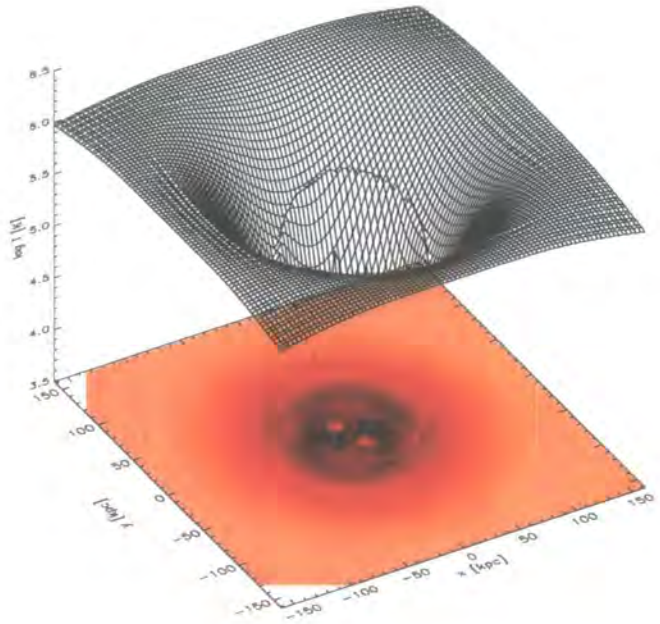

FIGURE 4.3.: Three dimensional representation of the projected spectroscopic-like temperature. The surface height is the value of the temperature. The sound wave climbing the temperature gradient is clearly visible.

In Figure 4.2 I show results for a selected output. Emissivity and spectroscopic-like temperature maps are shown in the top-left and top-right panels, respectively. The emissivity plot shows depressions in the surface brightness associated with the bubble cavities. Sound waves that are detaching from the edge of the last generation bubbles are still found close to the bright, over-pressured ring around the bubbles. For this reason they are not clearly visible in the image. A previous generation of bubbles has generated the sound waves visible as small jumps in the emissivity profile, at about 70 kpc from the centre (see bottom-left panel). The temperature profile clearly shows the sound waves as narrow bumps in the gradient (at the same distance of 70 kpc; see bottom-right panel). Figure 4.3 shows a three dimensional realisation of the projected temperature. The height of the surface is the value of the temperature for the given pixel of the image below. A sound wave, climbing the

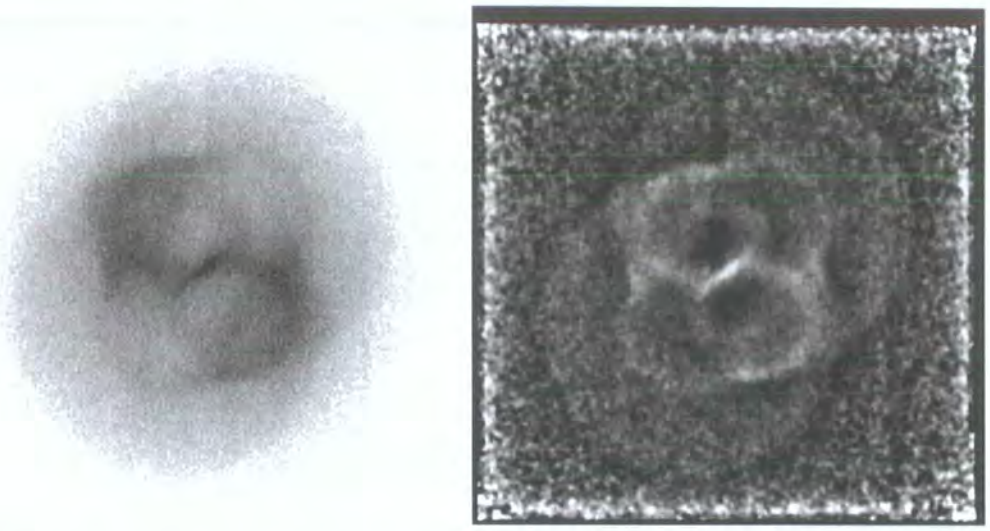

FIGURE 4.4.: The observation of the cluster made with X-MAS (left panel) as saved in the event file. The contrast has been enhanced and the colour table inverted to better show the details. Two recent bubbles are visible, aligned from the top-left to the bottom-right. Two older bubbles are still visible aligned from the top-right to the bottom-left. In the same direction, at larger radii, are two arcs concentric about the position of the older bubbles. They are at the position of the sound waves. In the right panel, I show the unsharped-mask version of the previous image. The same structures are recognisable.

temperature gradient, is clearly visible.

Figure 4.4 shows the synthetic *Chandra* observation. The "raw" image, as observed by the instrument, is shown in the left panel. The image was obtained with X-MAS considering the cluster to be at the redshift of Perseus ($z = 0.0183$), and exposing for 200 ks. Some major features are visible with the naked eye: two recent bubbles surrounded by brighter rings and aligned from the top-left to the bottom-right; two previous generation bubbles aligned from the top-right to the bottom-left; two arcs concentric about the position of the older bubbles marking a jump in the surface brightness. The image in the right panel is obtained by applying an unsharp-mask filter. Two smoothed version of the original image are created using different smoothing scales. The larger scale smoothing provides an image without small details; the image is then subtracted from the other. The bright edges of the

T	c_s	d_0	d_1	Δt
keV	km s^{-1}	kpc	kpc	yr
WITH RESPECT TO THE CLUSTER CENTRE				
3.5	979	59.9	45.0	1.63×10^7
5.0	1170	59.9	45.0	1.35×10^7
6.3	1306	59.9	45.0	1.21×10^7
WITH RESPECT TO THE BUBBLE CENTRE				
3.5	979	70.2	61.6	0.94×10^7
5.0	1170	70.2	61.6	0.78×10^7
6.3	1306	70.2	61.6	0.70×10^7

TABLE 4.1.: Results of the calculation of the duty cycle Δt for different measurements of the sound wave distance.

bubbles and the sound waves associated with the first AGN event are clearly enhanced by the filtering technique.

4.5. Measure of the duty cycle

Fabian et al. (2003a) give an estimate of the duty cycle of the central AGN of the *Perseus* cluster. Assuming that the ripples they observe are sound waves generated by intermittent activity of the central AGN and are moving at constant speed, they compute the duty cycle assuming a canonical value for the distance between sound waves. They measure a distance of about 11 kpc and impose a constant speed of 1170 km s^{-1} (this is the sound speed for a gas at the temperature of 5 keV). The duty cycle is then about 9.6×10^6 yr. In this section, I infer the duty cycle knowing *a priori* the initial position of the bubbles.

Initially, I do not take into account the position of the bubbles, but I refer all the measures to the cluster centre. Even knowing the time evolution of the sound waves from the simulation outputs, it is not a trivial task to distinguish them. Because of the different spatial

orientation of bubble pairs, the second generation of sound waves could partially overlap with the previous one. Looking in the direction of the bubble pair axis, sound waves are at the maximum distance from the cluster centre; looking in the orthogonal direction, they appear much closer to the centre. In order to avoid introducing more uncertainty, I measured positions along the pairs axes. Positions are measured by picking a point on a projected temperature image. The simulation output has been chosen making sure that the second generation of sound waves are well detached from the bubbles' edge.

The sound speed in a gas of temperature T is defined as

$$c_{\rm s} = \left(\gamma \frac{k_{\rm B} T}{\mu m_{\rm p}}\right)^{\frac{1}{2}} \tag{4.1}$$

where γ is the adiabatic index, $k_{\rm B}$ is the Boltzmann's constant, μ is the mean molecular weight in units of $m_{\rm p}$ ($\mu = 0.59$ for a fully ionised plasma), and $m_{\rm p}$ is the mass of a proton. The sound speed does not depend strongly on the temperature, and assuming a constant value should not introduce too large errors in our qualitative approach. I assume constant sound speeds for temperatures of 5 keV (see Fabian et al., 2003a) and 6.3 keV; these are 1170 km s^{-1} and 1306 km s^{-1}, respectively. The values of the duty cycle are then 1.35 and 1.21 $\times$ 10^7 yr (see Table 4.1), 21% to 63% higher than the duty cycle set in the simulation.

To see whether the initial position of the bubbles is relevant in measuring the distance between sound waves, I repeated the calculations as follow. I picked two new positions for the sound waves, and subtracted from them the initial position of the bubbles. With these new values and the same sound speeds, I recomputed the duty cycles. The result are values that are 6% to 30% lower than the duty cycle set in the simulation (see Table 4.1). For a mean sound speed of 979 km s^{-1}, I am able to obtain the simulation duty cycle. Adopting this value of the sound speed, the duty cycle inferred by Fabian et al. (2003a) becomes 1.2 $\times$ 10^7 yr. The large discrepancy in the computed values is basically due to the steep temperature gradient in the inner 100 kpc of the cluster, and a better, quantitative analysis should take into account the change in sound speed with radius. On the other hand, this

analysis demonstrates that knowing the initial position of the bubble that generated the sound wave can improve the accuracy of the measurements. Of course, this is impossible for observed clusters. However, the results confirm that this is a generally good method to estimate AGN duty cycles, and the estimates are accurate within a factor of two.

4.6. Conclusions

I have shown that sound waves can be detected in *Chandra* X-ray images. In order to do this, I made use of two new methods of simulation analysis. I computed realistic temperature maps using the definition of spectroscopic-like temperature. I then observed the cluster producing *Chandra*-like X-ray images with the code X-MAS. In both cases, sound waves are visible. I also inferred the duty cycle of AGN events from plots of the projected spectroscopic-like temperature. I conclude that, even in this highly idealised set-up, the determination of the duty cycle is accurate to within a factor of two. For more realistic data, at least the order of magnitude of the duty cycle should be easily accessible.

AGN bubbles in low mass halos

In this chapter I investigate whether the model of energy injection for quenching cooling-flows in clusters of galaxies presented in Chapter 3 is applicable to low mass halos, namely to halos corresponding to groups of galaxies and to individual galaxies. The presence of a mass limit below of which the energy injection is no longer effective is discussed in the light of recent results on semianalytic simulations of the luminosity function of galaxies.

5.1. Introduction

Semianalytic models (White and Frenk, 1991; Kauffmann et al., 1993; Cole et al., 1994) have been extensively used to understand the formation of galaxies in the hierarchical scenario of structure formation. The physics of gas cooling and merging included in these models is modelled with reasonable accuracy (Helly et al., 2003). Recently, high precision measurements of the luminosity function of galaxies have been obtained from large redshift survey such as the two-degree field galaxy redshift survey (2dFGRS) and the two micron all-sky survey (2MASS). These accurate measurements pose a challenge to the realism with which various physical processes are modelled with semianalytic codes. The processes that shape the luminosity function remain poorly understood; basic semianalytic models tend to overestimate the number of galaxies at both the faint and bright end.

In their model, Cole et al. (2000) found good agreement with the observed luminosity function when they adopted a cosmic baryon fraction $\Omega_b = 0.02$ and a model of the halo gas distribution containing a core radius that varied as a function of the amount of gas that had cooled in previous generations of halos. But, adopting recent determinations of the baryon fraction, $\Omega_b = 0.04$, the cooling radius is typically much larger than the core radius, whose evolution then has little effect on the cooling rate.

In their recent work, Benson et al. (2003) tried to match the observed luminosity function using different feedback recipes. They considered three processes whereby supernovae and stellar wind energy could affect the forming galaxy: the reheating of cold disk gas to the halo temperature; the expansion of the hot, diffuse halo gas; and the expulsion of cold disk gas from the halo. They show that re-heating of disk gas can flatten the faint end of the luminosity function. In Figure 5.1, I reproduce two plots from their paper. The plot in the left panel shows the result of disk gas reheating in star-forming galaxies. The faint end is well recovered by one of their models, but at the bright end the number of bright galaxies has increased compared with simple models including only cooling of the gas. The plot in the right panel shows the model of energy injection by supernova and stellar winds. The

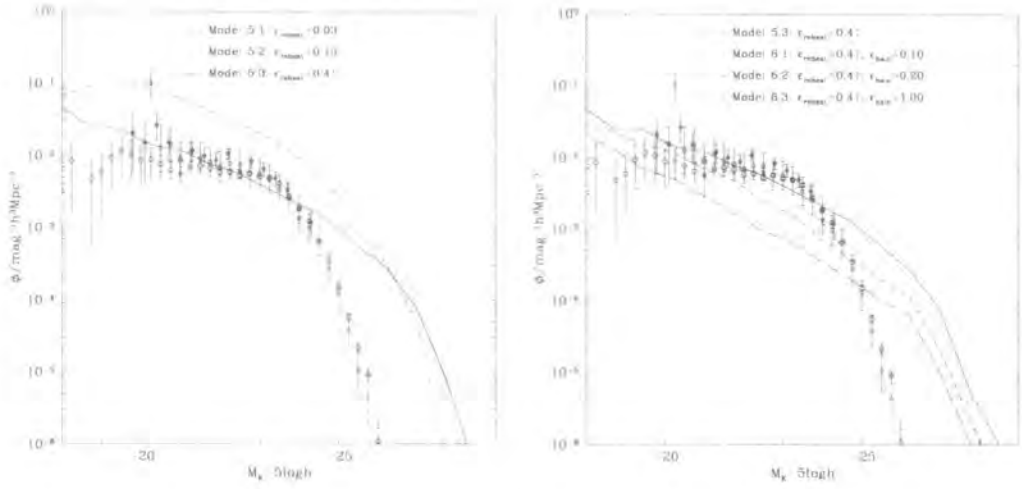

FIGURE 5.1.: **Left panel**: Galaxy luminosity function resulting from disk gas reheating in star-forming galaxies. **Right panel**: Galaxy luminosity function resulting from supernova and stellar wind energy injection. Both plots are from Benson et al. (2003).

model that best recovers the faint end of the luminosity function is the one with the largest number of bright objects.

The complete ejection of cold gas from galaxies by strong stellar winds could also prevent the formation of low mass objects. However, the expulsion of gas increases the mean density of the ICM, enhancing the cooling of the ICM in clusters of galaxies resulting in an increase in the number of luminous massive galaxies. Introducing thermal conduction can solve this problem; thermal conduction is much more effective in massive halos because of its strong dependency on the temperature (see Section 1.4.1). However, Benson et al. (2003) required values of the thermal conduction well over the Spitzer rate in order to obtain the sharp cut-off at the bright end. Thus, thermal conduction is not the favoured mechanism to limit the luminosity of massive galaxies, and it could introduce other undesirable effects (e.g. Dolag et al., 2004, see also Section 1.4.1 for more details).

Energy injection from AGN has been shown to quench cooling-flows in the centre of massive halos. In Chapter 3, I discussed the problem of cooling-flows in a cluster of mass 3×10^{14} $M_{\odot}$, and showed that when the rate of energy injection equals the luminosity

	$\log M$	c	T	$t_{\rm sim}$	r_{200}
	$M_\odot$		K	Myr	kpc
S120	12.0	8.7453	7.4550×10^5	219	207.4
S125	12.5	7.8586	1.6061×10^6	312	304.6
S130	13.0	6.9719	3.4603×10^6	469	453.7
S135	13.5	6.0853	7.4550×10^6	656	648.2

TABLE 5.1.: Parameters of the four simulation: cluster mass and concentration, initial isothermal temperature, simulation time and virial radius (size of the box).

of the cluster the cooling-flow is completely suppressed. However, the AGN mechanical feedback of this model requires that the bubbles reach the cooling radius and mix with the surrounding gas. This mechanism works on a timescale of a fraction of the sound crossing time, which is shorter than or comparable to the timescale for cooling. What happens when the cooling timescale becomes shorter than AGN heating timescale is not clear, and the investigation of these cooling regimes is the aim of the work presented in this chapter. The goal of the simulations presented here is to investigate if in low mass halos cooling becomes so strong that AGN energy injection becomes ineffective at heating the halo core. In this chapter, I try to find a lower mass limit for the applicability of the AGN heating model and investigate whether this model is relevant to the shape of the bright end of the luminosity function of galaxies. In the next section, I describe the setup for the simulations. I show also the results of cooling-only simulations and the effect of energy injection. In Section 5.2.4 I discuss the significance of the mass threshold.

5.2. The simulations

5.2.1. Setup

I performed simulations of energy injection in different mass clusters using the isothermal setup presented in Chapter 3. At the initial time, the baryonic matter is in hydrostatic equilibrium in the potential well given by an analytic NFW dark matter profile. The dark matter profile is defined once the mass of the cluster, M, and the concentration parameter, c, are specified. I chose the concentration as a function of the mass of the cluster according to the equation[1]

$$c = 30.02 - 1.77 \log10(M/M_{\odot})$$

The temperature of the gas, T, is derived from the M-T relation, and scales as $M^{2/3}$:

$$T = T_0(M/M_0)^{2/3}$$

where $T_0 = 6.424$ keV and $M_0 = 10^{15}$ $M_{\odot}$ are in agreement with values from Reiprich and Böhringer (2002). As in Chapter 3, the gas density distribution is

$$\rho_{\rm GAS}/\rho_0 = (1 + r/r_{\rm s})^{\eta/(r/r_{\rm s})},$$

where ρ_0 was chosen to satisfy the relation

$$M^{200}_{\rm GAS}/M^{200}_{\rm DM} = \Omega_{\rm b}/(\Omega_{\rm m} - \Omega_{\rm b}),$$

Here, $M^{200}_{\rm DM}$ ($M^{200}_{\rm GAS}$) denotes the dark matter (gas) mass contained within a sphere whose mean interior mass density is 200 times the critical density; the radius of the sphere is called r_{200}. I used $\Omega_{\rm m} = 0.3$ and $\Omega_{\rm b} = 0.04$ for the matter and baryon densities in units of $\rho_{\rm crit}$, and $h = 0.7$ for the Hubble parameter. The value of η depends on the temperature and the dark matter distribution through the relation (see Wu et al., 2000, for more details)

$$\eta = \alpha \left(\frac{k_{\rm B}T}{\mu m_{\rm p}}\right)^{-1}$$

[1]The equation is a fit to data from Eke et al. (2001) and was provided by Mike Balogh.

where $\alpha = 4\pi G\rho_s r_s^2$ (here, G is the gravitational constant, $\rho_s = \rho_{DM}(r_s)$, and $r_s = r_{200}/c$), k_B is the Boltzmann's constant, μ is the mean molecular weight in units of m_p, and m_p is the mass of the proton.

The size of the simulated volumes have been set to the value of the virial radius r_{200} of each cluster. In each simulation, the maximum resolution is therefore equal to the same fraction of the virial radius. The simulation time has been chosen to be proportional to the cooling time as follows. The cooling time depends on the temperature, T, of the gas through the relation

$$t_{cool} \sim \frac{T}{\Lambda(T)} \sim \frac{T}{T^{1/2}} = T^{1/2}$$

for Bremsstrahlung cooling. The virial radius scales as $M^{1/3}$, where M is the mass of the halo, and the temperature is proportional to M/R. That is, $T \sim M^{2/3}$ and $t_{cool} \sim M^{1/3}$. Thus, we take the simulation time to scale as $M^{1/3}$. This is not the optimal choice, however, since the cooling of the smaller halos is mainly due to line cooling rather than Bremsstrahlung, but the only disadvantage of this scaling is that these simulations will be run for a larger number of cooling times than the ones for the more massive halos. For each simulation, a metallicity $Z = Z_\odot$ was adopted. The simulation parameters are tabulated in Table 5.1.

5.2.2. *Cooling-only simulations*

Four simulations without bubble injection were performed to establish the effect of cooling on the deposition of cold mass at the centre of the halo and on the luminosity of the cluster. They will be labelled hereafter as S120, S125, S130 and S135 according to the mass of the halo. In Figure 5.2, I show the results of these simulations. In the left panel, the emissivity of the cluster as function of time is plotted. For the clusters modelled in S120 and S125, the initial luminosity is very high, and drops quickly in the first $0.2t_{sim}$ time interval. This is a consequence of the large concentration of the halo and the associated high central gas density. The central gas cools very quickly down to the minimum temperature allowed by

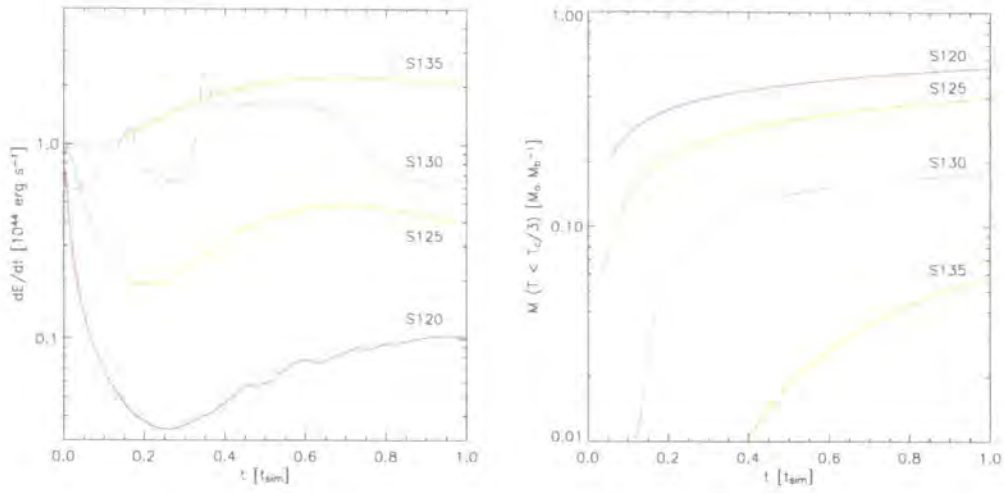

FIGURE 5.2.: **Left panel**: The emissivity of the clusters as function of the time in cooling-only simulations. **Right panel**: Fraction of mass that has cooled below one-third of the initial temperature as a function of time for simulations S120, S125, S130 and S135. Time is in units of the simulation time in both plots.

the cooling routine (10^4 K). The cooling timescale is shorter than the dynamical timescale and the surrounding gas takes a long time to collapse into the centre. This can be observed in the behaviour of the total emissivity. The inflow of mass starts to increase the energy radiated after $0.2t_{\rm sim}$. At this time, a cooling-flow is established and the inflow of hotter gas keeps the emissivity almost unchanged.

In the case of simulations S130 and S135, the initial cooling phase is much more gentle. S130 presents a quite irregular emissivity curve due to the temperature regime in which the gas finds itself. The initial halo temperature is close to the minimum of the cooling function (see Figure 1.3). Gas that is cooling slowly at the start suddenly cools faster once its temperature falls in the regime of line cooling. For simulation S135 the transition is very smooth. The cooling timescale is sufficiently long that the gas has time to balance the pressure loss with an inflow. The core increases its density and therefore its cooling rate. The emissivity slowly increases until the balance between cooling and inflow of hotter gas is reached.

	$r_{\rm b}$	$t_{\rm b}$	$t_{\rm inj}$	$E_{\rm b}$
	kpc	yr	yr	erg
B120	4	10^6	10^5	1.6×10^{56}
B125	4	10^6	10^5	6.4×10^{56}
B130	5	10^6	10^5	1.6×10^{57}
B135	5	10^6	10^5	5.0×10^{57}

TABLE 5.2.: Parameters of the energy injection: bubble radius, duty cycle, injection time and bubble energy.

In the right panel of Figure 5.2, I plot the fraction of cold gas in the cluster. The fraction is defined as the amount of gas that has cooled below one-third of the initial temperature, normalised to the baryon mass. The reader should notice that although the mass that has cooled at the end of the simulations is higher for the cluster S135, the cooling is much more effective in low mass halos. As I will discuss below, the ratio between the cooling radius and the virial radius is larger for smaller mass halos.

5.2.3. Energy injection simulations

The four simulations presented above have been re-run adding energy injection. I refer hereafter to the simulations with bubbles of energy as B120, B125, B130 and B135 according to the mass of the clusters. Energy injection follows the scheme discussed in Chapter 3. Bubbles of energy are injected at a rate that matches the luminosity of the cluster. The values of the luminosity have been taken from the cooling-only simulation and are the mean values of the energy emitted by the cluster in the cooling-flow regime. In order to avoid the large drop of emissivity observed in S120 and S125, the mean luminosities have been calculated in the time interval 0.5–$1.0t_{\rm sim}$. The values of the energy injected in each bubble are tabulated in Table 5.2.

Recent observations have placed some constraints on the value of the duty cycle[2] in cluster of galaxies (e.g. Fabian et al., 2003a), but it is still not clear whether there is any dependence on halo mass, inflow rate or any other parameter related to the AGN feedback. One could argue that the cooling timescale should dictate the duty cycle if the mass inflow rate is responsible for switching on the central AGN. That is, the mass flowing towards the central black hole could trigger accretion onto the black hole itself and the consequent emission of energy. For small mass halos, this could result in a large number of AGN events producing small bubbles of energy. The buoyant force acting on these bubbles would be quite small, and bubbles would rise very slowly through the ICM. However, if the precession angle of the central black hole is small, subsequent AGN episodes would keep inflating the same bubbles until the bubbles effectively detach from the centre of the halo. I adopted a different strategy: the duty cycle, $t_{\rm b}$, was fixed to the same value of 10^6 yr in all the simulations. In the light of the previous discussion, if this duty cycle is longer than expected for small halos, then the bubbles would be more energetic, their size would be larger and the bubbles would move faster through the ICM. That is, energy would be deposited at larger radii. One could expect that in this case it would be easier to quench the cooling-flow. I will show in the next section that even with this approach cooling-flows in the smaller halos that I simulated can not be quenched.

I show the results of the bubble simulations in Figure 5.3. The dashed lines, taken from Figure 5.2, are from the cooling-only simulations. Cooling is strongly suppressed in simulation B135, and no material as cooled below the temperature threshold (the simulation has run just for half the time). In the case of simulation B130, energy injection offsets the cooling initially (relative to simulation S130) but, at later times, the amount of cold material approaches that found in S130. In simulations B120 and B125 energy injection does not offset the initial cooling phase. In simulation B125, the energy injection keeps the amount of cold material constant at later times. A more detailed discussion of the results

[2]The time interval between two consecutive AGN active states.

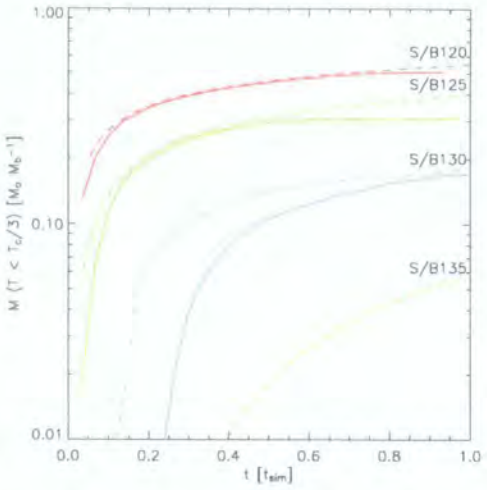

FIGURE 5.3.: Fraction of mass that has cooled below one-third of the initial temperature as a function of time for simulations with energy injection, B120, B125, B130 and B135. Dashed lines are the results of the cooling-only simulations S120, S125, S130 and S135. Time is in units of the simulation time.

is given in the next section.

5.2.4. Qualitative overview of the simulations

Plots of the temperature on a plane through the centre of the simulated clusters are shown in Figure 5.4, 5.5, 5.6 and 5.7 (for simulations B120, B125, B130 and B135, respectively). Simulations B135 has not been run to the end, but at the time of $0.5t_{\rm sim}$ the mass that has cooled below the temperature threshold is 300 times less than in simulation S135.

In Figure 5.4, snapshots from simulation S120 are shown. The cooling time is so short that a large amount of the volume has cooled down to the minimum temperature of 10^4 K (dark region in the plots). The radius of this region is larger than one-quarter of the virial radius of the halo at the end of the simulation and therefore much larger than the radius at which bubbles are injected. The bubbles remain close to the centre of the cluster possibly because of the inflow of material and the small buoyant forces, and do not affect the cooling region much (see discussion is Section 5.3). The recurrent injection of energy creates two

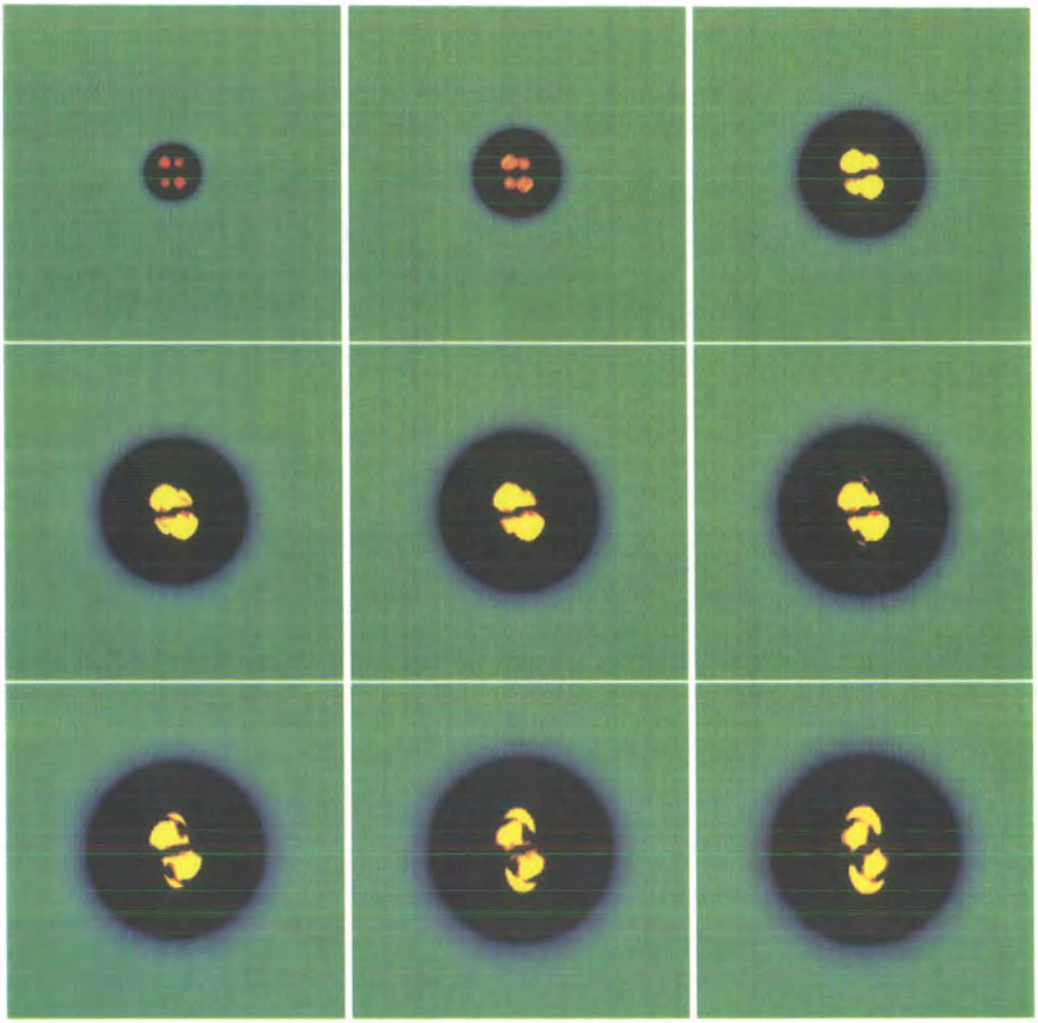

FIGURE 5.4.: Simulation B120. Plots of the temperature on a plane through the centre of the halo.

large regions of hot gas that are the result of multiple small bubbles injected close to each other. This suggests that adopting a shorter duty cycle would not affect the net result that most of the gas cools. Hereafter, I refer to these regions with the term *hyper-bubbles*. The two hyper-bubbles seen in the pictures do not detach from the centre, but seem to shrink with time. It could be that the inflow of material is increasing the pressure around the bubbles or that some of the bubble energy is leaking. A similar behaviour is observed in the central phases of simulation B125.

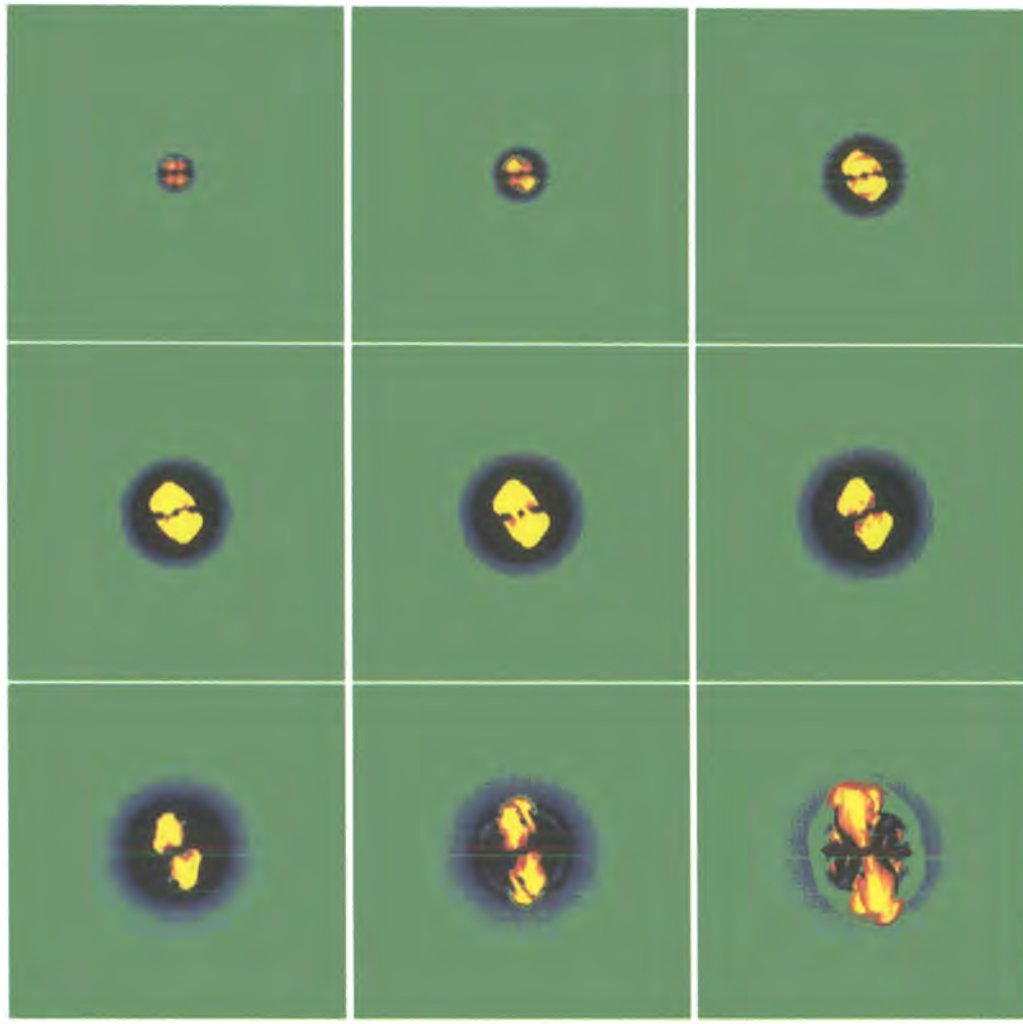

FIGURE 5.5.: Simulation B125. Plots of the temperature on a plane through the centre of the halo.

A slightly different behaviour is in simulation B125 at late times (Figure 5.5). Cooling is still strong as shown by the region that has cooled, and the injection of many bubbles close to the centre has inflated two lobes of high temperature gas. These hyper-bubbles can now detach from the centre and rise buoyantly through the ICM. The high energy material can reach the border of the cooling region, and seems able to heat some gas. The morphology of the hyper-bubbles is interesting. The long filaments of colder gas seen inside the bubbles could be interfaces created when many small bubbles came together; or it could be that the

filaments are the result of Rayleigh-Taylor instabilities at the hyper-bubble surface.

In simulation B130 (Figure 5.6), bubbles seem to rise earlier and faster. However, there is still some overlap of the regions of energy injection. The expansion of the core gas increases the density, and therefore the cooling rate, of the gas surrounding the bubbles. Interfaces of cold gas are entrained between bubbles and cold gas is able to flow to the centre of the cluster as can be seen in the fifth and sixth pictures. As shown in Figure 5.3, the amount of cold gas is not dissimilar to the amount of cold gas in the cooling-only simulation.

Simulation B135 (Figure 5.7) presents very similar morphologies to the case shown in Chapter 3. Bubbles of energy rise through the ICM to large radii and deposit their energy well beyond the cold core. Up-lifting central gas, they efficiently suppress the cooling in the centre of the halo.

5.3. Discussion

The luminosity-function of galaxies shows a change of slope at the characteristic mass of $\sim 10^{12}$ $M_{\odot}$. Semianalytic models are able to reproduce the flat profile at the faint end of the luminosity function with feedback from stars and supernovae. However, this form of feedback can not explain the sharp cut-off at the bright end. AGN energy injection is the primary candidate for quenching the cooling and therefore reducing the luminosity of these bright objects. This process has been shown in Chapter 3 to work well on large mass halos in which the cooling timescale is much longer than feedback timescale. When applied to small halos ($M \lesssim 10^{12}$ $M_{\odot}$), however, this process fails to be effective when the radius of the region that has cooled is well beyond the radius of the region in which energy injection happens. Semianalytic modelling requires the threshold at which AGN heating ceases to be effective to be $M = 10^{12}$ $M_{\odot}$. Winds and super-wind feedback produce luminosity functions that show good agreement with the observed luminosity function up to this mass.

In my simulations, the effect of energy injection on the cooling-flow changes dramati-

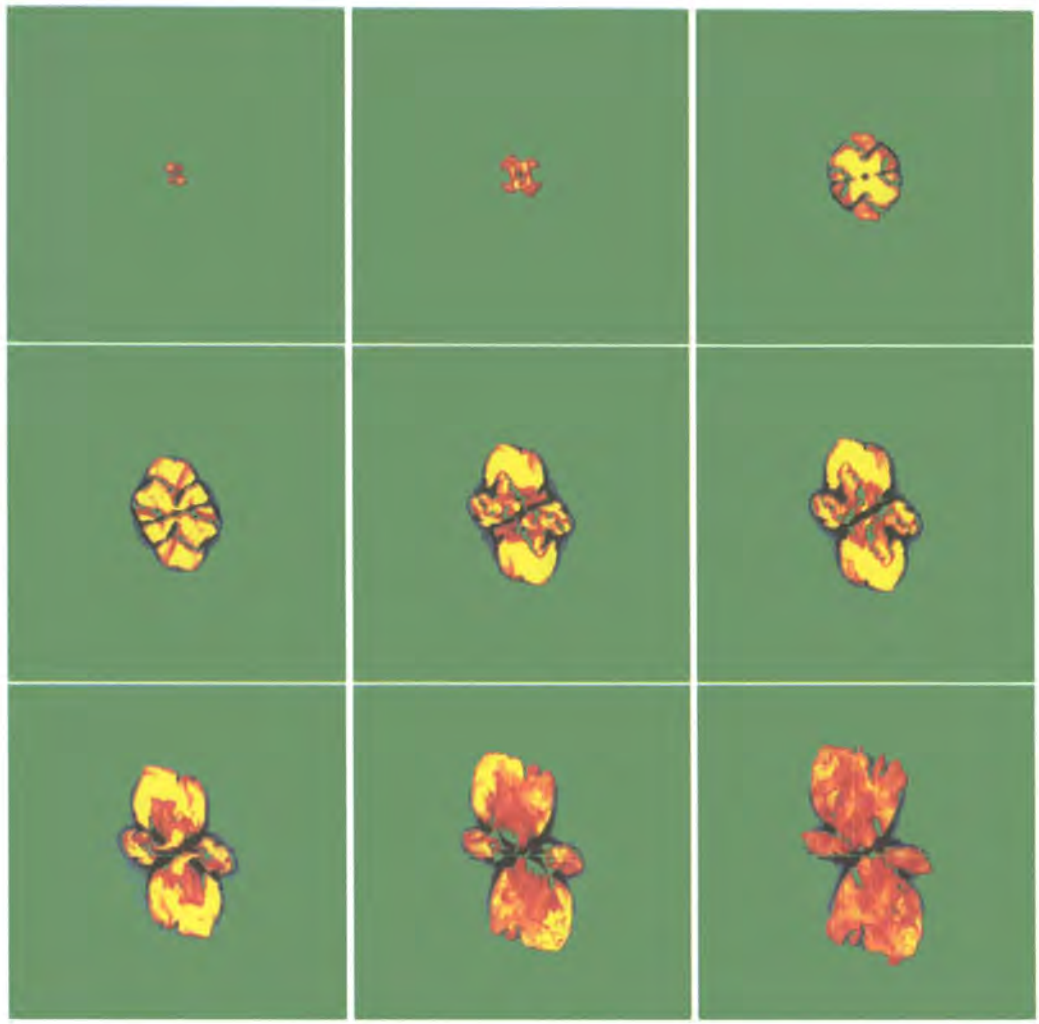

FIGURE 5.6.: Simulation B130. Plots of the temperature on a plane through the centre of the halo.

cally with the mass of the halo. It is clear from the data that the energy injection from AGN stops quenching the cooling at a characteristic mass of 10^{13} $M_{\odot}$. This value is somewhat larger than implied by the semianalytic models, but the simulations are still quite idealised. Moreover, the bubbles in the lower mass halo appear to behave differently than in the other cases. Bubbles do not seem to move much from their injection position. As mentioned in Section 5.2.4, this could be caused by the drag force exerted by the inflowing material. I am giving some more physical justifications of this in the following paragraphs.

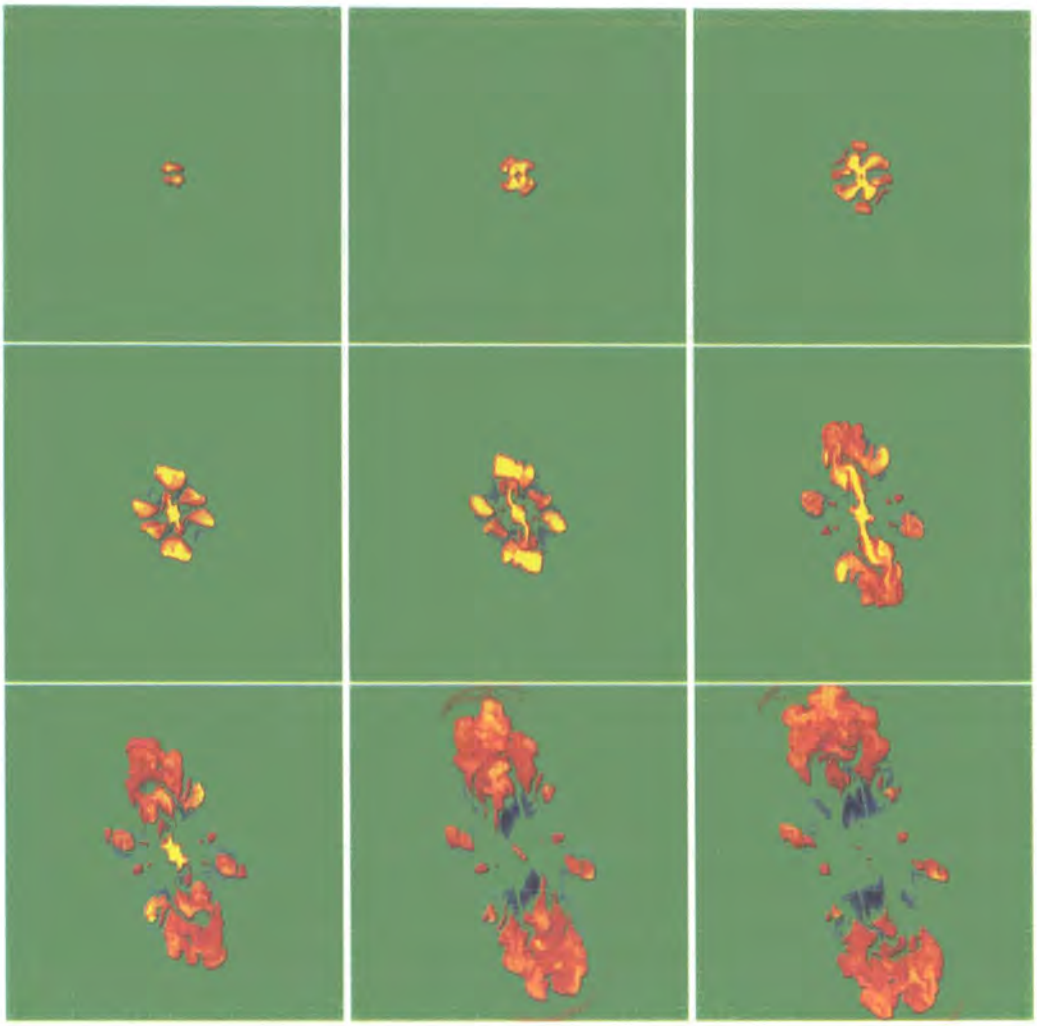

FIGURE 5.7.: Simulation B135. Plots of the temperature on a plane through the centre of the halo.

The motion of bubbles is determined by the balance between buoyancy and drag forces with the ICM (e.g. Churazov et al., 2001). The buoyant force experienced by a bubble of radius R at distance r is

$$F_{\rm b} = \frac{4\pi}{3} R^3 \rho_{\rm a} g_{\rm eff}$$

where $\rho_{\rm a}$ is the ambient density (the bubble density is supposed to be much smaller and neglected), and $g_{\rm eff}$ is the gravitational force field as measured in a frame of reference

attached to the ambient medium at distance r. The effective gravitational field, $g_{\rm eff}$, differs from the actual one only when the system deviates from hydrostatic equilibrium. In a cooling runaway its value is close to zero. The drag force is

$$F_{\rm d} = \frac{\pi}{2} C R^2 \rho_{\rm a} v_{\rm b}^2$$

assuming that the bubble velocity, $v_{\rm b}$, is much lower than the sound speed. C is the drag force coefficient, and depends on the bubble shape. Following Churazov et al. (2001) I adopted a value of $C = 0.75$. Equating the two forces gives

$$v_{\rm b}^2 = \frac{8 g_{\rm eff} R}{3C}$$

The pressure profile in simulation B120 (Figure 5.8, top-left panel) is quite shallow within a tenth of the virial radius. This indicates that there is no hydrostatic equilibrium. The value of $g_{\rm eff}$ is related to the pressure gradient by the following equation:

$$\frac{{\rm d}p}{{\rm d}r} = -\rho_{\rm a} g_{\rm eff}$$

An rough estimate of it from the pressure profile in simulation B120 gives values of one order of magnitude smaller than g. Therefore, the drag force can be large enough to keep the bubbles close to the cluster centre and prevent energy deposition above the cooling radius.

An estimate of g at the bubbles location for the dark matter profile of simulation B120 gives values of the bubble velocity of 20 and 95 km s^{-1} for $g_{\rm eff}$ equal to $0.01g$ and $0.2g$ respectively. The radial velocity profile in Figure 5.9 shows that the inflow velocity in the inner part of the halo is 50 to 100 km s^{-1}, that is greater than or comparable to the bubble speed. In such a regime one should expect the bubble not to move to larger radii (see also Nusser et al., 2006).

This could be the result of a too idealised approach to the problem. It is quite difficult to set up initial conditions for a model cluster which have such a short central cooling time. In Chapter 3 I discussed the problem of over-luminosity of the simulated cluster, that

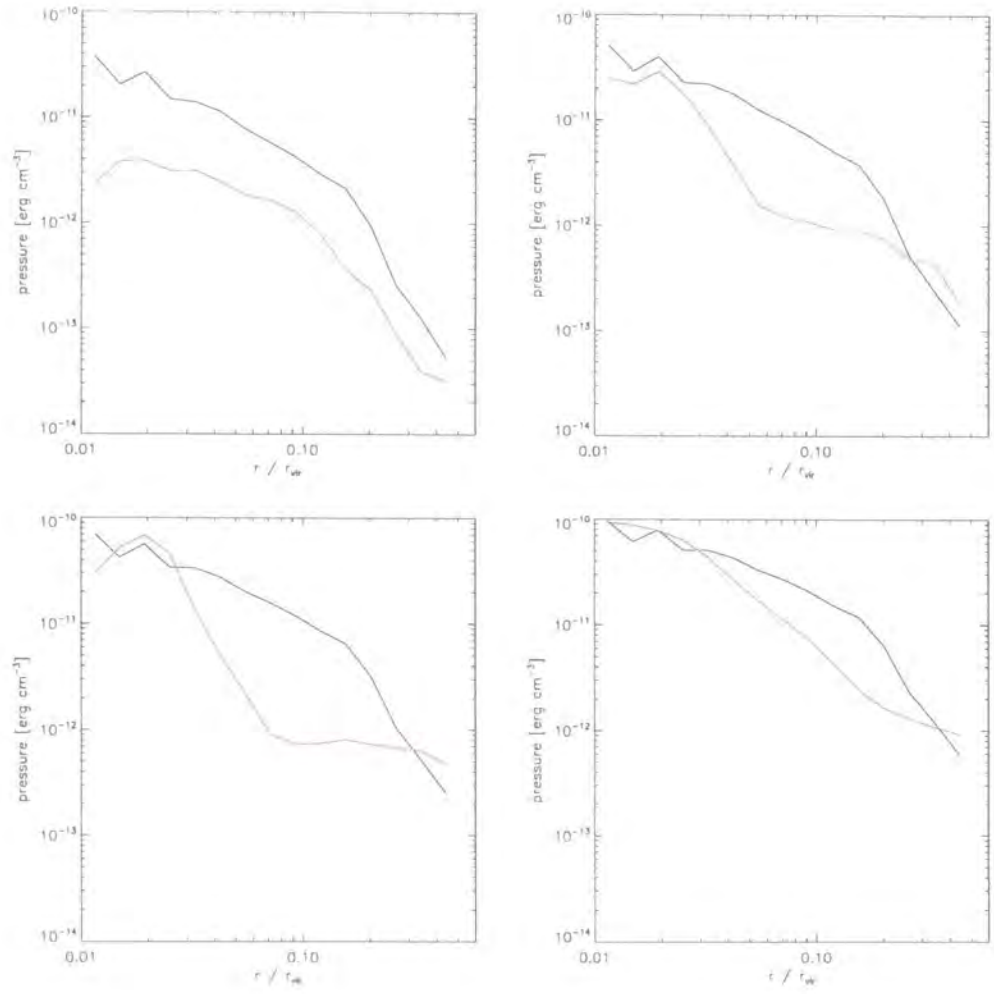

FIGURE 5.8.: From top-left to bottom-right, pressure profiles in simulations B120, B125, B130, B135. The black lines are the profiles at the beginning of the simulations, and the red lines are the profiles at the end.

was solved in Chapter 4 adopting observed temperature and density profiles. This approach would not be of great help in the case of the low mass halos discussed in this chapter. Using a centrally declining temperature profile, rather than a constant temperature, would need higher central density to balance the gravity. Therefore, cooling would increase making the problem even worse. A solution of the problem could be following the formation of the

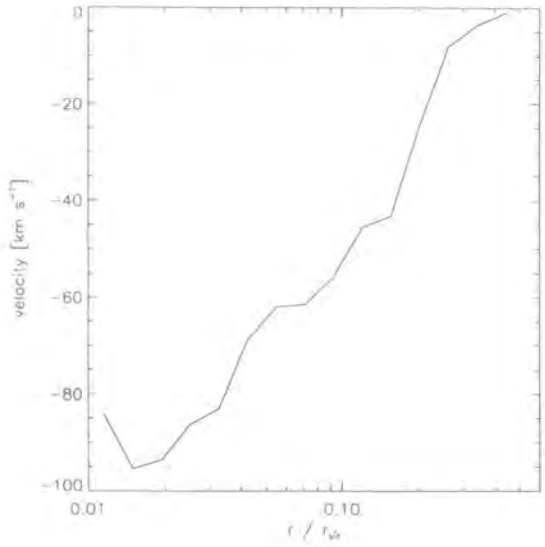

FIGURE 5.9.: Radial velocity profile for simulation S120.

halo starting from the collapse of an initial density fluctuation. Gas would be heated to the virial temperature until the cooling will start to dominate. This approach has more physical motivations (i.e. collapse is how first structures form in the Universe) and it is worth it to be explored in the future. Nevertheless, it introduces some technical problems (i.e. the inflow of matter from the simulation box boundaries) that could require more expensive simulations (i.e. larger box).

The model could also be improved if:

- feedback from other physical processes is included to smooth the transition from effective to ineffective quenching of the cooling;
- a different feedback scheme for AGN heating is adopted;
- halos are formed in a cosmological environment.

In the first case, winds and super-wind feedback would need to be stronger than what seems to be required to flatten the faint end of the luminosity function. This has the drawback

that the knee of the function would be smoothed as shown in the right panel of Figure 5.1. Changing the scheme of AGN energy injection by following the evolution of jets could reduce the characteristic halo mass at which heating ceases to be effective. With these models, energy is deposited in the surrounding gas when the kinetic energy of the jets is transformed into thermal energy by shock heating. Using jets has the advantage that part of the cold gas in the centre of the halo is evacuated, reducing the total luminosity. The third solution is more appealing. Forming halos in proper cosmological simulations adds an extra ingredient to the recipe: the halo dynamics. Halos in cosmological simulations are not isothermal or in hydrostatic equilibrium, and the gas can be heated by mergers with other sub-halos. Even though the dynamical timescale of these halos is longer than the cooling timescale, cooling should not be as catastrophic as in the idealised model presented in this chapter. As a result, the energy injection could be more effective in quenching cooling in the range of masses I simulated.

5.4. Conclusions

I presented simulations of AGN energy injection in cooling halos with mass ranging from 10^{12} $M_{\odot}$ to $10^{13.5}$ $M_{\odot}$. The variety of behaviour shown by the simulations is striking. There is an evident breakdown of the heating process at a mass of 10^{13} $M_{\odot}$ and below. Energy injection does not prevent the gas from cooling catastrophically in these halos. As discussed above, the value of the characteristic mass could be lowered if a more sophisticated model was adopted. The existence of a characteristic mass at which the AGN heating stops being effective is important for explaining the shape of the luminosity function of galaxies. Its value would mark the transition between star feedback and AGN feedback, and could reconcile the shallow slope of the faint end of the luminosity function with the sharp cut-off at the bright end. Improvements of the model are clearly necessary. Nevertheless, the simulations presented in this chapter give an indication that AGN heating is not effective in low mass halos.

AGN bubbles in a cosmological environment

In Chapter 3 I investigated the role of recurrent AGN activity in quenching cluster cooling-flows. As many other authors (e.g. Brüggen and Kaiser, 2001; Quilis et al., 2001; Basson and Alexander, 2003), I performed simulations of a model cluster with spherical dark matter potential. In this chapter, I improve the model by injecting bubbles of energy in a cluster formed in a cosmological simulation. The setup of the initial conditions is described in Section 6.2 and 6.3. In Section 6.4, I report on the results of a cooling-only simulation. In Section 6.5 I discuss the effect of energy injection. The initial conditions for the simulations were provided by Adrian Jenkins.

6.1. Introduction

The model cluster used in Chapter 3 has a constant dark matter mass distribution, and initially isothermal baryonic gas in hydrodynamical equilibrium. This model can give interesting insight into the relation between cooling-flows and AGN bubbles. Models like this are often used to investigate parameter space—in this case, a range of bubble energies—because of their ease to set up and simulate. However, the evolution of a real cluster is much more complex. In the hierarchical model of structure formation, large halos are built up by mergers of smaller halos. The mass infall perturbs the structure of the original ICM. The cooling flow could be temporarily suppressed and the gas in the core heated as gravitational energy is released and shocks are dissipated in the merging process (Gómez et al., 2002; Burns et al., 2004). Moreover, the motion of the ICM could influence the evolution of AGN bubbles. Collimated pairs of bubbles could rise buoyantly in different directions or be quickly disrupted by shear motion of the surrounding gas. It is therefore important to simulate AGN bubbles in an appropriate environment to confirm our previous results, and to investigate the presence of structural features observed in recent X-ray images.

While I was running the simulations, other authors have presented simulations of AGN bubbles in a cosmological environment (Brüggen et al., 2005). They mainly focus on the role of viscosity in dissipating bubble induced sound waves. They also show in their paper that sound waves, as well as bubble cavities, can be observed in *Chandra*-like X-ray images of their cluster. I will refer to this paper in the next sections, comparing their setup with mine, and underlining the differences between the two approaches.

The simulations presented in this chapter have not been completed due to the poor performance of the gravity solver of FLASH; I discussed this issue in Chapter 2. The preliminary results presented here seem to confirm that the model of energy injection of Chapter 3 works well.

6.2. Initial conditions

The cluster of galaxies selected for the simulations of AGN bubbles has been chosen according to the following criteria: its mass should be about 10^{14} $M_\odot$ and the cluster should not undergo any major mergers for a period of several cooling times. These choices may sound arbitrary, but they are mainly motivated by two reasons: (1) a cluster with similar properties has already been simulated and used to investigate a large range of bubble energies; the results presented in Chapter 3 provide an appropriate term of comparison and a reference for the choice of the energy to be injected. In addition, the cooling time at the centre of the cluster should be short enough to create a cooling flow; (2) if the effects of recurrent AGN events over a large timescale are to be simulated as in Chapter 3, then, the cluster must be in a relatively quiescent state over the simulation timescale. For the selection of the cluster and the setup of the initial conditions, I proceeded in this way: (1) the cluster was selected from an N-body, dark matter only, low resolution simulation; (2) the region in the initial conditions containing all the particles that are in the cluster at the final redshift was identified; (3) a high resolution grid was put on top of this region, and new initial conditions including baryonic matter were generated. I describe in details these steps in the following paragraphs.

The first N-body simulation was run with GADGET. The initial conditions contain 200^3 particles in a box of size 100 comoving Mpc. The initial redshift is $z = 63$, and the adopted cosmology is ΛCDM with matter fraction $\Omega_m = 0.25$, cosmological constant $\Omega_\Lambda = 0.75$, and Hubble parameter $h = 0.73$. A large sample of clusters was identified by a friends-of-friends algorithm (Davis et al., 1985) at redshift $z = 0$. A few clusters of mass about 10^{14} $M_\odot$ were traced back in time looking (qualitatively) at their merger history, and one of theme was selected the re-simulation. This cluster forms quite early in a relatively quiet environment. It undergoes its last major merger at redshift close to one, and looks quite relaxed afterwards. At redshift $z = 0$ the mass of the cluster is 1.0×10^{14} $M_\odot$ and its virial radius is 930 kpc. For a cluster of this mass, the cooling time is shorter than the Hubble

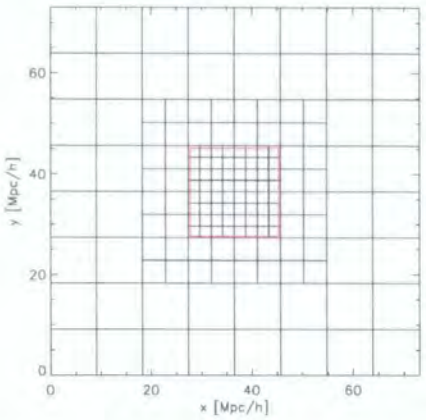

FIGURE 6.1.: Initial conditions. Structure of the AMR grid for the resimulation of the chosen cluster. The red box is the high resolution region containing all the particles belonging to the cluster at redshift $z = 0.0$.

time and the cluster should host a cooling flow in its core.

The initial conditions for the re-simulations were generated increasing the resolution of the region containing all the particles that make up the cluster at the final redshift. The initial AMR grid was set up starting from a cubic mesh covering the high resolution region which was put at the centre of the box. The mesh has the equivalent resolution of a uniform grid of 256^3 cells covering the entire box, and contains all the particles identified as belonging to the cluster at the present redshift. Its volume is just $1/64$ the simulation volume. Coarse grids were added ensuring that no jump of more than one level of refinement occurs from one block to another. With these criteria, the mesh at the 4th level of refinement covers the entire volume. The size of this mesh is 64^3 cells, allowing a coarsest linear resolution of 1.5625 Mpc. To each cell of the initial grid, a dark matter particle is assigned, the mass particle changing by a factor of eighth between contiguous levels of refinement. The mass of the high resolution particles is 2.57×10^9 $M_\odot$. The initial conditions for SPH gas particles (not used in the FLASH runs) were also computed. For each particle (dark matter and gas) the initial displacement and velocity are derived from the Zel'dovich

approximation and, for each cell, the initial baryon overdensity is specified. In Figure 6.1 I show the block structure of the initial conditions. Each block contains 8^3 cells. The red box represents the high resolution volume.

The refinement criteria adopted for the simulations is based on the dark matter particle density (see Section 2.5.1). One can work out that the finest grid needed to resolve the centre of the cluster should have a resolution of about 6 kpc. Thus, the maximum level of refinement should be 12. This means that an equivalent resolution of 16384^3 cells is reached and the size of one cell is 6.1 kpc. The region in which bubbles of energy are injected is then covered by 7-8 cells on the finest grid, and no extra refinement is necessary.

6.3. The simulations

The simulations I intended to perform would follow the formation of the cluster, the establishment of a cooling-flow in its core and the injection of bubbles of energy. The simulations are: (1) a simulation of the adiabatic formation of the cluster from the initial redshift to $z = 0.5$; (2) a resimulation from $z = 0.5$ with cooling to investigate the cooling-flow in the core of the cluster; (3) the final simulation with cooling and energy injection. As argued in Section 2.6, the scaling of the code is not optimal, and the gravity solver is very slow because of the heavy parallel communication it needs. Because of this, the gravity solver is ten times slower than the hydrodynamics solver. Therefore, a work-around is needed to reduce the computational costs. I explored two different strategies:

- I ran the adiabatic simulation with FLASH reducing the maximum level of refinement; the simulation was then restarted switching on the cooling;
- I ran the adiabatic simulation with GADGET and restarted it with FLASH including cooling and energy injection.

I describe in detail the two approaches in the following paragraphs.

As shown in the previous section, the resolution necessary to resolve the region of energy injection is equivalent to 16384^3 cells (12 levels of refinement). Because of the poor performance of the gravity solver of FLASH, the simulations would take too long to reach a suitable redshift to restart from. Therefore, I limited the maximum number of refinement levels to 8 (corresponding to the equivalent resolution of a uniform grid of size 1024^3 cells), and I ran the adiabatic simulation up to redshift $z = 0.5$. The gas profile was presenting an extended core, and the central density was too low. As a result, the central cooling time was of the order of 3 Gyr. This is clearly a resolution issue: the resolution is too low to resolve properly the cluster centre and the sub-halos that will end forming the cluster. If the sub-halos do not contain low entropy material, they do not contribute to the build-up of the central density peak of the cluster. Restarting the simulation allowing the resolution to increase did not give any improvement. I decided to abandon this strategy, for the moment, and to adopt a methodology similar to the one of Brüggen et al. (2005). In summary, they proceeded as follow:

1. they selected a cluster from a database of GADGET simulations; the simulation was ran including cooling, star formation and feedback;

2. they extracted a small box (2 Mpc on a side) centred on the cluster, and resimulated it imposing outflow and isolated boundary conditions for the edge of the volume and the multigrid Poisson solver, respectively;

3. they did not include cooling claiming it is not relevant on the simulated timescale of 140 Myr.

My method is somewhat different, and can be summarised as follows:

1. I ran the simulation with GADGET with refined initial conditions in adiabatic mode;

2. I mapped the entire volume on the AMR grid of FLASH and restarted the simulation with cooling;

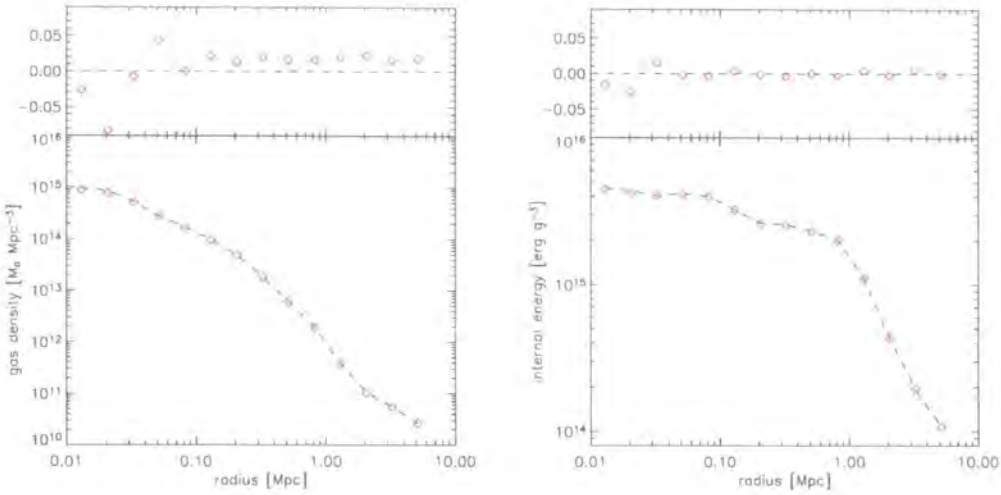

FIGURE 6.2.: **Left panel**: Initial density profiles. **Right panel**: Initial internal energy profiles. Dashed lines are the profiles computed from the GADGET output at redshift $z = 0.5$, and diamonds are the profiles after the mapping onto the AMR grid.

Although the configuration of Brüggen et al. (2005) allows a grid with higher resolution, it does not permit long runs to be performed. This is because the method does not resolve the large scale gravitational field, and because one can not ignore inflow or outflow of matter at the boundaries. In my case, I have no limitation on the simulated timescale a cosmological volume much larger than the cluster is evolved imposing periodic boundary conditions. I chose this setup mainly because it is suitable for simulating recurrent AGN events. I show the original and remapped density and internal energy profiles of the cluster in Figure 6.2. The internal energy profile is well recovered, as shown by the values of the residuals. The density profile shows an almost constant difference of about 2%. A possible explanation of this is that the density at the centre of the cell is not an accurate estimate of the mean density in the cell because the density profile is very steep. However, one should expect the average density to be higher than the central value if the density profile is positively curved. The plot in Figure 6.3 (left panel) shows the density profile of the cluster after mapping the gas on the grid using a different algorithm. Cells are subdivided in 7^3 sub-cells. The value of the density is then the average of the values computed on the cells of

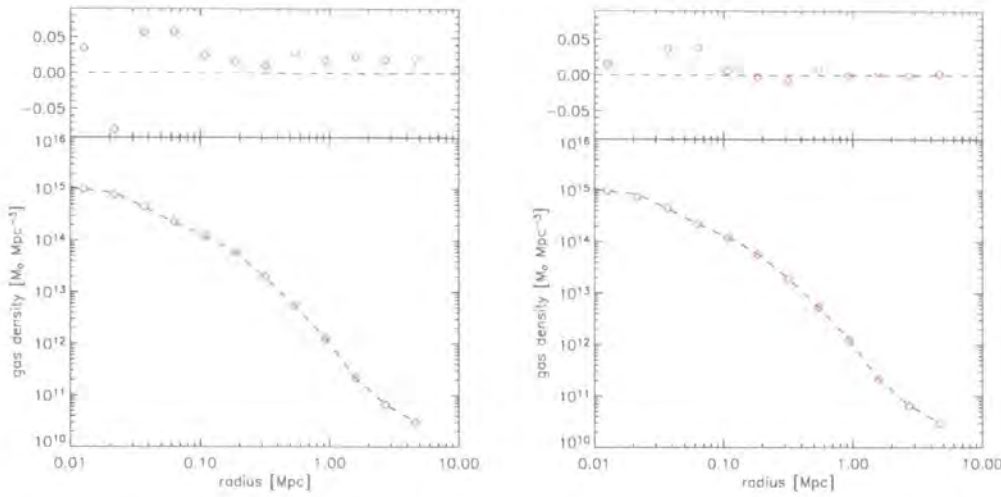

FIGURE 6.3.: **Left panel**: Initial density profiles mapping the density on sub-zones with rescaling. **Right panel**: Initial density profiles mapping the density on sub-zones without rescaling. Dashed lines are the profiles computed from the GADGET output at redshift $z = 0.5$, and diamonds are the profiles after the mapping into the AMR grid.

the sub-grids rescaled to conserve the total mass. The result is not dissimilar from what was obtained with cell centred mapping and, as expected, the density values are slightly larger. Another explanation is that the density is underestimated in low density regions, so that the total mass rescaling always increases the density in collapsed objects. The plot in Figure 6.3 (right panel) shows the density profile obtained with the sub-zones mapping, but not scaled by the total mass. The error is considerably reduced. The reason why the mapping is not accurate in low density regions is not yet clear and further investigations will be necessary. In Figure 6.4, I show the distribution of dark matter particles (left) and gas density (right) on a slice 20 Mpc thick, after the mapping into FLASH. Collisionless particles are colour coded to show the different masses. For a better visualisation, they are also sampled to have the same mean number density everywhere in the volume. The gas density panel shows the maximum value of the density along each line of sight to enhance the visibility of substructures.

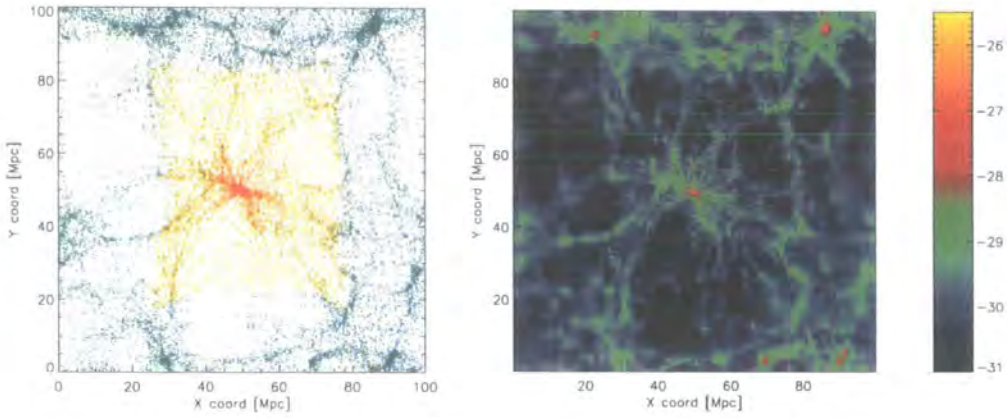

FIGURE 6.4.: Dark matter particles (left) and gas density (right) on a slice of 20 Mpc through the centre of the simulation box. Particles are colour coded to show different mass resolution, and, for the sake of clarity, adaptively reduced in number to have the same mean number density across all levels of refinement. The density is in g cm^{-3} and its maximum along each line of sight is plotted to enhance the visibility of substructures.

During the simulation, I allowed the code to refine the grid, as required, only in a small box around the cluster. The size of the box was set to contain at least all the particles belonging to the cluster at the final redshift. This ensure that substructures in the vicinity of the cluster are not under-refined. According to the definition of the AMR grid, the refinement propagates out from this region, but no extra fine blocks were added to over-densities far from the cluster. That is, the resolution decreases with the distance from the cluster, and distant, small-scale structures are not resolved. It is important to realise that at large distances the effect of gravity comes mainly from large-scale structures and hence the lower resolution will not affect it. The initial grid is fully refined and contains 9217 blocks (4719104 cells) just after the mapping of the gas particles. After a few time-steps, the total number of blocks is reduced to 4250 (2176000 cells) as a result of the derefinement of the external region. This gives an improvement of the speed in the code of almost 40%.

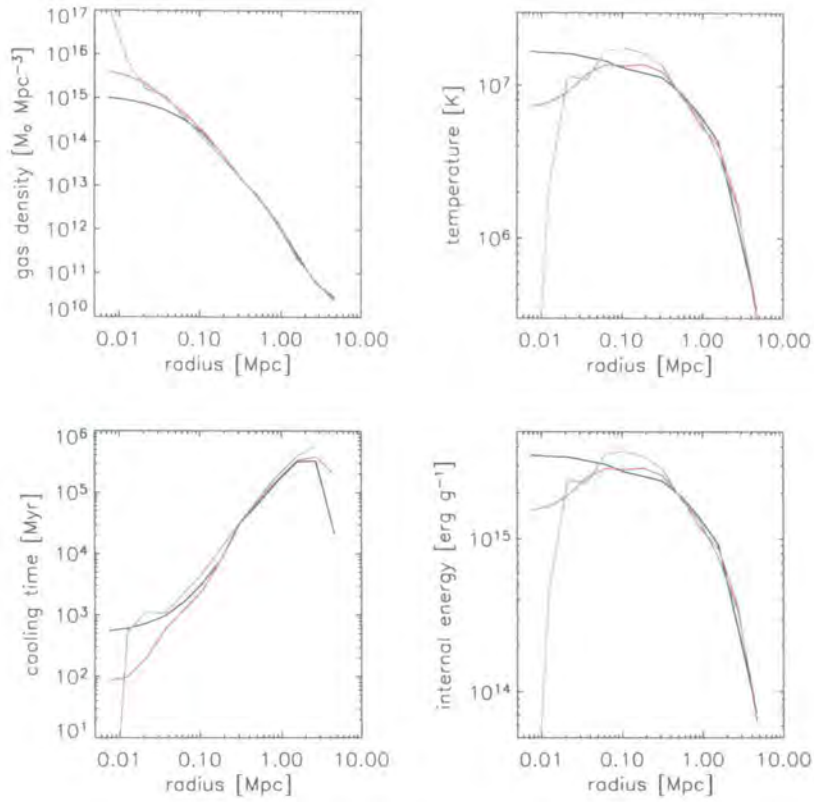

FIGURE 6.5.: Evolution of the density, temperature, cooling time and internal energy profiles. Black, red and blue lines are at time 9.6, 10.2 and 11.4 Gyr, respectively. The red line profiles are taken just before the large drop of the core temperature.

6.4. Cooling only simulation

I first performed a simulation without AGN bubbles to investigate the cooling-flow of the cluster. The simulation was ran assuming an adiabatic gas to $z = 0.5$ and with cooling included from $z = 0.5$ to $z = 0.184$. I discussed above the inaccuracy of the density profile mapping on the AMR grid. The slightly higher density introduced by the density mapping should give rise to a slightly over-pressurised gas. In fact, the gas distribution oscillates around the equilibrium configuration during the first part of the simulation. After a time interval of 0.94 Gyr, the gas settles in the gravitational potential, and starts to cool.

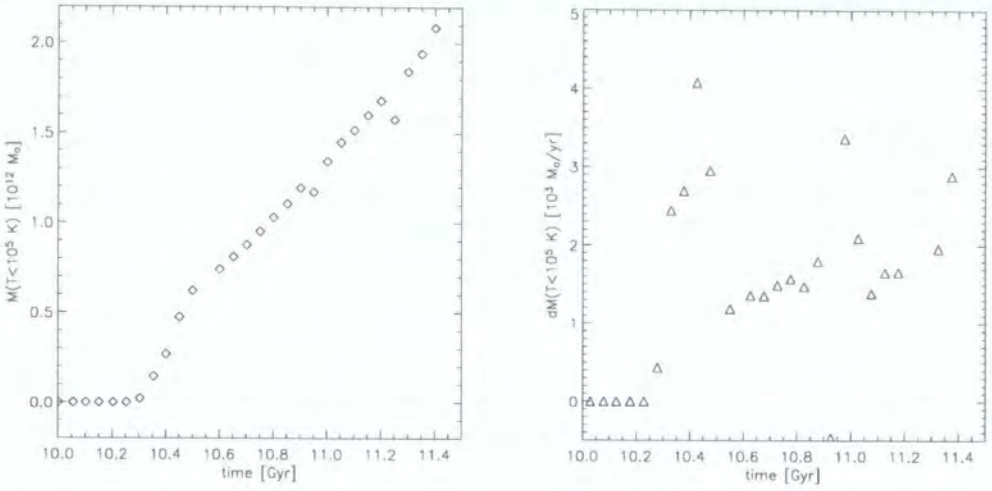

FIGURE 6.6.: Time evolution of the total mass cooling below the temperature of 10^5 K. The left panel shows the amount of mass in the core below the temperature threshold. In the right panel I plot the cooling rate on the same timescale.

The cooling time in the cluster core (the inner 100 kpc) is shorter than 1 Gyr when the calculation reaches redshift $z = 0.376$. It eventually becomes shorter than 200 Myr in some regions close to the core, but this is not visible in the spherically averaged profiles plotted in Figure 6.5. Here, I show shell-averaged profiles the density, the temperature, the cooling time and the internal energy at three different times. The black lines are the profiles after the cluster settles, at the time $t = 9.6$ Gyr ($z = 0.376$). The red lines are the profiles soon after the cooling phase has started, and before the temperature of the core has dropped below one-third of the initial temperature. The blue lines are the profiles at redshift $z = 0.184$—the end of the run. At this redshift, the core is denser by a factor of $\sim 10^{3-4}$, as a result of the catastrophic cooling phase.

The loss of pressure support, caused by the cooling of the inner part of the cluster, results in an inflow of matter towards the centre. The inflow will eventually become steady as observed by McCarthy et al. (2004), and cold matter will keep accumulating in the core. This is clearly seen in Figure 6.5, where the density is increasing the most inside a radius of 20 kpc. I computed the amount of cold gas as a function of time and I show this in the

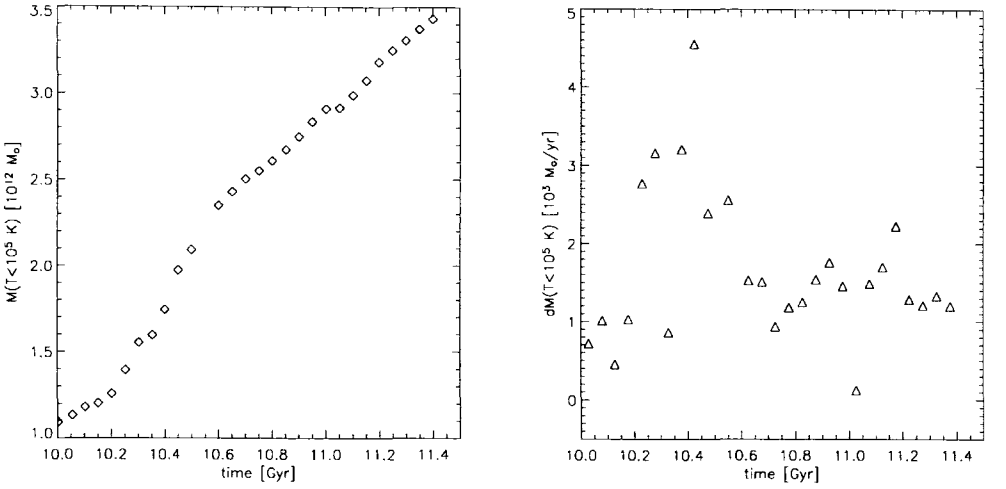

FIGURE 6.7.: Time evolution of the mass inside a radius of 100 kpc. The left panel shows the amount of mass in the inner 100 kpc. The right panel shows the cooling inflow rate on the same timescale.

left panel of Figure 6.6. In roughly 1 Gyr the amount of gas that has cooled below 10^5 K is 2×10^{12} $M_\odot$. This implies a mean inflow rate of 2000 $M_\odot$ yr^{-1}. In the right panel of the same figure, I plot the mass deposition rate as a function of time. The inflow of mass is clearly shown in Figure 6.7. The mass inside the inner 100 kpc increases with time at a rate up to 3000 $M_\odot$ yr^{-1} and peaks at the time when the core collapses in a run-away cooling process. After the cooling-flow reaches a steady configuration, the mass inflow rate keeps a mean value close to 1300 $M_\odot$ yr^{-1}. It is interesting to notice that the cooling-flow is already ongoing between 10.0 and 10.2 Gyr while the temperature of the core is rapidly dropping.

6.5. Energy injection simulation

The injection of bubbles of energy mimics the behaviour of a precessing central black hole. Pairs of bubbles are injected symmetrically about the centre of the cluster, each pair with a random orientation. The distance from the centre at which the bubbles are injected is kept

constant (12 kpc), resembling an AGN that periodically switches on with the same power. The duty cycle is 10^7 yr while energy is injected during a time interval of 3×10^5 yr. The duty cycle value is based on observational evidence as suggested by Fabian et al. (2003a). The initial radius of the bubbles is 10 kpc. The total energy injected per bubble matches the luminosity of the cluster in its cooling-flow phase, and it is 2.5×10^{59} erg. All these quantities are expressed in physical units.

Energy injection is switched on at the time of 10.14 Gyr, before the cooling starts its run-away process. A physical feedback scheme to couple the cooling-flow with the AGN activity is not yet included in the simulation. However, it is sensible to start the energy injection when the central density exceeds some threshold value. This resembles the feedback mechanism adopted by many authors (e.g. Springel et al., 2005; Kawata and Gibson, 2005). The density in the centre of the cluster in the adiabatic simulation is 0.034 cm^{-3}. I chose to start to inject energy when the central density changes by a factor of four. Thus, the density is about 0.14 cm^{-3} when the first pair of bubbles is injected. The reader should be aware that this density threshold is only used to trigger the injection of the first pair of bubbles and is not intended as a model of feedback. The next bubbles are injected periodically with a time interval given by the duty cycle discussed above.

One numerical problem is that the injection timescale is comparable with the time-step of the simulation. That is, the time-step could overshoot the initial and final energy injection times. To match the initial and final time of the AGN activity I modified the time-stepping criteria. This ensures that the energy injected is exactly what is specified in the parameters file. However, the energy could be injected in just one time-step. In this case, the bubble would have the density of the ICM and the energy just injected would be quickly dissipated by radiative cooling. Moreover, the resulting temperature (and therefore pressure) would be much lower than what is observed. In this case, the bubbles do not expand and do not have any appreciable effect on the ICM. Therefore, I put a constraint of the time-step ensuring that several time-steps are performed during the injection time interval.

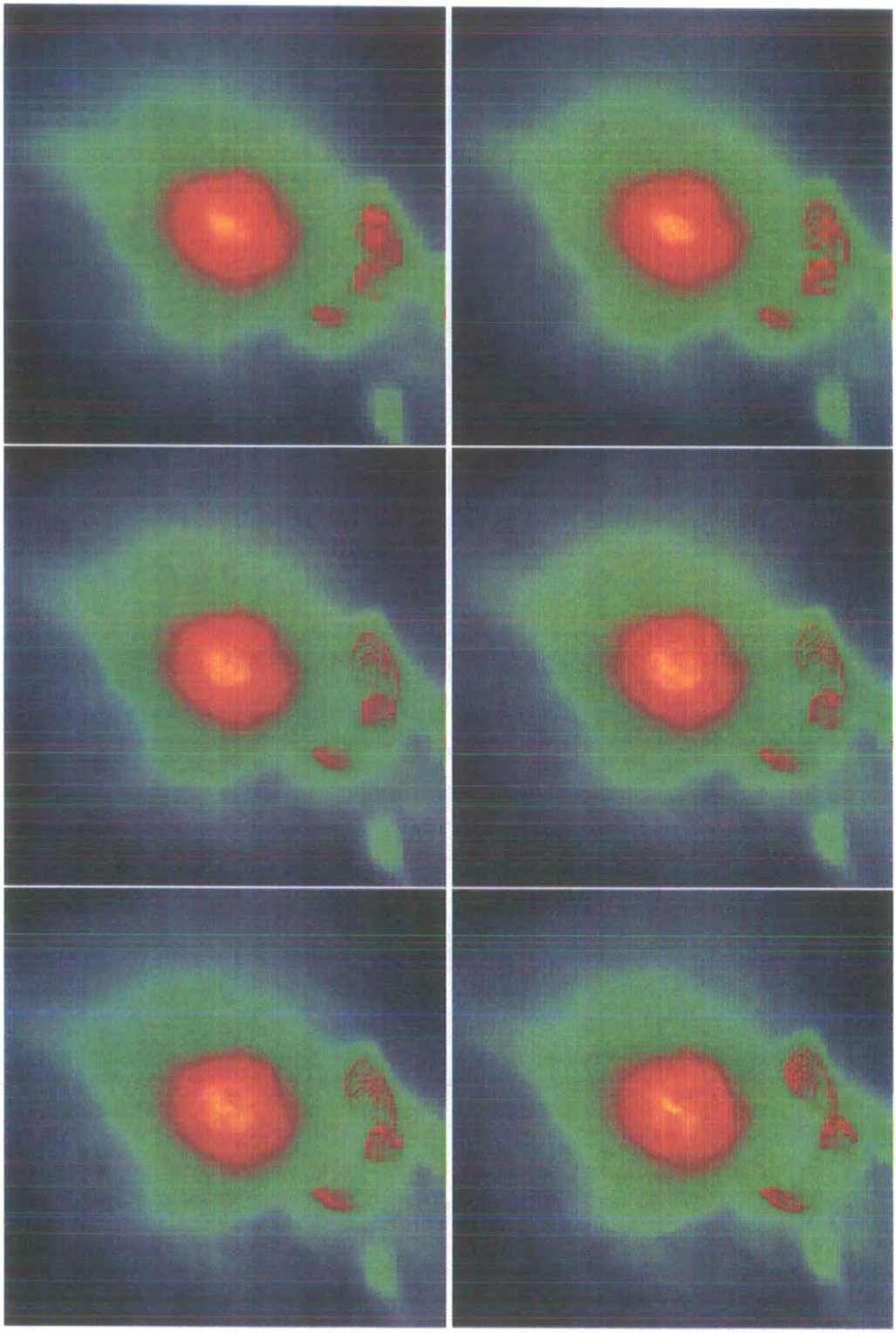

FIGURE 6.8.: Evolution of bubbles in the cosmological simulation. The surface brightness is plotted for a box of 1.56 Mpc side centred on the cluster.

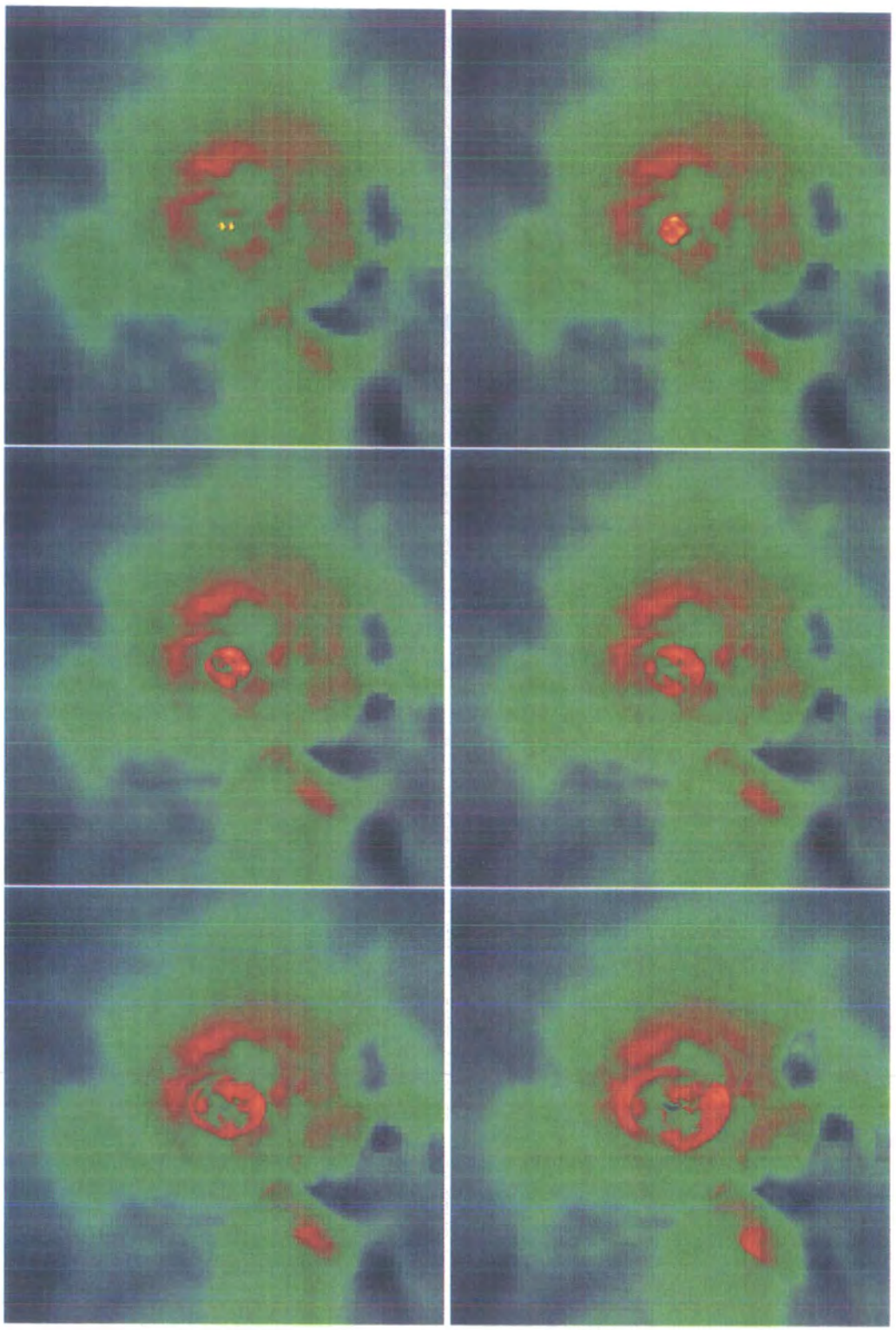

FIGURE 6.9.: Evolution of bubbles in the cosmological simulation. The spectroscopic-like temperature is plotted for a box of 1.56 Mpc side centred on the cluster. The plotted temperature is in the range $[10^2, 3 \times 10^7]$ K.

This seems to solve the problem, but further tests should be made. One improvement that I am actually working on (not presented in this work) is to simulate the formation of bubbles with jets. Exploring a large sample of parameters, one could put some constraints on the final density and internal energy of the bubbles. These values could be used to define bubble parameters in simulations similar to the ones presented in this work.

6.6. Results

Even though the simulation has not yet finished, some interesting results have been achieved. I would point the attention of the reader to the images shown in Figure 6.8 and 6.9. In Figure 6.8, the emissivity of the cluster is plotted. The emissivity is not the usual bolometric emissivity proportional to $\rho^2 T^{1/2}$, but comes from the integration along the line of sight of the emitted energy computed by the cooling routine; in this setup, the variation of energy due to cooling and heating processes calculated by the cooling routine is stored at each time-step for each cell. In the top-right image the cavities produced by the bubbles' expansion are clear. One can also recognise the brighter edges of the bubbles.

In Figure 6.9, I show the spectroscopic-like temperature of the cluster (see Chapter 4). The temperature maps show clearly the evolution of the sound wave generated by the expansion of the bubbles. It is interesting to notice that the sound wave that appears almost spherical in the emissivity images, does look fragmented in the temperature projections. It could be that crossing high density regions does not enhance the temperature of the dense gas much, and, since the temperature of this regions is lower, it slows down the wave. These arc features look like the ripples in the images of the *Perseus* cluster.

The preliminary results show that the cooling-flow may be quenched by the AGN bubbles, but the simulation has not ran longer enough. What I observe is that the amount of cold gas is constant over the simulation time. However, it is not clear whether the cooling-flow will be completely suppressed as shown in the simulations of Chapter 3 or the mass inflow will be just reduced. These issues can not be clarified yet and will be investigated

in the future with a major improvement of the code performance. Another aspect of the simulation that can not be investigated yet is the evolution of the bubbles in the cluster ambient. Bubbles may be quickly disrupted by the motion of the ICM and deposit their energy much closer to the centre than expected. Simulation with resolution higher than the one adopted in this work would help to clarify this issues.

6.7. Conclusions

I presented a first attempt to simulate AGN heating in a cluster of galaxies that forms in a cosmological environment. The results have been affected by the poor code performance. However, it seems that the heating process is working during the short timescale that has been simulated. Once the performance of the code will be improved I will be provided with a powerful tools to investigate the AGN heating of "live" clusters.

Conclusions and future work

Chandra and *XMM-Newton* X-ray telescopes have been providing a wealth of data on galaxy clusters. The ICM presents complex structures that comes from the merger history of clusters and from the activity of their central AGN. In this work I focused on the heating process related to the evolution of buoyant bubbles generated by the central AGN. I also investigated the formation and the evolution of sound waves and the possibility of determining the duty cycle of the AGN by using measurements of sound wave positions. In the next section I summarise the results presented in this thesis. In Section 7.2 I discuss the work that has to be done to improve the heating model used this far.

7.1. Summary

7.1.1. Quenching cluster cooling-flows

In Chapter 3, I discussed the effect of recurrent AGN mechanical heating of the ICM with respect to the cooling of the gas in a model cluster. This process is effective in quenching the cooling-flow in the simulated model by up-lifting cold material in the centre of the cluster and mixing it with high entropy material at radii larger than or comparable to the cooling radius. The gas cooling-flow is completely stopped once the mean energy injection rate matches the emissivity of the cluster. I proved that the turbulent motion of the gas, necessary for an effective mixing of low and high entropy material, would not affect filamentary Hα structures observed in the *Perseus* cluster. Moreover, bubble motion through the ICM would not have any major effect on a pre-existent metallicity gradient.

7.1.2. Detection of sound waves and determination of AGN duty cycle

In Chapter 4, I investigated the detectability of sound waves in X-ray images of clusters and the accuracy in the determination of the duty cycle of the AGN by using the relative position of sound waves. I adopted a model cluster with temperature and mass density profiles similar to the ones of the *Perseus* cluster and I showed that sound waves generated in my simulation can be detected in deep *Chandra* images by using synthetic X-ray observations of the simulated cluster taken with the code X-MAS. The determination of the duty cycle from simulations shows that not knowing the initial position of the bubbles that generated the sound waves can lead to an estimate of its value that is a factor of two larger.

7.1.3. Characteristic mass of AGN heating

In Chapter 5, I simulated a sample of halos with different mass to investigate the low mass limit for the applicability of the AGN heating model of Chapter 3. I showed that there is an evident breakdown of the heating process at the characteristic mass of 10^{13} $M_{\odot}$. Semi-

analytic models require this threshold to be at a mass of $\sim 10^{12}$ $M_{\odot}$. I argued how possible improvements of my model would lower the mass threshold.

7.2. Future work

7.2.1. Cosmological simulations

In Chapter 6, I applied the AGN heating process to a cluster formed in a cosmological environment. I argued about the code performance problem in solving the Poisson's equation on the adaptive grid. In the near future, I will focus on the speed-up of the gravitational solver. As pointed out in Chapter 6, the use of the FFT technique would give an improvement in performance. However, a speed-up of the code can be achieved by modifying the time-stepping scheme. One should notice that, in most of the cases, the Courant time-step limiter provides a value of the time-step that is shorter than the time-step needed by the gravity solver. Thus, much less time-steps are required to evolve the gravity. Then, a performance improvement would be achieved by evolving the hydrodynamics and the gravity with different time-steps. I will explore this way in the future.

7.2.2. Jets and bubble formation

In my model, I have been injecting bubbles of energy with values of energy, radius and distance from the centre resembling observed values. Bubbles are injected without explicitly taking into account their formation history. The next step is to model AGN jets in order to investigate how bubble formation is related to the jet parameters: radius, power and mass ejection rate. I will run a set of simulations to cover a large volume of the parameter space and to study the interaction of the jet with the surrounding gas.

7.2.3. Feedback scheme

Once the formation of bubbles is linked to the parameters of the jet, the next step would be to investigate how the jet depends on the properties of the ICM. Modelling the accretion of a black hole, the ejection of material and its interaction with the ICM would require the simulation of a large range of length-scales. This is still a computational challenge even using the most powerful computers currently available. The model should link the gas physical properties "in the vicinity" of the black hole to the energy feedback from the black hole without simulating the physics of the black hole itself. Such a *black box* model would be the best way of mode-ling a realistic AGN feedback in clusters of galaxies.

Mathematics

A.1. Hyperbolic systems of conservation laws

Let us consider a three-dimensional system of n partial differential equations

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}^{\alpha}(\mathbf{u})}{\partial x^{\alpha}} = 0 \tag{A.1}$$

where $\alpha = 1, 2, 3$, $\mathbf{u} : R \times R^3 \rightarrow R^n$ is the n-dimensional vector of unknowns and $\mathbf{F}^{\alpha} = (\mathbf{f}, \mathbf{g}, \mathbf{h}) : R^n \rightarrow R^n$ are functions of $\mathbf{u}$ called fluxes. Einstein's summation is assumed on Greek indexes, so that $\partial \mathbf{F}^{\alpha}(\mathbf{u})/\partial x^{\alpha} = \partial \mathbf{f}(\mathbf{u})/\partial x^1 + \partial \mathbf{g}(\mathbf{u})/\partial x^2 + \partial \mathbf{h}(\mathbf{u})/\partial x^3$. If the initial condition $\mathbf{u}(x^{\alpha}, 0) = \mathbf{u}_0$ is given, then *Cauchy's problem* (or initial value problem)

is defined. The above system can be rewritten as

$$\frac{\partial \mathbf{u}(x^\alpha, t)}{\partial t} + \mathcal{A}^\beta(\mathbf{u}) \frac{\partial \mathbf{u}(x^\alpha, t)}{\partial x^\beta} = 0 \tag{A.2}$$

where $\mathcal{A}$ are the Jacobian matrices defined by

$$\mathcal{A}^\alpha = \frac{\partial \mathbf{F}^\alpha(\mathbf{u})}{\partial \mathbf{u}}$$

System A.1 is called *strictly hyperbolic* if the Jacobian matrices $\mathcal{A}^\alpha$ can be diagonalised with real and distinct eigenvalues (the *characteristic velocities*), λ_k^α $(k = 1, \ldots, n)$, and the corresponding eigenvectors form a complete set in R^n. Equation A.1 defines a *system of conservation laws*. Given a volume $V \in R^3$ with boundary ∂V, its integral form is

$$\frac{\partial}{\partial t} \int_V \mathbf{u} \, \mathrm{d}V + \int_{\partial V} \mathbf{F}(\mathbf{u}) \cdot \mathbf{n} \, \mathrm{d}S = 0$$

where $\mathbf{n}$ is the unitary vector orthogonal the surface $\mathrm{d}S$. Thus, the time variation of $\mathbf{u}$ is equal to the losses through the boundary ∂V, and, in the case of the hydrodynamics laws, this implies conservation of energy, momentum and mass. This is an important property of those numerical algorithm designed to solve a hyperbolic system of conservation laws. That is, in the absence of sources, those quantities that have to be conserved, are exactly conserved in the finite difference form of the differential equations.

A.2. Definition of Riemann's problem

Given discontinuous initial conditions

$$\mathbf{u}(x^\alpha, 0) = \begin{cases} \mathbf{u}_\mathrm{L} \text{ if } x^\alpha < x_0^\alpha \\ \mathbf{u}_\mathrm{R} \text{ if } x^\alpha > x_0^\alpha \end{cases}$$

where x_0^α $(\alpha = 1, 2, 3)$ is the position of the discontinuity, the initial value problem for the hyperbolic system of conservation laws A.1 is called *Riemann's problem*.

Riemann's problem is solved at the interface between cells, once the values of $\mathbf{u}_\mathrm{L}$ and $\mathbf{u}_\mathrm{R}$ are specified. These are the values of $\mathbf{u}$ at the left and right side of the interface, computed

interpolating cell centered values at the left and right of the interface. The solution of Riemann's problem gives the fluxes at each cell interface used to update $\mathbf{u}$ in the time interval $[t, t + \Delta t]$. Restricting our discussion to the x direction ($\alpha = 1$), the numerical flux is then

$$\mathbf{f} = \frac{1}{2}\left[\mathbf{f}(\mathbf{u}_\mathrm{L}) + \mathbf{f}(\mathbf{u}_\mathrm{R}) - \sum_{k=1}^{n} |\lambda_k^1| \Delta\omega_k \mathbf{R}_k^1\right]$$

where λ_k^1 and $\mathbf{R}_k^1$ are, respectively, the eigenvalues and the k-right eigenvectors of the Jacobian matrix $\mathcal{A}^1$ (see previous section) calculated in $\mathbf{u} = (\mathbf{u}_\mathrm{L} + \mathbf{u}_\mathrm{R})/2$. The quantities $\Delta\omega_k$ are the jumps in the local characteristic variables through each interface. They are obtained from the relation

$$\mathbf{u}_\mathrm{R} - \mathbf{u}_\mathrm{L} = \sum_{k=1}^{n} \Delta\omega_k \mathbf{R}_k^1$$

Resolution convergence tests

I tested the convergence of the code by running simulation S2.0 at increased resolution. In Figure B.1, I compare the total energy evolution for simulation S1.0 (6 levels of refinement, resolution of 7.6 kpc) and S2.0 (6 (7.6 kpc), 7 (3.8 kpc), 8 (1.9 kpc) and 9 (0.9 kpc) levels of refinement). The codes were run for different lengths of time because of limitations in the computational time and memory available. The differences between the four versions of the S2.0 run are small in comparison to the relative evolution of the S1.5 and S2.0 simulations. The figure also shows that the radiated energy does not increase systematically with resolution. Comparing the total radiated energy at $t = 0.45$ Gyr, the four versions

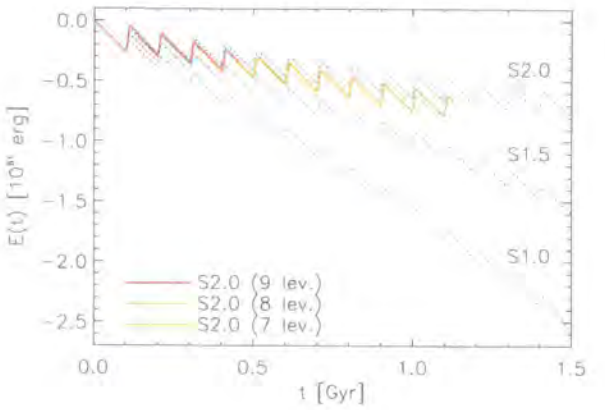

FIGURE B.1.: As Fig. 3.1, the energy evolution is shown for simulations S1.0, S1.5 and S2.0 run with 6 levels of refinment (dotted lines). I compare these with simulation S2.0 run at different resolutions (7, 8 and 9 refinement levels).

of the S2.0 simulation differ by less than 3%. At the last common output time, the 6 and 7 level simulations differ by 4%. I also compared the density and entropy profiles of the simulations at this output time, and found a similar level of agreement. I therefore chose to run the code with a maximum refinement level of 6, this being a good compromise between speed and the desired accuracy for the relevant quantities.

Simulation S0.0 was also used to check the convergence of energy conservation as the maximum number of levels of refinement changes. With just 4 refinement levels, the total energy has an error of $\sim 1\%$, while with 6 levels the error is less than 0.1%. I decided to run the simulations with a maximum number of 6 levels. The equivalent resolution of a fixed-grid Eulerian code is 256^3.

Up-lift and turbulence appear to play the most significant role in generating entropy. Low entropy material is drawn out of the core by the up draft generated as the bubble starts to rise. This material actually overtakes the bubble material itself as vortices shred the bubble. Thus, the vorticity generates entropy and redistributes energy both by mixing the hot bubble material with the surrounding ICM and by mixing the lowest entropy material with higher entropy ICM.

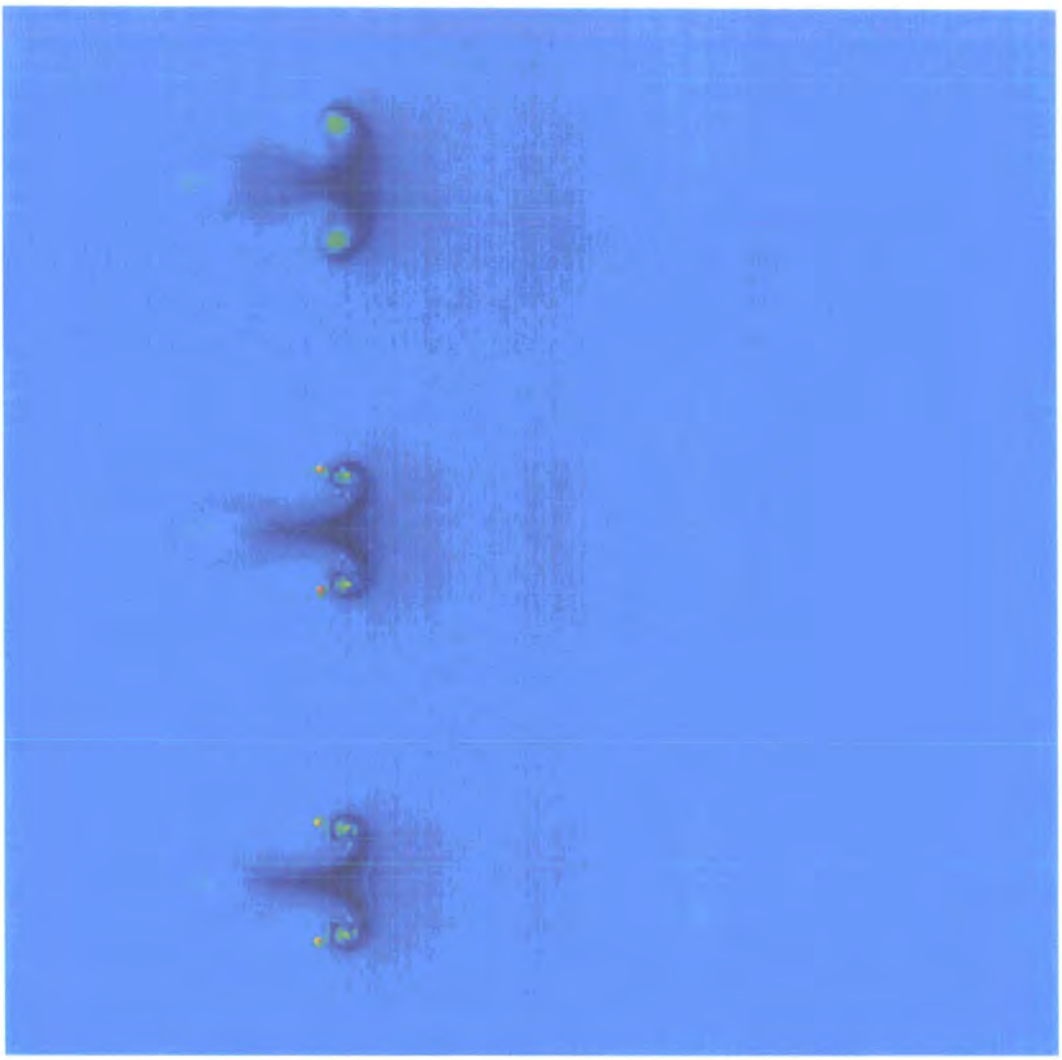

FIGURE B.2.: Comparison between high and low resolution simulations of one bubble. At the top, a simulation with the resolution adopted in the paper; in the center, a high resolution simulation degraded to the simulation adopted in the paper; at the bottom, the same simulation at maximum resolution.

I was concerned that the effectiveness of this mixing might depend on the resolution of the simulation. My simulations clearly resolve the flow around the bubble, but cannot track the mixing of material down to molecular scales. If such high resolution were to prove essential, everyone working in this area could abandon any hope of ever obtaining realistic results from cosmological hydrodynamic simulations. Fortunately, the key properties of

mixing and entropy generation in turbulent flows are usually dominated by the largest scale eddies (e.g. Tennekes and Lumley, 1972). An energy cascade is generated that transfers the energy to smaller and smaller vortices, until the energy is eventually dissipated by molecular viscosity. This is the basis of Kolmogorov (1941) seminal work on turbulence. In sufficiently high resolution simulations, the turbulent cascade is terminated by numerical mixing within a single cell, but the large scale properties of the flow should be modelled well.

The practical result of this is that it is not necessary to resolve the mixing process down to molecular scales. It is possible for the numerical resolution to truncate the turbulent cascade at larger scales without adversely affecting the global dissipation and mixing of the flow. Of course, it is very hard to establish what counts as "sufficiently high resolution". All I can do simulate the flow at a range of resolutions and test whether the key results converge.

To further demonstrate this, I ran some test simulations of single bubbles, ignoring cooling (so as to optimise the computation effort) and using a smaller computational domain. With this set-up, I can directly compare the flow around a bubble at the fiducial resolution with a calculation performed with an order of magnitude better resolution. In Figure B.2 I show a qualitative comparison between the two simulation. At the top, the bubble at the resolution adopted in Chapter 3; at the bottom, the high resolution simulation; in the center, the same simulation but with degraded resolution. The snapshots are taken at the same time. In all simulations, the bubble moves at the same speed through the ICM, as shown in the picture by their position. The large eddies break into smaller structures once the resolution is increased. But, when the simulations are degraded to the same resolution, the structure of the flow is almost indistinguishable and the overall evolution of the bubble is not significantly affected.

Bibliography

S. W. Allen, R. W. Schmidt, and A. C. Fabian. The X-ray virial relations for relaxed lensing clusters observed with Chandra. MNRAS, 328:L37–L41, Dec. 2001.

A. Babul, M. L. Balogh, G. F. Lewis, and G. B. Poole. Physical implications of the X-ray properties of galaxy groups and clusters. MNRAS, 330:329–343, Feb. 2002.

R. Ball, J. O. Burns, and C. Loken. The radio properties of cD galaxies in Abell clusters. II - The VLA sample. AJ, 105:53–66, Jan. 1993.

M. L. Balogh, A. Babul, and D. R. Patton. Pre-heated isentropic gas in groups of galaxies. MNRAS, 307:463–479, Aug. 1999.

J. F. Basson and P. Alexander. The long-term effect of radio sources on the intracluster medium. MNRAS, 339:353–359, Feb. 2003.

M. C. Begelman. Impact of Active Galactic Nuclei on the Surrounding Medium. In *ASP Conf. Ser. 240: Gas and Galaxy Evolution*, pages 363–+, 2001.

C. L. Bennett, A. J. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results. ApJ, 464:L1+, June 1996.

C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Grea-

son, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. ApJS, 148:1–27, Sept. 2003a.

C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission. ApJS, 148:97–117, Sept. 2003b.

A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole. What Shapes the Luminosity Function of Galaxies? ApJ, 599:38–49, Dec. 2003.

J. Binney. On the origin of the galaxy luminosity function. MNRAS, 347:1093–1096, Feb. 2004.

J. Binney and G. Tabor. Evolving Cooling Flows. MNRAS, 276:663–+, Sept. 1995.

S. Borgani, F. Governato, J. Wadsley, N. Menci, P. Tozzi, T. Quinn, J. Stadel, and G. Lake. The effect of non-gravitational gas heating in groups and clusters of galaxies. MNRAS, 336:409–424, Oct. 2002.

R. G. Bower, A. J. Benson, C. G. Lacey, C. M. Baugh, S. Cole, and C. S. Frenk. The impact of galaxy formation on the X-ray evolution of clusters. MNRAS, 325:497–508, Aug. 2001.

M. Brüggen and C. R. Kaiser. Buoyant radio plasma in clusters of galaxies. MNRAS, 325: 676–684, Aug. 2001.

M. Brüggen, M. Ruszkowski, and E. Hallman. AGN heating and dissipative processes in galaxy clusters. *ArXiv Astrophysics e-prints*, Jan. 2005.

J. O. Burns. The radio properties of cD galaxies in Abell clusters. I - an X-ray selected sample. AJ, 99:14–30, Jan. 1990.

J. O. Burns, P. M. Motl, M. L. Norman, and G. L. Bryan. On the Formation of Cool, Non-Flowing Cores in Galaxy Clusters Via Hierarchical Mergers. In *The Riddle of Cooling Flows in Galaxies and Clusters of galaxies*, 2004.

C. L. Carilli and G. B. Taylor. Cluster Magnetic Fields. ARA&A, 40:319–348, 2002.

E. Churazov, M. Brüggen, C. R. Kaiser, H. Böhringer, and W. Forman. Evolution of Buoyant Bubbles in M87. ApJ, 554:261–273, June 2001.

E. Churazov, R. Sunyaev, W. Forman, and H. Böhringer. Cooling flows as a calorimeter of active galactic nucleus mechanical power. MNRAS, 332:729–734, May 2002.

S. Cole, A. Aragon-Salamanca, C. S. Frenk, J. F. Navarro, and S. E. Zepf. A Recipe for Galaxy Formation. MNRAS, 271:781–+, Dec. 1994.

S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk. Hierarchical galaxy formation. MNRAS, 319:168–204, Nov. 2000.

S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, M. Colless, C. Collins, W. Couch, N. J. G. Cross, G. Dalton, V. R. Eke, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, K. Glazebrook, C. Jackson, A. Jenkins, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, B. A. Peterson, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: Power-spectrum analysis of the final dataset and cosmological implications. *ArXiv Astrophysics e-prints*, Jan. 2005.

P. Colella and P. R. Woodward. The Piecewise Parabolic Method (PPM) for gas-dynamical simulations. *J. Comp. Phys.*, 54:174–201, Apr. 1984.

C. S. Crawford, S. W. Allen, H. Ebeling, A. C. Edge, and A. C. Fabian. The ROSAT Brightest Cluster Sample - III. Optical spectra of the central cluster galaxies. MNRAS, 306:857–896, July 1999.

J. H. Croston, M. J. Hardcastle, and M. Birkinshaw. Evidence for radio-source heating of groups. *ArXiv Astrophysics e-prints*, Nov. 2004.

C. Dalla Vecchia, R. G. Bower, T. Theuns, M. L. Balogh, P. Mazzotta, and C. S. Frenk. Quenching cluster cooling flows with recurrent hot plasma bubbles. MNRAS, 355:995–1004, Dec. 2004.

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. ApJ, 292:371–394, May 1985. doi: 10.1086/163168.

S. De Grandi and S. Molendi. Metallicity Gradients in X-Ray Clusters of Galaxies. ApJ, 551:153–159, Apr. 2001.

S. De Grandi and S. Molendi. Temperature Profiles of Nearby Clusters of Galaxies. ApJ, 567:163–177, Mar. 2002.

K. Dolag, M. Jubelgas, V. Springel, S. Borgani, and E. Rasia. Thermal Conduction in Simulated Galaxy Clusters. ApJ, 606:L97–L100, May 2004.

C. Done and M. Gierlinski. Scaling black holes. I. Narrow-line Seyfert 1 variability. *ArXiv Astrophysics e-prints*, May 2005.

A. C. Edge. The detection of molecular gas in the central galaxies of cooling flow clusters. MNRAS, 328:762–782, Dec. 2001.

A. C. Edge and D. T. Frayer. Resolving Molecular gas in the Central Galaxies of Cooling Flow Clusters. ApJ, 594:L13–L17, Sept. 2003.

A. C. Edge, R. J. Wilman, R. M. Johnstone, C. S. Crawford, A. C. Fabian, and S. W. Allen. A survey of molecular hydrogen in the central galaxies of cooling flows. MNRAS, 337: 49–62, Nov. 2002.

J. A. Eilek. Radio Galaxies in Cooling Cores. In *The Riddle of Cooling Flows in Galaxies and Clusters of galaxies*, 2004.

V. R. Eke, J. F. Navarro, and M. Steinmetz. The Power Spectrum Dependence of Dark Matter Halo Concentrations. ApJ, 554:114–125, June 2001.

S. Ettori, A. C. Fabian, and D. A. White. ROSATPSPC observations of the outer regions of the Perseus cluster of galaxies. MNRAS, 300:837–856, Nov. 1998.

A. C. Fabian. Cooling Flows in Clusters of Galaxies. ARA&A, 32:277–318, 1994.

A. C. Fabian and P. E. J. Nulsen. Subsonic accretion of cooling gas in clusters of galaxies. MNRAS, 180:479–484, Aug. 1977.

A. C. Fabian, J. S. Sanders, S. W. Allen, C. S. Crawford, K. Iwasawa, R. M. Johnstone, R. W. Schmidt, and G. B. Taylor. A deep Chandra observation of the Perseus cluster: shocks and ripples. MNRAS, 344:L43–L47, Sept. 2003a.

A. C. Fabian, J. S. Sanders, C. S. Crawford, C. J. Conselice, J. S. Gallagher, and R. F. G. Wyse. The relationship between the optical Hα filaments and the X-ray emission in the core of the Perseus cluster. MNRAS, 344:L48–L52, Sept. 2003b.

G. J. Ferland, K. T. Korista, D. A. Verner, J. W. Ferguson, J. B. Kingdon, and E. M. Verner. CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra. PASP, 110: 761–778, July 1998.

W. Forman, P. Nulsen, S. Heinz, F. Owen, J. Eilek, A. Vikhlinin, M. Markevitch, R. Kraft, E. Churazov, and C. Jones. Reflections of AGN Outbursts in the Gaseous Atmosphere of M87. *ArXiv Astrophysics e-prints*, Dec. 2003.

B. Fryxell, K. Olson, P. Ricker, F. X. Timmes, M. Zingale, D. Q. Lamb, P. MacNeice, R. Rosner, J. W. Truran, and H. Tufo. FLASH: An Adaptive Mesh Hydrodynamics

Code for Modeling Astrophysical Thermonuclear Flashes. ApJS, 131:273–334, Nov. 2000.

P. L. Gómez, C. Loken, K. Roettiger, and J. O. Burns. Do Cooling Flows Survive Cluster Mergers? ApJ, 569:122–133, Apr. 2002.

A. Gardini, E. Rasia, P. Mazzotta, G. Tormen, S. De Grandi, and L. Moscardini. Simulating Chandra observations of galaxy clusters. MNRAS, 351:505–514, June 2004.

F. Haardt and P. Madau. Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background. ApJ, 461:20–+, Apr. 1996.

F. Haardt and P. Madau. Modelling the UV/X-ray cosmic background with CUBA. In *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, 2001.

J. C. Helly, S. Cole, C. S. Frenk, C. M. Baugh, A. Benson, C. Lacey, and F. R. Pearce. A comparison of gas dynamics in smooth particle hydrodynamics and semi-analytic models of galaxy formation. MNRAS, 338:913–925, Feb. 2003.

R. M. Johnstone, A. C. Fabian, and P. E. J. Nulsen. The optical spectra of central galaxies in southern clusters Evidence for star formation. MNRAS, 224:75–91, Jan. 1987.

J. S. Kaastra. SRON, Leiden report. 1992.

J. S. Kaastra, C. Ferrigno, T. Tamura, F. B. S. Paerels, J. R. Peterson, and J. P. D. Mittaz. XMM-Newton observations of the cluster of galaxies Sérsic 159-03. A&A, 365:L99–L103, Jan. 2001.

C. R. Kaiser, G. Pavlovski, E. C. D. Pope, and H. Fangohr. The stability of buoyant bubbles in the atmospheres of galaxy clusters. MNRAS, 359:493–503, May 2005.

N. Kaiser. Evolution of clusters of galaxies. ApJ, 383:104–111, Dec. 1991.

G. Kauffmann, S. D. M. White, and B. Guiderdoni. The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. MNRAS, 264:201–+, Sept. 1993.

D. Kawata and B. K. Gibson. Self-regulated active galactic nuclei heating in elliptical galaxies. MNRAS, 358:L16–L20, Mar. 2005.

A. N. Kolmogorov. *First Course in Turbulence.* C. R. Acad. Sci. URSS, 30, 301, 1941.

E. Komatsu and U. Seljak. Universal gas density and temperature profile. MNRAS, 327: 1353–1366, Nov. 2001.

M. J. Ledlow and F. N. Owen. 20 CM VLA Survey of Abell Clusters of Galaxies. VI. Radio/Optical Luminosity Functions. AJ, 112:9–+, July 1996.

R. Löhner. An adaptive finite element scheme for transient problems in CFD. *Comp. Meth. Appl. Mech. Eng.*, 61:323–338, Apr. 1987.

J. Magorrian, S. Tremaine, D. Richstone, R. Bender, G. Bower, A. Dressler, S. M. Faber, K. Gebhardt, R. Green, C. Grillmair, J. Kormendy, and T. Lauer. The Demography of Massive Dark Objects in Galaxy Centers. AJ, 115:2285–2305, June 1998.

M. Markevitch. The L X-T Relation and Temperature Function for Nearby Clusters Revisited. ApJ, 504:27–+, Sept. 1998.

M. Markevitch, T. J. Ponman, P. E. J. Nulsen, M. W. Bautz, D. J. Burke, L. P. David, D. Davis, R. H. Donnelly, W. R. Forman, C. Jones, J. Kaastra, E. Kellogg, D.-W. Kim, J. Kolodziejczak, P. Mazzotta, A. Pagliaro, S. Patel, L. Van Speybroeck, A. Vikhlinin, J. Vrtilek, M. Wise, and P. Zhao. Chandra Observation of Abell 2142: Survival of Dense Subcluster Cores in a Merger. ApJ, 541:542–549, Oct. 2000.

H. Mathis, G. Lavaux, J. M. Diego, and J. Silk. On the formation of cold fronts in massive mergers. MNRAS, 357:801–818, Mar. 2005.

P. Mazzotta, M. Markevitch, A. Vikhlinin, and W. R. Forman. Chandra Observations of Cold Fronts in Clusters of Galaxies. In *ASP Conf. Ser. 257: AMiBA 2001: High-Z Clusters, Missing Baryons, and CMB Polarization*, pages 173–+, 2002.

P. Mazzotta, E. Rasia, L. Moscardini, and G. Tormen. Comparing the temperatures of galaxy clusters from hydrodynamical N-body simulations to Chandra and XMM-Newton observations. MNRAS, 354:10–24, Oct. 2004.

I. G. McCarthy, M. L. Balogh, A. Babul, G. B. Poole, and D. J. Horner. Models of the Intracluster Medium with Heating and Cooling: Explaining the Global and Structural X-Ray Properties of Clusters. ApJ, 613:811–830, Oct. 2004.

B. R. McNamara, M. Wise, P. E. J. Nulsen, L. P. David, C. L. Sarazin, M. Bautz, M. Markevitch, A. Vikhlinin, W. R. Forman, C. Jones, and D. E. Harris. Chandra X-Ray Observations of the Hydra A Cluster: An Interaction between the Radio Source and the X-Ray-emitting Gas. ApJ, 534:L135–L138, May 2000.

B. R. McNamara, M. W. Wise, P. E. J. Nulsen, L. P. David, C. L. Carilli, C. L. Sarazin, C. P. O'Dea, J. Houck, M. Donahue, S. Baum, M. Voit, R. W. O'Connell, and A. Koekemoer. Discovery of Ghost Cavities in the X-Ray Atmosphere of Abell 2597. ApJ, 562:L149–L152, Dec. 2001.

B. R. McNamara, P. E. J. Nulsen, M. W. Wise, D. A. Rafferty, C. Carilli, C. L. Sarazin, and E. L. Blanton. The heating of gas in a galaxy cluster by X-ray cavities and large-scale shock fronts. Nature, 433:45–47, Jan. 2005.

R. Mewe, J. R. Lemen, and G. H. J. van den Oord. Calculated X-radiation from optically thin plasmas. VI - Improved calculations for continuum emission and approximation formulae for nonrelativistic average Gaunt factors. A&AS, 65:511–536, Sept. 1986.

J. J. Monaghan. Smoothed particle hydrodynamics. ARA&A, 30:543–574, 1992.

O. Muanwong, P. A. Thomas, S. T. Kay, and F. R. Pearce. The effect of cooling and preheating on the X-ray properties of clusters of galaxies. MNRAS, 336:527–540, Oct. 2002.

J. F. Navarro, C. S. Frenk, and S. D. M. White. Simulations of X-ray clusters. MNRAS, 275:720–740, Aug. 1995.

J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. ApJ, 462:563–+, May 1996.

J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. ApJ, 490:493–+, Dec. 1997.

P. E. J. Nulsen, B. R. McNamara, M. W. Wise, and L. P. David. The cluster-scale AGN outburst in Hydra A. *ArXiv Astrophysics e-prints*, Aug. 2004.

A. Nusser, J. Silk, and A. Babul. Suppressing cluster cooling flows by multiple AGN activity. *ArXiv Astrophysics e-prints*, Feb. 2006.

R. W. O'Connell and B. R. McNamara. The Fate of Matter in Cooling Flows. AJ, 98: 180–+, July 1989.

H. Omma, J. Binney, G. Bryan, and A. Slyz. Heating cooling flows with jets. MNRAS, 348:1105–1119, Mar. 2004.

F. N. Owen, J. A. Eilek, and W. C. Keel. Detection of large Faraday rotation in the inner 2 kiloparsecs of M87. ApJ, 362:449–454, Oct. 1990.

F. N. Owen, J. A. Eilek, and N. E. Kassim. M87 at 90 Centimeters: A Different Picture. ApJ, 543:611–619, Nov. 2000.

J. R. Peterson, F. B. S. Paerels, J. S. Kaastra, M. Arnaud, T. H. Reiprich, A. C. Fabian, R. F. Mushotzky, J. G. Jernigan, and I. Sakelliou. X-ray imaging-spectroscopy of Abell 1835. A&A, 365:L104–L109, Jan. 2001.

J. R. Peterson, S. M. Kahn, F. B. S. Paerels, J. S. Kaastra, T. Tamura, J. A. M. Bleeker, C. Ferrigno, and J. G. Jernigan. High-Resolution X-Ray Spectroscopic Constraints on Cooling-Flow Models for Clusters of Galaxies. ApJ, 590:207–224, June 2003.

E. Pointecouteau, M. Arnaud, J. Kaastra, and J. de Plaa. XMM-Newton observation of the relaxed cluster A478: Gas and dark matter distribution from $0.01R_{200}$ to $0.5R_{200}$. A&A, 423:33–47, Aug. 2004.

T. J. Ponman, D. B. Cannon, and J. F. Navarro. The thermal imprint of galaxy formation on X-ray clusters. Nature, 397:135–137, 1999.

T. J. Ponman, A. J. R. Sanderson, and A. Finoguenov. The Birmingham-CfA cluster scaling project - III. Entropy and similarity in galaxy systems. MNRAS, 343:331–342, July 2003.

V. Quilis, R. G. Bower, and M. L. Balogh. Bubbles, feedback and the intracluster medium: three-dimensional hydrodynamic simulations. MNRAS, 328:1091–1097, Dec. 2001.

E. Rasia, P. Mazzotta, S. Borgani, L. Moscardini, K. Dolag, G. Tormen, A. Diaferio, and G. Murante. Mismatch between X-Ray and Emission-weighted Temperatures in Galaxy Clusters: Cosmological Implications. ApJ, 618:L1–L4, Jan. 2005.

T. H. Reiprich and H. Böhringer. The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters. ApJ, 567:716–740, Mar. 2002.

C. S. Reynolds, S. Heinz, and M. C. Begelman. The hydrodynamics of dead radio galaxies. MNRAS, 332:271–282, May 2002.

S. Roychowdhury, M. Ruszkowski, B. B. Nath, and M. C. Begelman. Entropy "Floor" and Effervescent Heating of Intracluster Gas. ApJ, 615:681–688, Nov. 2004.

M. Ruszkowski and M. C. Begelman. Heating, Conduction, and Minimum Temperatures in Cooling Flows. ApJ, 581:223–228, Dec. 2002.

M. Ruszkowski, M. Brüggen, and M. C. Begelman. Cluster Heating by Viscous Dissipation of Sound Waves. ApJ, 611:158–163, Aug. 2004a.

M. Ruszkowski, M. Brüggen, and M. C. Begelman. Three-Dimensional Simulations of Viscous Dissipation in the Intracluster Medium. ApJ, 615:675–680, Nov. 2004b.

P. Salomé and F. Combes. Cold molecular gas in cooling flow clusters of galaxies. A&A, 412:657–667, Dec. 2003.

J. S. Sanders, A. C. Fabian, and R. J. H. Dunn. Non-thermal X-rays, a high abundance ridge and fossil bubbles in the core of the Perseus cluster of galaxies. *ArXiv Astrophysics e-prints*, Mar. 2005.

A. J. R. Sanderson, A. Finoguenov, and J. J. Mohr. Possible AGN Shock Heating in the Cool Core Galaxy Cluster Abell 478. *ArXiv Astrophysics e-prints*, Dec. 2004.

C. L. Sarazin. *X-ray emission from clusters of galaxies*. Cambridge Astrophysics Series, Cambridge: Cambridge University Press, 1988, 1988.

R. W. Schmidt, A. C. Fabian, and J. S. Sanders. Chandra temperature and metallicity maps of the Perseus cluster core. MNRAS, 337:71–78, Nov. 2002.

N. Soker, E. L. Blanton, and C. L. Sarazin. Hot Bubbles in Cooling Flow Clusters. ApJ, 573:533–541, July 2002.

D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. ApJS, 148: 175–194, Sept. 2003. doi: 10.1086/377226.

J. Spitzer. *Physicl of Fully Ionized Gases*. New York: Interscience, 1962.

V. Springel. The cosmological simulation code GADGET-2. *ArXiv Astrophysics e-prints*, May 2005.

V. Springel, N. Yoshida, and S. D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Astronomy*, 6:79–117, Apr. 2001.

V. Springel, T. Di Matteo, and L. Hernquist. Black Holes in Galaxy Mergers: The Formation of Red Elliptical Galaxies. ApJ, 620:L79–L82, Feb. 2005.

G. Tabor and J. Binney. Elliptical Galaxy Cooling Flows Without Mass Drop-Out. MNRAS, 263:323–+, July 1993.

T. Tamura, J. S. Kaastra, J. R. Peterson, F. B. S. Paerels, J. P. D. Mittaz, S. P. Trudolyubov, G. Stewart, A. C. Fabian, R. F. Mushotzky, D. H. Lumb, and Y. Ikebe. X-ray spectroscopy of the cluster of galaxies Abell 1795 with XMM-Newton. A&A, 365:L87–L92, Jan. 2001.

H. Tennekes and J. L. Lumley. *First Course in Turbulence*. First Course in Turbulence, Cambridge: MIT Press, 1972, 1972.

T. Theuns, J. Schaye, S. Zaroubi, T. Kim, P. Tzanavaris, and B. Carswell. Constraints on Reionization from the Thermal History of the Intergalactic Medium. ApJ, 567:L103–L106, Mar. 2002.

L. M. Voigt and A. C. Fabian. Thermal conduction and reduced cooling flows in galaxy clusters. MNRAS, 347:1130–1149, Feb. 2004.

G. M. Voit and T. J. Ponman. Signatures of Galaxy Formation in the Intracluster Medium. ApJ, 594:L75–L78, Sept. 2003.

G. M. Voit, G. L. Bryan, M. L. Balogh, and R. G. Bower. Modified Entropy Models for the Intracluster Medium. ApJ, 576:601–624, Sept. 2002.

G. M. Voit, M. L. Balogh, R. G. Bower, C. G. Lacey, and G. L. Bryan. On the Origin of Intracluster Entropy. ApJ, 593:272–290, Aug. 2003.

R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov, and A. Dekel. Concentrations of Dark Halos from Their Assembly Histories. ApJ, 568:52–70, Mar. 2002.

S. D. M. White and C. S. Frenk. Galaxy formation through hierarchical clustering. ApJ, 379:52–79, Sept. 1991. doi: 10.1086/170483.

C. J. Willott, S. Rawlings, K. M. Blundell, and M. Lacy. The emission line-radio correlation for radio sources using the 7C Redshift Survey. MNRAS, 309:1017–1033, Nov. 1999.

P. R. Woodward and P. Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comp. Phys.*, 54:115–173, Apr. 1984.

K. K. S. Wu, A. C. Fabian, and P. E. J. Nulsen. Non-gravitational heating in the hierarchical formation of X-ray clusters. MNRAS, 318:889–912, Nov. 2000.