A modified cooling algorithm for semi-analytic modelling

Benoit Fournier[1], Peter A. Thomas[1★], Bruno M. Henriques[2]
[1]Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK
[2]Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching b. München, Germany

Draft - 7 February 2019

ABSTRACT

The cooling algorithm used by many existing semi-analytic models is rather simplisitic and does not directly predict the lumnosity of the intra-halo medium. We eliminate these deficiencies by modifying the density profile of the hot gas to match that of observed halos, and by eliminating the dependence upon timestep.

Our results . . .

Key words: methods: numerical – galaxies: formation – galaxies: evolution

EDITING LOCKS

Please add your name against each file if you are editing it (to prevent conflicts). If a file is locked contact the owner to seek permission to edit.

- paper.tex:
- method.tex:
- results.tex:
- conc.tex:
- ihm.bib: (no need to lock if just adding extra entries)

1 INTRODUCTION

2 METHOD

In this section we first describe the existing L-Galaxies cooling algorithm and its deficiencies. We then present our new method before comparing the two for different ratios of cooling to dynamical time.

2.1 The existing cooling algorithm

The cooling model that is used in the current version of L-Galaxies is described in detail in the supplementary material that accompanies Henriques et al. (2014). We reiterate the essential points here.

2.1.1 *The singular isothermal sphere*

For simplicity, the density profile of the gas, shown in the top panel of Fig. 1, is assumed to be that of a singular isothermal sphere (hereafter, SIS),

$$\rho_g = f_g \frac{200\rho_c}{3}\frac{1}{x^2}, \tag{1}$$

where $f_g = M_g/M_{200c}$ and $x = r/r_{200c}$. Here M_g is the gas mass and M_{200c} the total mass within a radius r_{200c} that encloses an average density of 200 times the critical density ρ_c. Of course, it is recognised that having an infinite density at the centre of the halo is not an accurate representation of real halos, but in practice, this is expected to make little difference to the cooling rates as calculated below.

In an SIS then the gas mass contained within radius r is

$$m_g = f_g M_{200c} x. \tag{2}$$

The specific energy of the gas is constant and equal to

$$\epsilon = \frac{3k_BT}{2\mu m_H}, \tag{3}$$

where k_B is the Boltzmann constant, T is the gas temperature and μm_H is the mean mass per particle in the gas. The equation of hydrostatic support leads quickly to the relation

$$\frac{k_BT}{\mu m_H} = \frac{1}{2}{v_c}^2 = \frac{GM_{200c}}{2r_{200c}} = \frac{400\pi}{3}G\rho_c {r_{200c}}^2, \tag{4}$$

where v_c is the circular speed of the halo.

2.1.2 *The cooling and dynamical times*

We take the cooling rate per unit volume of the gas to be ${n_g}^2\Lambda(T,Z)$, where $n_g = \rho_g/\mu m_H$ is the number density of particles in the gas and the cooling function, Λ, is a function of temperature and metallicity, Z. The specific cooling rate of the gas is then

$$\frac{d\epsilon}{dt} = -\frac{n_g}{\mu m_H}\Lambda, \tag{5}$$

★ E-mail: P.A.Thomas@sussex.ac.uk

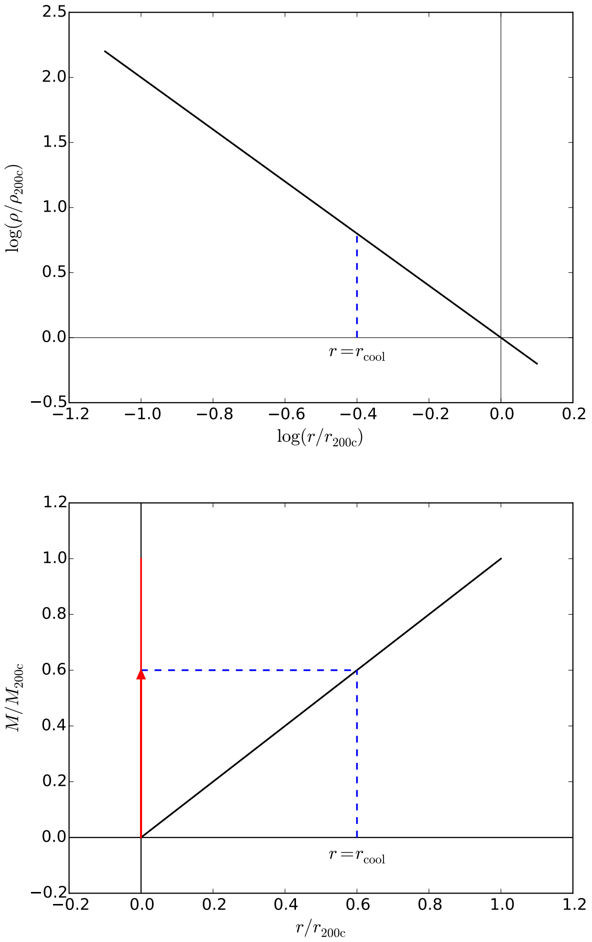

Figure 1. The singular isothermal sphere: upper panel - density profile; lower panel: mass profile. The cooling radius, $r_{\rm cool}$, is the radius at which the cooling time is equal to the dynamical time of the halo.

which leads to a characteristic cooling timescale for the gas of

$$t_{\rm cool} \equiv \frac{\epsilon}{|{\rm d}\epsilon/{\rm d}t|} = \frac{3k_{\rm B}T}{2n_{\rm g}\Lambda}. \tag{6}$$

We can write this as

$$t_{\rm cool} = \tau_{\rm cool}\frac{x^2}{f_{\rm g}}, \tag{7}$$

where

$$\tau_{\rm cool} = \frac{9\mu m_{\rm H}k_{\rm B}T}{400\rho_{\rm c}\Lambda}. \tag{8}$$

The adiabatic sound speed in the gas is $(5k_{\rm B}T/3\mu m_{\rm H})^{1/2} \approx v_{\rm c}$. Hence the characteristic dynamical time for gas to flow a distance r is

$$t_{\rm dyn} = \frac{r}{v_{\rm c}} = \tau_{\rm dyn}x, \tag{9}$$

where $\tau_{\rm dyn} = r_{\rm 200c}/v_{\rm c}$.

2.1.3 The cooling model

Consider first the case where $\tau_{\rm cool} \ll f_{\rm g}\tau_{\rm dyn}$. Then all the gas within a radius $r_{\rm 200c}$ can cool within a dynamical time. The latter limits the rate at which can flow into the halo to replace that which has cooled, however, and so the maximum (i.e. most negative) cooling rate is taken to be

$$\dot{M}_{\rm g,max} = -\frac{M_{\rm g}}{\tau_{\rm dyn}} = -\frac{M_{\rm 200c}}{\tau_{\rm dyn}}f_{\rm g}. \tag{10}$$

The key element of the algorithm is what to do when the cooling time for the halo is longer than the dynmaical time $\tau_{\rm cool} \gtrsim f_{\rm g}\tau_{\rm dyn}$. One might naively expect that the gas in the centre of the halo will be able to cool to low temperatures more effectively than that at larger radii, leading to a positive temperature gradient in the gas as a function of radius. That in turn would give a density increase at the halo centre (to maintain pressure support) and a central luminosity spike. However, X-ray observations of quasi-steady cooling in groups and clusters of galaxies (e.g. Ettori et al. 2013) show only mild temperature gradients so that, to a good approximation, $T(r) \approx$ constant, as in the SIS. Instead the gas is thought to have a multiphase structure (i.e. there is a mixture of densities at any given radius) leading to a multiphase cooling flow model (Thomas 1988). In such a model, the mean temperature of the gas remains fixed also in time and only the density changes as gas cools out of the flow.

In the existing L-Galaxies model, this is implemented by finding the radius, $r_{\rm cool}$, at which $t_{\rm cool}(r_{\rm cool}) = \tau_{\rm dyn}$ and setting the cooling rate to be

$$\dot{M}_{\rm g} = -\frac{m_{\rm g}(r_{\rm cool})}{\tau_{\rm dyn}} = \dot{M}_{\rm g,max}x_{\rm cool}, \tag{11}$$

where $x_{\rm cool} = r_{\rm cool}/r_{\rm 200c}$. Now, from Equation 8, we see that $x_{\rm cool} = (\tau_{\rm dyn}f_{\rm g}/\tau_{\rm cool})^{1/2}$. Hence the final expression for the cooling rate is

$$\dot{M}_{\rm g} = \dot{M}_{\rm g,max}\,\min\left[1, \left(\frac{\tau_{\rm dyn}f_{\rm g}}{\tau_{\rm cool}}\right)^{\frac{1}{2}}\right]. \tag{12}$$

Over timestep Δt the amount of gas that has cooled and been deposited is estimated assuming the cooling rate to be constant over the timestep, but restricting the total amount of gas cooled to be no larger than that initially present:

$$|\Delta M_{\rm g}| = \min\left[M_{\rm g}, |\dot{M}_{\rm g}|\Delta t\right]. \tag{13}$$

We can rewrite this in terms of the hot gas mass fraction. Starting with initial gas fraction $f_{\rm g0}$, the remaining gas fraction after time Δt is

$$f_{\rm g} = f_{\rm g0}\,\max\left[0, 1-\frac{\Delta t}{\tau_{\rm dyn}}, 1-\left(\frac{\tau_{\rm dyn}f_{\rm g0}}{\tau_{\rm cool}}\right)^{\frac{1}{2}}\frac{\Delta t}{\tau_{\rm dyn}}\right]. \tag{14}$$

From Equation 12 one can obtain an estimate of the bolometric luminosity of the hot gas (in practice mostly coming out in the X-ray band) of

$$L = \epsilon|\dot{M}_{\rm g}|. \tag{15}$$

2.1.4 Problems with the model

The first problem with the model is that the total amount of mass cooled is only correctly estimated from Equation 13 when the timestep is sufficiently small that $f_{\rm g} \approx f_{\rm g0}$. This is an unnecessary restriction introduced to minimise execution time in the code and can easily be replaced by integrated versions of the cooling equations that use a variable cooling rate as the density of the remaining gas decreases over time. Equation 12 can be rewritten as

$$\dot{f}_{\rm g} = -\begin{cases} \dfrac{f_{\rm g}}{\tau_{\rm dyn}}, & \tau_{\rm cool} \leq \tau_{\rm dyn} f_{\rm g}; \\ \dfrac{f_{\rm g}^{\frac{3}{2}}}{(\tau_{\rm dyn}\tau_{\rm cool})^{\frac{1}{2}}}, & \tau_{\rm cool} > \tau_{\rm dyn} f_{\rm g}. \end{cases} \tag{16}$$

When $\tau_{\rm cool} > \tau_{\rm dyn} f_{\rm g0}$ the lower of these expressions integrates to give

$$f_{\rm g} = f_{\rm g0}\left(1 + \left(\frac{\tau_{\rm dyn} f_{\rm g0}}{\tau_{\rm cool}}\right)^{\frac{1}{2}} \frac{\Delta t}{2\tau_{\rm dyn}}\right)^{-2}. \tag{17}$$

The condition $\tau_{\rm cool} \leq \tau_{\rm dyn} f_{\rm g}$ is trickier, because after some of the gas has cooled, the cooling time drops and the cooling rate switches from the upper to the lower expression:

$$f_{\rm g} = \begin{cases} f_{\rm g0} e^{-\Delta t/\tau_{\rm dyn}}, & \Delta t \leq t_{\rm eq}; \\ \dfrac{\tau_{\rm cool}}{\tau_{\rm dyn}}\left(1 + \dfrac{\Delta t - t_{\rm eq}}{2\tau_{\rm dyn}}\right)^{-2}, & \Delta t > t_{\rm eq}; \end{cases} \tag{18}$$

where $t_{\rm eq} = \tau_{\rm dyn} \ln(\tau_{\rm dyn} f_{\rm g0}/\tau_{\rm cool})$.

Fig. 2 shows the difference between the two approaches for the cases where $\tau_{\rm cool}/\tau_{\rm dyn} f_{\rm g} = 0.2$ and 2. The amount of gas cooled is the same for timesteps that are short compared to the dynamical time but, whereas the fixed cooling rate scheme exhausts all the gas after one dynamical time, the variable cooling rate approach leaves residual gas even at late times. For this point onwards in the paper, we will use a variable cooling rate.

We note in passing that it could be argued that the cooling radius should be determined not by the location where the cooling time of the gas equals the dynamical time of the halo, $\tau_{\rm dyn}$, but rather the local dynamical time, $t_{\rm dyn}$. That would have the effect of modifying Equation 18 as follows:

$$f_{\rm g} = \begin{cases} f_{\rm g0} e^{-\Delta t/\tau_{\rm dyn}}, & \Delta t \leq t_{\rm eq}; \\ \dfrac{\tau_{\rm cool}}{\tau_{\rm dyn}}\left(1 + \dfrac{\Delta t - t_{\rm eq}}{3\tau_{\rm dyn}}\right)^{-3}, & \Delta t > t_{\rm eq}; \end{cases} \tag{19}$$

In practice, the two are very similar.

A more serious object to the cooling model is that it seems to depend only upon the behaviour of the gas within the cooling radius: why should gas outside that radius not contribute? A direct integration of the expected luminosity in an SIS profile leads to an infinite result, hence the use of Equation 15 which again depends only upon the emission wihtin the cooling radius. In the following section we develop a new version of the model that does not have this deficiency.

2.2 The new cooling algorithm

Fundamental to our new cooling algorithm is the use of a density profile for the intrahalo gas that more closely mimics that of observed clusters. In particular, we use a model that has a finite central density. That then gives a finite luminosity and we derive the rate of mass depostion from this luminosity, rather than the other way around.

2.2.1 The isothermal-beta model

We again assume that the total matter content of the halo can be approximated by an SIS. However, we adopt a density profile for the gas that tends to a constant within core radius a:

$$\rho_{\rm g} = \frac{\rho_0}{1+y^2}, \tag{20}$$

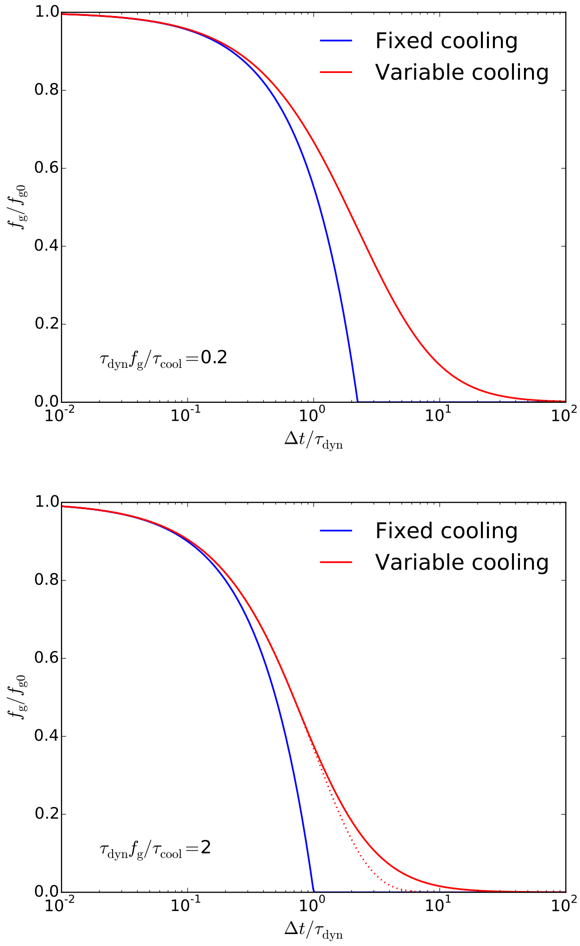

Figure 2. The amount of gas remaining in the SIS model after time Δt using a fixed cooling rate (lower, blue curve) and variable cooling rate (upper, red curve), for the cases where $\tau_{\rm cool}/\tau_{\rm dyn} f_{\rm g} = 0.2$ (upper panel) and 2 (lower panel). The dotted red line in the lower panel shows the exponential branch of the solution given in Equation 18.

where $y = r/a$. We call this the isothermal-beta model as it is derived from a class of density profiles $\rho_{\rm g} \propto (1+y^2)^{-\beta}$, with $\beta = \frac{2}{3}$, giving the usual $1/r^2$ dependence of the isothermal sphere at large radii, $r \gg a$.

The mass profile of the gas is

$$m_{\rm g} = 4\pi\rho_0 a^3(y - \arctan y) = M_{\rm g}\frac{y - \arctan y}{Y - \arctan Y}, \tag{21}$$

where $Y = y(r_{\rm 200c})$. Hence the gas fraction within $r_{\rm 200c}$ is

$$f_{\rm g} = \frac{4\pi\rho_0 a^3(Y - \arctan Y)}{M_{\rm 200c}}. \tag{22}$$

We again assume an iosthermal temperature for the gas. This is not strictly valid because the core gas would not be in hydrostatic equilibrium within an SIS potential at this temperature. However, we know that the background potential does not itself truly follow an SIS and, as the core is small compared to the virial radius, this seems a reasonable approximation to adopt.

The luminosity of the halo gas (integrating now out to infinite radius) is

$$L = \int_0^\infty {n_\mathrm{g}}^2 \Lambda\, 4\pi r^2\, \mathrm{d}r = \frac{\rho_0^2 \Lambda\, \pi^2 a^3}{(\mu m_\mathrm{H})^2}. \tag{23}$$

2.2.2 *The cooling model*

In a quasi-static cooling flow, in which the gas cools at constant pressure, then the mass-deposition rate is (Thomas 1988)

$$\dot{M}_\mathrm{g} = \frac{L}{H}, \tag{24}$$

where $H = 5k_\mathrm{B}T/2\mu m_\mathrm{H}$ is the specific enthaply of the gas. The reason for using enthalpy instead of energy is that the parcels of cooling gas remain in pressure equilibrium with their surrondings and therefore have pressure work done on them as they cool.

Combining Equation 24 with Equations 8, 22 and 23 gives

$$\dot{f}_\mathrm{g} = -\frac{{f_\mathrm{g}}^2}{{\tau_\mathrm{cool}}'}, \tag{25}$$

where

$$\begin{aligned} {\tau_\mathrm{cool}}' &= \frac{3\mu m_\mathrm{H} k_\mathrm{B} T}{20\pi\rho_\mathrm{c}\Lambda}\frac{(Y-\arctan Y)^2}{Y^3} && (26)\\ &= \frac{20}{3\pi}\frac{(Y-\arctan Y)^2}{Y^3}\tau_\mathrm{cool}. && (27)\end{aligned}$$

As for the SIS, the quasi-static cooling flow assumptions break down when the cooling time at the edge of the halo is less than the dynamical time. We therefore impose the same limit on the cooling rate, meaning that the expression for the gas fraction again splits into two cases. For ${\tau_\mathrm{cool}}' \geq \tau_\mathrm{dyn} f_\mathrm{g0}$,

$$f_\mathrm{g} = f_\mathrm{g0}\left(1 + \frac{f_\mathrm{g0}\Delta t}{{\tau_\mathrm{cool}}'}\right)^{-1}; \tag{28}$$

whereas for ${\tau_\mathrm{cool}}' < \tau_\mathrm{dyn} f_\mathrm{g0}$,

$$f_\mathrm{g} = \begin{cases} f_\mathrm{g0} e^{-\Delta t/\tau_\mathrm{dyn}}, & \Delta t \leq t_\mathrm{eq};\\ \frac{{\tau_\mathrm{cool}}'}{\tau_\mathrm{dyn}}\left(1 + \frac{\Delta t - t_\mathrm{eq}}{\tau_\mathrm{dyn}}\right)^{-1}, & \Delta t > t_\mathrm{eq}; \end{cases} \tag{29}$$

where $t_\mathrm{eq} = \tau_\mathrm{dyn}\ln(\tau_\mathrm{dyn} f_\mathrm{g0}/{\tau_\mathrm{cool}}')$. For typical values of $Y \approx 10$ *****Need to show this somewhere*****, this gives cooling times that are much shorter in the existing model, ${\tau_\mathrm{cool}}' \approx 0.15\tau_\mathrm{cool}$.

There is one important distinction between Equations **??** and 28: in the latter, the fraction of gas cooled is independent of the dynamical time of the halo (as it should be). This can potentially lead to very different estimates of the rate of deposition of hot gas in the cooling flow regime. Fig. 3 shows the amount of gas cooled using the existing model (solid, red curves), the existing model modified to use a variable cooling rate (dotted, red curves) and for the new model (solid, blue curves), using a value of $Y = 10$. We see that the amount of cooled gas is less than in the existing SIS model both in the limit of short and long cooling times. However, when $\tau_\mathrm{dyn} f_\mathrm{g} \approx {\tau_\mathrm{cool}}'$) then the cooling rate has approximately doubled leading to greater deposition of gas for short tiemsteps.

2.3 Summary

We have shown that the existing L-Galaxies cooling model, based on a singular isothermal sphere model for the hot gas density profile, is deficient in two ways: (i) for large time-steps, $\Delta t \gg \max(\tau_\mathrm{dyn}, (\tau_\mathrm{dyn}\tau_\mathrm{cool}/f_\mathrm{g})^{1/2})$, it cools too much gas, and (ii)

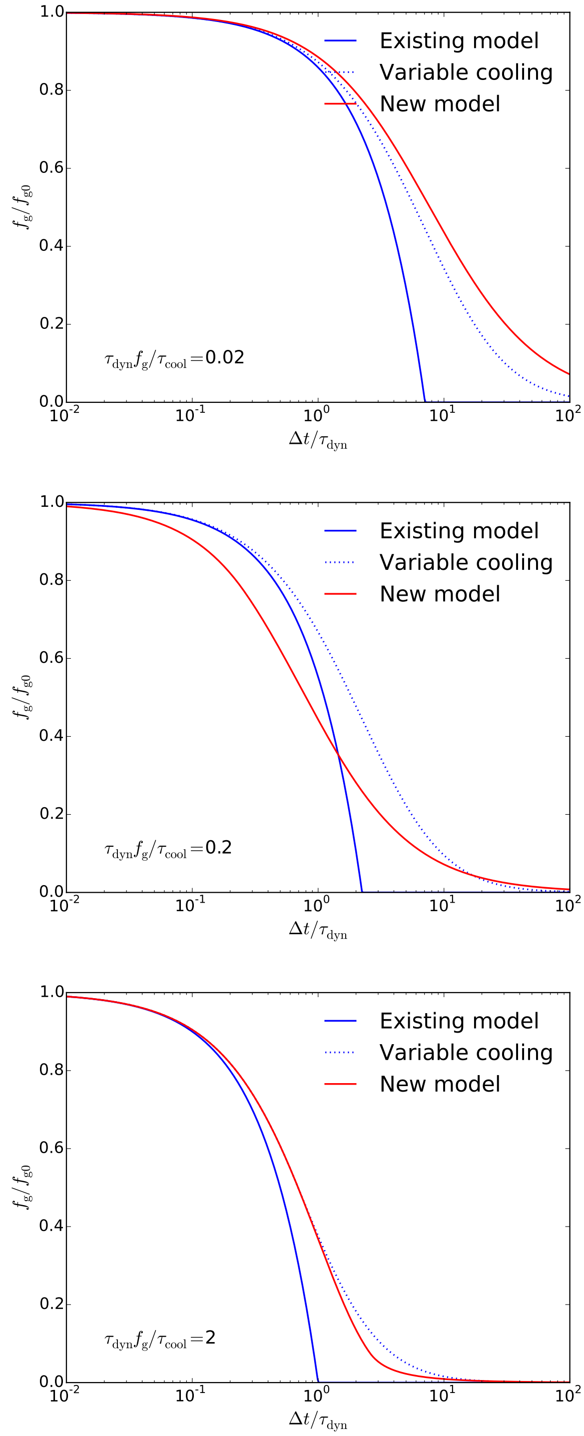

Figure 3. The amount of gas remaining as a function of time using the existing SIS model with a fixed cooling rate (solid, blue curve) and variable cooling rate (dotted, blue curve), and for the new isothermal-beta model model (solid, red curve), for the cases where $\tau_\mathrm{cool}/\tau_\mathrm{dyn} f_\mathrm{g} = 0.02$ (upper panel), 0.2 (middle panel) and 2 (lower panel). We have taken $Y = 10$.

the cooling rate in the slow-cooling regime $\tau_{\rm cool} > \tau_{\rm dyn} f_{\rm g}$) is dependent upon the dynamical time of the halo.

We overcome these problems by modifying the density profile of the hot gas to be an isothermal-beta model which has a finite centra density. The cooling rate of the gas is then self-consistently calculated and used to determine the rate of mass deposition in the cooling flow regime.

The new model differs in two important respects from the old:

- The hot gas is never fully depleted, no matter how large a timestep is used (and, unlike the exisiting sheme, using many small timesteps to cover a fixed time interval gives the same result as using one larger one).
- When the cooling time is very long (as in groups and clusters of galaxies) then the absence of a central density cusp means that the rate of mass deposition is much smaller in the new model than in the old one.

The first of these items is largely a numerical issue, but the latter can have implications for the amoutn of feedback required to suppress star-formation in the largest galaxies, as we will show below.

3 RESULTS

4 CONCLUSIONS

ACKNOWLEDGEMENTS

The authors contributed in the following way to this paper. BF undertook the vast majority of the coding, data analysis and drafting of the paper and figures. PAT supervised BF, devised the new cooling method and wrote that portion of the paper. BH aided with the interpretation of the results. All three authors helped to proof-read the text.

PAT acknowledges support from the Science and Technology Facilities Council (grant number ST/L000652/1).

REFERENCES

Ettori S., Donnarumma A., Pointecouteau E., Reiprich T., Giodini S., Lovisari L., Schmidt R., 2013, SSRv, 177, 119

Henriques B., White S., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2014, MNRAS, accepted, astro-ph/1410.0365

Thomas P., 1988, MNRAS, 235, 315