Chaotic cold accretion in giant elliptical galaxies heated by AGN cosmic rays

Chaoran Wang,1⋆ Mateusz Ruszkowski,1† H.-Y. Karen Yang2‡
1 Department of Astronomy, University of Michigan, 1085 S. University Avenue, 311 West Hall, Ann Arbor, MI 48109, USA
2 Department of Astronomy, University of Maryland, College Park, MD 20742, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Black hole feedback plays a central role in shaping the circumgalactic medium (CGM) of elliptical galaxies. We systematically study the impact of plasma physics on the evolution of ellipticals by performing three-dimensional non-ideal magneto-hydrodynamic simulations of the interactions of active galactic nucleus (AGN) jets with the CGM including magnetic fields, and cosmic rays (CRs) and their transport processes. We find that the physics of feedback operating on large galactic scales depends very sensitively on plasma physics operating on small scales. Specifically, we demonstrate that: (i) in the purely hydrodynamical case, the AGN jets initially maintain the atmospheres in global thermal balance. However, local thermal instability generically leads to the formation of massive cold disks in the vicinity of the central black hole in disagreement with observations; (ii) including weak magnetic fields prevents the formation of the disks because local B-field amplification in the precipitating cold gas leads to strong magnetic breaking, which quickly extracts angular momentum from the accreting clouds. The magnetic fields transform the cold clouds into narrow filaments that do not fall ballistically; (iii) when plasma composition in the AGN jets is dominated by CRs, and CR transport is neglected, the atmospheres exhibit cooling catastrophes due to inefficient heat transfer from the AGN to CGM despite Coulomb/hadronic CR losses being present; (iv) including CR streaming and heating restores agreement with the observations, i.e., cooling catastrophes are prevented and massive cold central disks do not form. The AGN power is reduced as its energy is utilized efficiently.

Key words: galaxies: magnetic fields – galaxies: evolution – galaxies: elliptical and lenticular, cD – galaxies: active

1 INTRODUCTION
In the centers of galaxy clusters, groups, and giant elliptical galaxies, radiative cooling time of the hot gaseous halos can be much shorter than the Hubble time. When this is the case, and in the absence of heating sources, a classic cooling flow will form (Fabian 1994) leading to high star formation rates (SFRs). However, the observed SFRs are at least one order of magnitude lower than that predicted by the pure cooling flow (e.g., O’Dea et al. 2008). It has been widely accepted that the feedback from active galactic nuclei (AGN) is the most plausible solution to the cooling flow problem. The energetic AGN feedback is able to maintain the overall thermal equilibrium of hot gaseous halos and prevents excessive gas cooling and star formation in galaxy clusters, groups, and giant elliptical galaxies (see McNamara & Nulsen 2007, for a review). Most AGN found in the centers of these massive systems at low redshift are associated with bipolar jet outflows launched by the central supermassive black holes (SMBHs). The AGN jet inflates radio-emitting bubbles that generate X-ray cavities in the hot halo. How exactly the jets and bubbles transfer energy to the hot halo remains unclear. Various possibilities have been proposed and extensively studied in the literature, including: mixing of jet-inflated bubble filled with ultra-hot gas (e.g., Brunetti & Kaiser 2002; Gilkis & Soker 2012; Hillel & Soker 2014, 2016, 2017; Yang & Reynolds 2016); dissipation of jet-induced sound waves and weak shocks (e.g., Fabian et al. 2003; Ruszkowski et al. 2004a,b; Fabian et al. 2005, 2006, 2017; Sanders & Fabian 2008; Zweibel et al. 2018; Fabian et al. 2003, 2005, 2017; Forman et al. 2007; Randall et al. 2015; Barai et al. 2016; Yang & Reynolds 2016; Li et al. 2017; Martizzi et al. 2019), and turbulent dissipation of gas motions (e.g., De Young 2010; Zhuravleva et al. 2014); and heating by cosmic rays...
(CRs) escaping from the jet-inflated bubbles (e.g., Guo & Oh 2008; Enßlin et al. 2011; Ruszkowski et al. 2008; Wiener et al. 2013b; Fujita & Ohira 2013; Fujita et al. 2013; Jacob & Pfrommer 2017a,b; Ruszkowski et al. 2017; Ehler et al. 2018). However, no consensus has been reached and the dominant mechanism(s) is(are) uncertain.

Despite adopting various numerical methods, simulations of AGN feedback and cooling flow problem often reveal formation of a massive long-lived cold disk in the vicinity of the central SMBH (e.g., Li & Bryan 2014a; Eisenreich et al. 2017; Qiu et al. 2018; Wang et al. 2019). However, the observed molecular gas rarely appears to be rotationally supported in either elliptical galaxies (Young et al. 2019) or central cluster galaxies (Russell et al. 2019). Moreover, these simulations produce overly massive disks compared to the few disks found in the observations. The formation of such massive disks poses further problems for the feeding of the central SMBH and AGN feedback. In our previous work (Wang et al. 2019), the fueling of large amounts of cold gas from the disk breaks the self-regularity of the AGN feedback, thus causing the central region of the halo to overheat. This results in an isentropic core, which is inconsistent with the observations.

Radio observations show that the jet-inflated bubbles emit synchrotron radiation, indicating the presence of relativistic electrons. However, it has been shown that the emitting electrons plus magnetic field have less pressure than required to maintain pressure balance between the bubble interior and the ambient hot halo gas (Dunn & Fabian 2004). Therefore, the non-radiating component has to be present inside the bubbles to provide the required pressure, e.g., ultra-hot gas, magnetic fields, and CR protons. However, the relative contribution of each of these non-thermal components is not well constrained by the observations.

CR-dominated AGN jet can potentially solve the disk problem mentioned above. CRs heat the hot halo gas via Coulomb, hadronic interactions (Yoast-Hull et al. 2013), and streaming instability (Zweibel 2013, for a review) as they are transported out from the jet-inflated bubbles. Streaming instability allows CRs to scatter on the self-excited Alfvén waves and to stream down their pressure gradient along the magnetic field at or above the Alfvén speed. CR streaming down the pressure gradient along the B-field results in CRs doing work on the gas and effectively heating it. This additional mode of heating could help to heat the gas at larger distances from the center and eliminate catastrophic cooling and cold disk formation at late times. However, recent simulations of a single short-duration jet in the galaxy cluster (Yang et al. 2019) demonstrate that CR-dominated jets can efficiently uplift the hot halo due to high buoyancy of the CR fluid filling the bubbles and larger cross-sections of the raising bubbles. This could remove gas with the shortest cooling time near the cluster center and potentially limit the formation of the disk in the long term. Interestingly, while more efficient heating in long term is expected, Yang et al. (2019) show that the CR-dominated jet makes the intracluster medium (ICM) more prone to the development of thermal instabilities in the center due to inefficient mixing and heating on short timescales.

Motivated by the above findings, in this paper we investigate how CR feedback affects the long-term evolution of the hot halo. Using magneto-hydrodynamic (MHD) simulations, we show that when CR transport and heating due to the streaming instability are included, CR-dominated jet can maintain global thermal equilibrium of the hot halo in elliptical galaxies and, at the same time, prevent the formation of the massive cold disks.

### 2 METHODOLOGY

We perform three-dimensional MHD simulations using the FLASH code (Fryxell et al. 2000; Dubey et al. 2008) with the directional unsplit staggered mesh MHD solver (Lee & Deane 2009; Lee 2013). The simulation domain is a statically-refined cubic box with of \( L_{\text{box}} \approx 250 \) kpc. The domain is refined by a set of nested cubic regions with increasing refinement levels. The entire region is refined by \((64)^3\) root grids and the central \( L_{\text{box}}/2^n\) wide regions have \( n \) additional nested levels of refinement \((n = 1, ..., 4)\) with the refinement level increasing with the decreasing distance from the domain center. Therefore, the smallest cell is \( L_{\text{min}} = L_{\text{box}}/(64)^2 \approx 0.244 \) kpc wide and the highest refinement region is \( L_{\text{min}}/2^4 \approx 16 \) kpc wide. We adopt the diode boundary conditions. That is, the gas can only flow out and all variables have zero gradient at the computational box boundary.

We follow our previous work (Wang et al. 2019) in setting up the initial conditions. The density and temperature profiles of the hot gas are initially in hydrostatic equilibrium. The gravitational potential includes contributions from stars, dark matter, and the central SMBH. Our initial conditions agree with NGC 5044, which is observed to have extended multiphase gas in the center (Werner et al. 2014; as explained below, our simulations develop cold phase component over time due to local thermal instability). Following Ruszkowski et al. (2007), we include tangled magnetic fields with the following power spectrum:

\[
B_k \propto k^{-11/6} \exp \left( -\left( \frac{k}{k_0} \right)^4 \right).
\] (1)

where \( k_0 = 10^2(2\pi/L_{\text{box}}) \). In order to generate magnetic fields characterized by the above spectrum, we first calculate real-space magnetic fields by inversely Fourier transforming the above power spectrum. We then renormalize the average field strength such that the plasma \( \beta \approx 100 \). We then Fourier transform the B-fields, clean B-field divergence in Fourier space and then perform inverse Fourier transform to obtain real-space B-fields. We repeat this procedure until the magnetic fields become divergence-free.

The key physical processes that we consider are radiative cooling, feedback from old stars, CR transport and heating, and AGN feedback. For radiative cooling, we use the tabulated cooling functions from Sutherland & Dopita.
Cosmic ray heating of ellipticals

Figure 1. Columns from left to right: projected gas density of KINETIC, K-MAG, CRCP, and CRSH runs. Within each column the slices are plotted in chronological order from top to bottom. The density is projected along the $y$-axis within the central 60 kpc-wide cube. The jet is along the $z$-axis.

(1993) assuming one solar metallicity (Buote et al. 2003; Komiyama et al. 2009). For old star feedback, we use the model described in Wang et al. (2019). In short, we consider the mass loss from evolved stars and thermal energy injection from type Ia supernovae. The ejecta from the evolved stars are assumed to be thermalized by the random stellar motions and the energy and mass are injected at a constant rate.

In the following sections, we describe the included CR physics (section 2.1) and the AGN feedback model (section 2.2).

2.1 CR physics

The CRs are modelled as a relativistic fluid with adiabatic index $\gamma_{cr} = 4/3$. We follow the evolution of the CR pressure $p_{cr}$ (or energy density $e_{cr} = p_{cr}/(\gamma_{cr} - 1)$) as a function of space and time. The CR implementation includes the following processes: (i) advection with the gas, (ii) transport of CRs along the magnetic field lines due to Alfvénic streaming (CR streaming/drift speed is determined by the streaming instability) and the associated energy transfer from CRs to the gas, and (iii) heating of gas by the CRs via Coulomb and hadronic interactions. We solve the MHD equations in the following form:

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

(2)

$$\frac{\partial (\rho \mathbf{u}_g)}{\partial t} + \nabla \cdot (\rho \mathbf{u}_g \mathbf{u}_g - \frac{\mathbf{B} \mathbf{B}}{4\pi}) + \nabla p_{tot} = \rho \mathbf{g}$$

(3)
\[
\frac{\partial e_{\text{tot}}}{\partial t} + \nabla \cdot \left( e_{\text{tot}} + p_{\text{tot}} \right) \mathbf{u}_g - \frac{\left( \mathbf{B} \cdot \mathbf{u}_g \right) \mathbf{B}}{4\pi} + \mathbf{F}_{\text{stream}} = \rho \mathbf{u}_g \cdot \mathbf{g} - C_{\text{rad}} - C_{\text{cr}, \text{C}} - C_{\text{cr}, \text{h}},
\]
where \( \rho \) is the gas density; \( \mathbf{u}_g \) is the gas velocity; \( \mathbf{B} \) is the magnetic field; and \( \mathbf{g} \) is the gravitational acceleration. The total pressure, \( p_{\text{tot}} \) is the sum of the gas thermal pressure, magnetic pressure and the CR pressure, \( p_{\text{tot}} = p + B^2/8\pi + p_{\text{cr}} \); and the total energy density \( e_{\text{tot}} \) is the sum of the gas internal energy, kinetic energy, magnetic energy, and the CR energy \( e_{\text{tot}} = \rho/(\gamma_{\text{gas}} - 1) + \rho u_g^2/2 + B^2/8\pi + e_{\text{cr}} \), where \( \gamma_{\text{gas}} = 5/3 \) is the adiabatic index of the ideal gas. The three terms on the right hand side of Eq.5, \( C_{\text{cr}, \text{C}}, C_{\text{cr}, \text{h}}, \) and \( C_{\text{cr}, \text{s}} \), account for the CR fluid energy losses due to Coulomb, hadronic interactions, and streaming heating, respectively. \( C_{\text{cr}, \text{C}} \) and \( C_{\text{cr}, \text{h}} \) are given by (Yoast-Hull et al. 2013):

\[
C_{\text{cr}, \text{C}} = 4.93 \times 10^{-19} \left( \frac{n_{\text{cr}}}{\text{cm}^{-3}} \right) \left( \frac{n_e}{\text{cm}^{-3}} \right) \text{erg} \cdot \text{cm}^{-3} \cdot \text{s}^{-1},
\]

\[
C_{\text{cr}, \text{h}} = 8.56 \times 10^{-19} \left( \frac{\gamma^5}{3} \right) \left( \frac{n_{\text{cr}}}{\text{cm}^{-3}} \right) \left( \frac{n_p}{\text{cm}^{-3}} \right) \text{erg} \cdot \text{cm}^{-3} \cdot \text{s}^{-1},
\]

where \( \gamma = 3 \) is the average Lorentz factor of CRs, \( n_p = \rho/(\mu_p m_p) \) is the gas proton number density, and \( n_e = \rho/(\mu_e m_p) \) is the gas electron number density. \( n_{\text{cr}} \) is the CR number density which is:

\[
n_{\text{cr}} = \frac{n - 4}{n - 3} E_{\min},
\]
where \( E_{\min} = 1\text{GeV} \) is the low energy cut-off of CR energy, and \( n = 4.5 \) is the slope of CR momentum distribution. In the self-confined picture, CRs stream with respect to the gas and get scattered on self-excited MHD waves. This leads to the transfer of CR energy to the gas. The heating rate is (Wiener et al. 2013a):

\[
C_{\text{cr}, \text{s}} = |\mathbf{u}'_A \cdot \nabla p_{\text{cr}}|,
\]
where \( \mathbf{u}'_A \) is the modified Alfvén velocity. We assume:

\[
\mathbf{u}'_A = \begin{cases} 
\text{min} (a_A, 200 \text{ km/s}), & r > 1 \text{ kpc} \\
0, & r \leq 1 \text{ kpc}
\end{cases}
\]

The size of the central region where the streaming heating is suppressed is comparable to the size of the base of the AGN jet. We treat the central 1 kpc region as a subgrid region and switch off streaming heating there to avoid artificial heating. In this approximation the streaming velocity cannot exceed than 200 km/s, which helps to save computational cost. We verified that this constraint limits the heating in
only in a very small fraction of cells and does not affect the conclusions presented below.

The terms on the right hand side of Eq.4, $C_{\text{rad}}$ and $C_{\text{net}}$, represent the radiative cooling losses of the gas and the net total energy loss of the composite gas plus CR fluid, respectively. As during hadronic collisions 5/6 of the CR energy escapes from the gas due to gamma-ray and neutrino emission, the net loss of the energy is

$$C_{\text{net}} = \frac{5}{6} C_{\text{cr}}.$$

i.e., 1/6 of the CR hadronic loss rate is used to heat the thermal gas, and the remaining 5/6 are permanently lost from the system.

The CR streaming flux $F_{\text{stream}}$ in Eq.4 and Eq.5 is

$$F_{\text{stream}} = -\left(\epsilon_{\text{cr}} + p_{\text{cr}}\right)\alpha A_{\text{cr}} (B \cdot \nabla \epsilon_{\text{cr}}).$$

To avoid infinitely fast variations of $\nabla \cdot F_{\text{stream}}$ at the local extrema of $\nabla \epsilon_{\text{cr}}$, we follow the method of Sharma et al. (2009) and regularize the streaming flux by a scale height $h_{\epsilon} = 10$ kpc:

$$F_{\text{stream}} = -\left(\epsilon_{\text{cr}} + p_{\text{cr}}\right)\alpha A_{\text{cr}} \tan \left(\frac{h_{\epsilon} B}{|B|} \nabla \epsilon_{\text{cr}}\right).$$

### 2.2 AGN feedback

For the fueling of the AGN, we adopt the cold mode accretion model, which has been used in many numerical studies in the literature (e.g., Gaspari et al. 2012, 2013; Li & Bryan 2014a,b; Yang & Reynolds 2016). The cold gas ($T < 10^5$ K) within the central $r < 0.5$ kpc region is accreted at the rate $M_{\text{acc}} = M_{\text{ac}}/5$ Myr, where $M_{\text{ac}}$ is the mass of cold gas within $r < 0.5$ kpc. The accretion-powered AGN feedback injects gas and CRs. The gas feedback is purely kinetic with kinetic power ($E_{\text{kin}}$) given by

$$E_{\text{kin}} = \frac{1}{2} \dot{M}_{\text{acc}} c^2 f_{\text{cr}} M_{\text{ac}} c^2;$$

and the CR feedback power ($E_{\text{cr}}$) given by

$$E_{\text{cr}} = \epsilon f_{\text{cr}} \dot{M}_{\text{acc}} c^2,$$

where $c$ is the speed of light, $\epsilon$ is the feedback efficiency, $\eta$ is the mass loading factor and $f_{\text{cr}}$ is the CR fraction of the AGN injected energy. We adopted $\epsilon = 3 \times 10^{-4}$ and $\eta = 1$ for all simulations in this work and $f_{\text{cr}} = 0.8$ for all simulations that included CRs. The AGN ejecta is launched via bipolar jets. The jet launching base is the innermost 8 cells and the jet precresses around the $z$-axis with the precession angle of 10 degrees and period of 10 Myr.

### 3 RESULTS

We now systematically discuss the results from our simulations. In this section we comment on simulations with progressively more complex physics of the CGM plasma, i.e., we begin with a purely hydrodynamical run (KINETIC run), then discuss its generalization to include magnetic fields (K-MAG run) and CRs including Coulomb and hadronic/pionic losses (CRCP run), and finally comment on the run that extends the above physics to include CR transport via the streaming instability and the associated with it CR heating of the CGM (CRSH run). Key parameters distinguishing these runs are shown in Table 1, where $f_{\text{cr}}$ denotes the fraction of the CR energy in the AGN jet.

### 3.1 AGN feedback in the hydrodynamical case

In the purely hydrodynamical AGN feedback case, the atmosphere goes through cycles of cold gas precipitation and AGN outbursts in the first 1.2 Gyr. At the beginning of each cycle, AGN outburst is triggered by either (i) central gas cooling due to short central cooling time, or (ii) infall of cold gas that condenses out of previously uplifted low entropy gas. The first cycle is initiated by the short central cooling time. The AGN jets perturb the halo triggering thermal instabilities. Consequently, cold gas condenses out of the halo and forms extended structures (see panel a1 in Fig.1). Some of the cold clumps fall toward the center, swirl around the central SMBH and eventually get accreted by it. A fraction of cold clumps find themselves in the path of the AGN jet, get fragmented by it, and get either ablated or accreted by the SMBH. Cold gas condensation ends when the AGN heating makes the halo thermally stable. When all cold gas gets consumed or destroyed, the AGN shuts down and the halo returns to the quiescent state (see panel a2 in Fig.1) until the next cycle begins. During the precipitation cycles, the cold gas mass fluctuates between $\sim 10^6$ and $\sim 10^8$ $M_\odot$, and the AGN jet power typically ranges between $10^{42.5}$ to $10^{43.5}$ erg s$^{-1}$ (see top and middle panels in Fig.2). This cyclic evolution is a natural result of the self-regulated AGN feedback. Many previous studies of cold mode accretion AGN feedback in giant elliptical galaxies or galaxy clusters led to similar conclusions with regard to the behavior of cold gas mass and jet power fluctuations. The cyclic evolution only lasts for the first 1.2 Gyr, after which the simulation goes into the unphysical disk phase described below.

#### 3.1.1 Formation of a massive cold disk in the vicinity of the SMBH

In the third precipitation cycle accreting cold gas forms a rotationally-supported disk at $t = 1.2$ Gyr. Although the formation of such disks has been seen in many simulations (e.g., Li & Bryan 2014a; Eisenreich et al. 2017; Qiu et al. 2018; Wang et al. 2019), we consider the evolution after the disk formation to be unphysical due to the following two reasons.

(i) The disk quickly becomes too large and too massive. Since the onset of disk formation, the cold gas mass in the disk increases monotonically as shown in the top panel of Fig.2. At $t = 2$ Gyr, the disk grows to 3 kpc in radius (third row of the left column in Fig.1) and $10^8$ $M_\odot$ in mass. Such massive cold disk is not found in NGC5044 or in most of the observed elliptical galaxies (Young et al. 2011).

(ii) The feeding of the AGN from the disk leads to the overheating of the central few kpc region while hot gas continues to cool directly onto the disk. Fig.3 shows the average radial distribution of gas cooling time, $t_{\text{cool}}(r)$, over two 200 Myr-long intervals – one just before the disk formation, $t = 1 - 1.2$ Gyr (left panel) and the other one 600 Myr later, $t = 1.8 - 2$ Gyr (right panel); the dashed lines in both panels represent the initial $t_{\text{cool}}(r)$. Before the formation of the disk, the system is self-regulated, and the gas cooling and AGN heating are overall in balance with each other. Therefore, $t_{\text{cool}}(r)$ is scattered approximately symmetrically around the initial values. In this phase, the cold gas is accreted onto the SMBH and the cold gas is removed from the central regions.
However, in later phase, the rate of cold gas consumption by the SMBH and the jet power reach a plateau. This happens because a centrifugally supported disk is formed and the angular momentum is not sufficiently quickly removed from it. Furthermore, because the disk rotates about the jet axis, the AGN outflow cannot intercept the cold gas and is unable to reheat it (this type of reheating is present in the simulations that exhibit cyclic AGN behavior). Therefore, the AGN consistently heats the halo while the disk continues to grow in mass as the hot phase cools and adds mass to it. As shown in the right panel of Fig. 3, while \( t_{\text{cool}}(r) \) systematically increases away from the initial condition due to the AGN heating in the \( 3 \text{kpc} < r < 10 \text{kpc} \) region, it cascades down for \( r \lesssim 3 \text{kpc} \) because the gas directly cools onto the disk. Therefore, even though the AGN overheats the halo, it fails to stop the fast cooling of the gas. The whole system is no longer self-regulated and its thermal properties are systematically changed compared to the initial condition.

### 3.2 AGN feedback in the magneto-hydrodynamical case

We now discuss the impact of magnetic fields on the purely hydrodynamical simulations of the AGN feedback described above. Similarly to the KINETIC case, in the K-MAG case the atmosphere and the AGN undergo cyclic variations. In the K-MAG case thermal instability is also present and cold gas clouds decouple from the hot atmosphere. However, unlike in the KINETIC case, massive cold disk surrounding the SMBH does not form (c.f. last row of the first and second column in Fig. 1). Moreover, the atmosphere remains in a state of global thermodynamic equilibrium without overheating in the time-averaged sense throughout the simulation time rather than exhibiting overheating that occurs in the KINETIC case at about 1.2 Gyr. As we argue below, both of these differences can be attributed to the impact of magnetic tension acting on the clouds and magnetic breaking that leads to more efficient extraction of angular momentum and accretion of the cold clumps onto the central SMBH.

#### 3.2.1 Morphology of the precipitating clouds

In order to unravel the reason for the absence of cold disks in the K-MAG case, we now turn our attention to the morphology of the clouds. To this end, we compare how cold clumps evolve in the KINETIC and K-MAG runs and analyze one event of the infall of an isolated cold clump in each run. Fig. 4 shows snapshots of the projected density at the beginning of the infall of isolated clumps (top row) and at the time when the cold clumps reach the center (bottom row). Arrows point to the locations of the cold clouds when they are compact. While in the KINETIC case the cloud appears compact throughout its infall, in the K-MAG case the cloud is compact in the beginning and then becomes extremely filamentary as it reaches the center. The filaments are approximately aligned with the radial direction, which to first order coincides with the infall trajectory (see middle column). This difference in the morphology of the cold gas clouds between KINETIC and K-MAG cases strongly suggests that magnetic forces may be important. We note in passing that the clouds in the KINETIC case also become filamentary in those cases where they directly interact with the AGN outflow.

#### 3.2.2 Impact of magnetic fields on the cold phase morphology

Although the mean magnetic field in the CGM is dynamically weak (initial plasma \( \beta \approx 100 \)), this does not preclude a possibility that the B-field could be locally enhanced. In
Cosmic ray heating of ellipticals

Figure 4. Infall of an isolated cold clump in (from left to right column) KINETIC, KMAG, and CRSH run. In each column, the top row is the projected density snapshot at the time when the condensed cold clump starts to fall in and the bottom row corresponds to the epoch when the head of the clump reaches the center. Arrows point to the locations of the cold clouds when they are compact.

Figure 5. Three sets of pressure profiles: thermal pressure (blue), magnetic pressure in the regions occupied by the cold gas (red), and magnetic pressure averaged in spherical shells (grey). For each set, the profiles are plotted every 50 Myr from light to dark color. For clarity, the initial thermal and magnetic pressure profiles are shown as yellow and black dashed lines, respectively.

Figure 5 we show the profiles of the thermal pressure (in blue), magnetic pressure averaged in spherical shells (in gray), and the magnetic pressure in cold gas regions ($T < 10^5 \text{K}$). In each category, different lines correspond to different simulation times with darker lines corresponding to later times. As expected, the average plasma $\beta \gg 1$. Interestingly, the cold gas phase is spatially correlated with enhanced magnetic field strength.

3.2.3 Magnetic breaking and cold gas velocity distribution

The local B-field amplification occurs due to flux freezing during the condensation of the cold gas followed by the field amplification during cloud infall. This amplification could occur on the leading edge of the cloud as well as in the filamentary tail. As the clumps fall to the center, the strong magnetic fields are stretched and exert tension force in the direction anti-parallel to the clump velocity direction. This mechanism is clearly shown in Fig.6, which shows the gas density distribution on a plane containing the cold filament at two different times (see left and middle panels). The spatial orientation of this plane is shown in the upper right panel of Fig.6. The magnetic field lines are shown as arrows. As the cold clump is falling to the center, it bends and stretches the magnetic field lines. The field lines exert strong tension force on the cold clump. Any gas that is stripped from the cloud is...
Figure 6. Left and middle panel: distribution of gas density and magnetic field lines on the plane \( S' \) at the epoch of the formation of a single filament (left, \( t = 0.66 \) Gyr) and at the epoch when the filament head reaches the center (middle, \( t = 0.70 \) Gyr). The gas density and magnetic field lines are shown as color maps and stream lines, respectively. Upper right panel: the position of plane \( S' \) relative to the \( x, y, z \) axes of the simulation box. The normal vector and the intersection with the \( x-y \) plane of \( S' \) are labelled as \( \hat{n} \) and \( x' \), respectively. The plane \( S' \) slices the volume through the bulk of the filament, which is schematically shown as the orange dashed line.

Figure 7. Kinetic energy - magnetic tension “resistence” phase plot of the cold gas in K-MAG run. Color represents the cold gas mass. The distribution is averaged over snapshots sampled every 2 Myr within \( t = 0-2.5 \) Gyr. Dashed lines denote the damping time scale \( t_{\text{damp}} = 1, 10, 100 \) Myr and 1 Gyr. The damping timescale is defined as the timescale over which the work done by the magnetic tension force depletes the kinetic energy of the cold gas clumps. The magnetic tension force density is

\[
f_t = \frac{1}{4 \pi} \nabla \cdot (B B) \tag{15}
\]

We found that, on average, 75 per cent of the cold gas mass has \( f_t \cdot u_g < 0 \), i.e., that the magnetic tension force does neg-
KINETIC run (solid blue line) has a high velocity tail that is absent in the K-MAG case (dotted black line).

3.2.4 Ambipolar diffusion of magnetic fields

The impact of magnetic tension on the cloud velocity can be reduced when the magnetic fields are allowed to slide past the clouds. This can happen as a result of ambipolar diffusion since the clouds are only partially ionized. In order to assess the impact of this process, we estimate the ambipolar diffusion timescale on which the field sliding occurs, $t_{damp} \sim L^2 v_{in}/\nu_{H}$, where $L$ is the filament length, $v_{in} \sim 10^{-9} T_2^{1/2} n_{H}^{-1} s^{-1}$ is the ion-neutral collision frequency (De Pontieu et al. 2001), where $n_{H}$ is the number density of neutrals and $T_2 = T/10^2 K$, where $T$ is the cold phase temperature. We can compare this timescale to the dynamical time $t_{dyn} \sim R/u$, where $R$ is the distance of a cloud from the center and $u$ is its velocity,

$$\frac{t_{damp}}{t_{dyn}} \sim 1.3 \times 10^{-8} x(1 − x) T_2^{1/2} n_{H} L^2 u / R.$$ (16)

where $x$ is the ionization fraction, and $n_{H}$ is the number density of hydrogen. In realistic filaments most of the gas mass is likely to be in the molecular form (Ferland et al. 2008, 2009). Assuming that, for low ionization, recombination are balanced by CR ionization, we can estimate the ionization fraction as $x \sim K/n_{H}^{-1/2}$, where $K \sim 10^{-5} cm^{-3}/2$ (McKee et al. 1993; Padoan et al. 2000). This estimate for the ionization fraction assumes ionization rate $\dot{\zeta}_{H} = 10^{-17}s^{-1}$ Spitzer (1978). Since this ionization rate is uncertain and can be larger (e.g., Indriolo et al. (2010, 2018)), and $K \sim \zeta_{H}^{-1/2}$, the above estimate for the ionization fraction is likely a lower limit. For typical, if somewhat conservative, filament parameters $R \sim 10$ kpc, $L \sim 10$ pc (Fabian et al. 2008), $u \sim 200 km s^{-1}$, $n_{H} \sim 2 \times 10^{4} cm^{-3}$, $T \sim 10^{2}$ K, and $B \sim 50 \mu G$ (see Werner et al. (2013) for the estimates of the last three quantities), we obtain $t_{damp}/t_{dyn} \sim 1.4 \times 10^{3}$. This implies that ambipolar diffusion is unlikely to lead to significant magnetic line slippage past the cold filaments. While detailed predictions for the filament powering rates are beyond the scope of this work, we also estimated that unlike other mechanisms proposed to significantly contribute the filament powering (e.g., Fabian et al. (2011); Ferland et al. (2008); Churazov et al. (2013); Ruszkowski et al. (2018)), the ambipolar drift heating rate is unlikely to account for the observed amount of power emitted by the Hα filaments, which is consistent with relatively long ambipolar diffusion timescales. These conclusions become even stronger for the case of filaments composed predominantly of Hα-emitting gas of typical density of $\sim 30 \ cm^{-3}$ and $T \sim 10^{4}$K (Werner et al. 2013).

3.3 AGN feedback including magnetic fields and CRs

As mentioned in the introduction, AGN jets inflate cavities in the CGM that can be dominated by CRs. CR simulations of AGN feedback (Yang et al. 2019) suggest that such CR-filled bubbles could uplift more cold thermally unstable gas because they have lower inertia and flatten in the direction perpendicular to the radial direction, which increases their

Significant magnetic tension forces that are anti-parallel to the velocity vectors of the cold gas not only affect the morphology of that gas phase but also modify its velocity distribution. Fig.9. shows the average distribution of the magnitude of the cold gas velocity along the $x$-axis. To make an equivalent comparison between the runs, we only include the data for the first 1.2 Gyr, during which both the KINETIC and K-MAG runs experience similar cyclic evolution. This figure clearly demonstrates that the cold gas in

...
cross section. These short-duration simulations also show that there is not enough time to transfer the CR energy from the bubbles to the ambient CGM. We now extend the results presented in the above sections to include CR-dominated bubbles inflated by AGN jets in order to study their long-term impact on the CGM.

### 3.3.1 Simulations without CR streaming

We begin by including CRs and their losses via Coulomb and hadronic interactions with the CGM (run CRCP). These processes heat the thermal CGM. The CRCP run does not exhibit a cycling behavior seen in the KINETIC and K-MAG cases. Instead, it suffers from the cooling catastrophe despite the fact that the AGN continues to inject energy. Cold gas piles up in the center since the early times in the simulation ($t \approx 120$ Myr) and forms an overdense region that gradually grows with time (see third column in Fig.1). The cold gas mass increases monotonically as shown in the top panel of Fig.2. At $t \approx 1.4$ Gyr, the mass of the cold gas reaches $\sim 10^{10} M_\odot$. This is two orders of magnitude more than typical values seen in elliptical galaxies. Therefore, we stop the simulation at $t \approx 1.4$ Gyr. One fundamental difference between the CRCP case and the KINETIC and K-MAG cases is that the energy transfer from the bubbles to the CGM proceeds at different rates. While in the KINETIC and K-MAG cases the thermal bubble gas eventually mixes with the surrounding CGM and instantly transfers its energy to the CGM, in the CRCP case the CRs that come into contact with the CGM via Coulomb and hadronic interactions. Assuming that the bubbles are in pressure equilibrium with the CGM, we can estimate the heating timescale $t_{\text{heat}}$ using Eq. 6 and 7. For typical conditions in the cool core at 10 kpc away from the center ($n_p \sim 10^{-2}$ cm$^{-3}$, $T \sim 10^7$ K), we obtain $t_{\text{heat}} \sim 1$ Gyr, which is a significant fraction of the simulation duration and suggests that the energy transfer from CRs to the CGM should be inefficient.

---

**Figure 10.** Projected cold gas ($T < 10^5$ K) density (left) and CR energy in the CRCP (middle) and CRSH (right) cases at $t = 1.43$ Gyr. The projections are along the y-axis and cover the central 80 kpc-wide cubic region.

---

Fig.10 shows the projected distribution of the cold gas density (left panel) and CR energy density in the CRCP run (middle panel) at $t \approx 1.4$ Gyr. The two distributions are very similar, which indicates that the cooling gas from the outer parts of the cool core flows in and traps the CR fluid in the region occupied by the cold gas. Although energetic CRs are co-spatial with the cold gas, the inefficient CR heating is unable to reheat the cold gas, which has excessively short cooling times. Consequently, cold gas accumulates in the center and the CRCP run suffers from a cooling catastrophe. Similar result was found in the context of cluster simulations (Ruszkowski et al. 2017).

### 3.3.2 Simulations with CR streaming and heating

We now extend the CR simulations of AGN feedback discussed above to include the effect of the CR streaming instability (CRSH run). That is, we include CR streaming with respect to the thermal gas and the associated CR heating of the CGM. The right panel in Fig.10 shows the projected distribution of the CR energy density in the CRSH run. This figure clearly demonstrates that the spatial distribution of CRs can be much broader in this case compared to the one where CR transport is neglected (CRCP run; middle panel). The better dispersal of CRs throughout the CGM in the CRSH case compared to CRCP case leads to more efficient heating as the CRs can come into contact with more hot thermal gas. It is evident from the comparison of the middle and right panels in Fig.10 that the energy density of CRs is lower in the CRSH case compared to CRCP case. The decrease in the CR energy density is due to a combination of adiabatic expansion and streaming of CRs and the additional heating of the gas associated with the CR streaming instability. This heating timescale $t_{\text{heat, A}} \sim 3L/u_A$, where $L$ is the scalelength of the CR gradient where significant heating takes place. Periods of the most significant CR heating coincide with the AGN outbursts when the pressure gradients of CR escaping from the
jet regions are the strongest, and there is relatively little heating between AGN episodes when the CRs have had time to disperse substantially. Depending on the steepness of the CR pressure gradient, $t_{\text{heat,}\Lambda}$ can be significantly shorter than the heating timescale $t_{\text{heat}}$ corresponding to Coulomb and hadronic losses described above (e.g., $\sim 5 \times 10^7$ yr for $L \sim 1$ kpc and assuming $\beta = 10^2$). Similar effect was observed in the context of CR simulations of AGN feedback in galaxy clusters (Ruszkowski et al. 2017). Consequently, the cooling catastrophe that was present in the CRCP case does not occur. This is illustrated in Fig.11 that shows the temperature versus gas density phase space. In the left panel, that corresponds to the CRCP run, large quantities of dense and low temperature gas can be seen, whereas only a relatively very small amounts of precipitating gas can be seen in the right panel that corresponds to the CRSH case. Similarly to the K-MAG case, the CRSH case also exhibits cyclic behavior that was absent in the CRCP run (see orange line in Fig.2). The average temperature - density phase diagrams plotted over $t = 0.9 - 1.1$ Gyr in CRCP (left) and CRSH run (right). The averaging method is the same as that used by Fig.3. Color shows the gas mass.

Figure 11. The average temperature - density phase diagrams plotted over $t = 0.9 - 1.1$ Gyr in CRCP (left) and CRSH run (right). The averaging method is the same as that used by Fig.3. Color shows the gas mass.

Table 2. Components of AGN energy

| name   | $E_{\text{agn}}$(erg) | $E_{\text{cr}}$(erg) | $E_{\text{cr,used}}$(erg) | $E_{\text{cr,used}}/E_{\text{agn}}$ |
|--------|-----------------------|----------------------|--------------------------|-------------------------------------|
| CRCP   | $2.41 \times 10^{59}$ | $6.28 \times 10^{58}$ | $1.30 \times 10^{59}$    | $84\%$                             |
| CRSH   | $1.81 \times 10^{59}$ | $2.33 \times 10^{58}$ | $1.22 \times 10^{59}$    | $67\%$                             |

That the evolution of the CRCP run differs so substantially from that seen in the CRSH run suggests that CR streaming plays a decisive role in determining whether the CR-dominated AGN jets could efficiently deliver the energy to the CGM. To better understand how efficiently the AGN energy is supplied to the CGM, we analyze the following quantities in these two runs: total energy released by the AGN $E_{\text{agn}}$, total energy of the CR fluid that remains in the simulation box $E_{\text{cr}}$, and the CR energy transferred to the gas $E_{\text{cr,used}} = f_{\text{cr}}E_{\text{AGN}} - E_{\text{cr}}$ computed assuming negligible CR outflow through the box boundary. In principle, $E_{\text{cr,used}}$ should exceed the energy that is actually used to heat the gas because a fraction of CR energy that lost due to hadronic collisions is removed from the system via gamma ray and neutrino emission. However, we verified that the hadronic energy losses in both cases are small compared to other terms and therefore approximately all of $E_{\text{cr,used}}$ is utilized to heat the gas. Comparison of these quantities for the CRCP and CRSH runs at $t = 1.2$ Gyr (see Table 2) reveals that: (i) $E_{\text{agn}}$ is larger in CRCP run than that in CRSH run. This is likely caused by the weak coupling of the AGN energy to the gas. The AGN is trying to compensate for this weak heating efficiency by injecting more energy into the CGM. One manifestation of this is the fact that while in the CRCP run the AGN is continuously injecting energy, the CRSH run undergoes a quiescent phase at $t = 750$ Myr during which AGN releases zero energy as the atmosphere is relatively close to being in the state of global thermal equilibrium (the solid orange line in the middle panel of Fig.2); (ii) the above finding is consistent with the observation that, due to the additional streaming heating, CR energy is more efficiently utilized in CRSH run, i.e., $E_{\text{cr,used}}/(f_{\text{cr}}E_{\text{AGN}}) \approx 84\%$ and $\approx 67\%$ in the CRSH and CRCP cases, respectively.

The total CR energy transferred to gas is slightly higher in the CRCP (as the AGN is trying to compensate for the low heating efficiency), but the heating in this case still fails to balance cooling globally. However, the ability of the AGN...
to shut off catastrophic cooling of the atmosphere depends not only on the total amount of energy transferred to it but also on how uniform the heating is distributed and how much of the cooling gas can come into contact with CRs. As mentioned above, the spatial distribution of CRs is broader in the CRSH case which, together with the fact that the amount of the injected energy is better utilized in this case, explains why the cooling catastrophe is prevented when CR streaming effects are included.

The better utilization of the AGN energy is also evident from the bottom panel in Fig. 2. This figure shows the evolution of the total AGN energy $E_{AGN}$. The smallest values of $E_{AGN}$ correspond to the CRSH case, i.e., the smallest values of AGN energy suffice to offset cooling in a time average sense and prevent cooling catastrophe.

4 CONCLUSIONS

We performed MHD simulations of AGN feedback in elliptical galaxies including the effects of CR heating and streaming. In these simulations we followed the evolution for a long period of time (Gyrs). We found that non-ideal hydrodynamical effects have a profound impact on the evolution of elliptical galaxies. In other words, the evolution of the systems on 100 kpc scales depends sensitively on the physics governing the “microscopic” scales relevant to plasma physics. Specifically, we showed that

(i) in the purely hydrodynamical case, the AGN jet initially maintains the atmosphere in global thermal balance. However, local thermal instability eventually leads to the formation of massive cold disk in the vicinity of the central SMBH. This is a generic finding. Similar conclusion was found in other grid-based and smooth particle hydrodynamics simulations performed by other groups. Once the disk has formed, it continues to feed the central SMBH and the jet begins to overheat the central regions of the atmosphere and does not prevent fast cooling of the gas in the direction perpendicular to it close to the center. Thus the AGN-to-CGM energy coupling is weak due to the inflow-outflow geometry very close to the SMBH. This leads to further disk growth and excessive buildup of the cold gas in the very center, which is in disagreement with the observations;

(ii) including high-$\beta$ magnetic fields prevents the formation of the cold disks. The removal of the cold gas from the central regions is due to local B-field amplification in the cold precipitating gas and increased magnetic breaking acting on the cold blobs. The local field amplification occurs via flux freezing during the development of local thermal instability, which is followed by stretching of the fields caused by the motion of the cold gas clumps that dynamically decouple from the hot gas. The B-field line stretching leads to filamentary/cometary appearance of the cold clouds that do not fall ballistically onto the center. In a statistical sense, the magnetic tension vectors are anti-parallel to the cold cloud velocities, which leads to magnetic breaking of the cold clouds. Even though the magnetic field in both the cold and hot phases remains dynamically weak throughout the evolution, the tension forces acting on the cold clouds are capable of affecting their dynamics on timescales much shorter than the evolutionary timescale of the system. We also find that magnetic fields can narrow the velocity distribution of the cold phase (compared to that observed in the purely hydrodynamical case);

(iii) when plasma composition in the AGN jets is dominated by CRs, and CR transport is neglected, the atmospheres exhibit cooling catastrophes due to the inefficient heat transfer from the AGN to CGM despite Coulomb/hadronic CR losses being present. In this case, CRs continue to accumulate in the center due to the fact that CRs do not come into direct contact with sufficiently large fraction of the radiatively cooling atmosphere. CR energy is not utilized very efficiently due to the lack of transport and the fact that typical CR cooling timescales associated with Coulomb and hadronic interactions are a substantial fraction of the system lifetime;

(iv) including CR streaming and heating restores the agreement with the observations, i.e., cooling catastrophes are prevented, the gas does not overheat, and massive cold central disks do not form. The AGN power is reduced as its energy is utilized efficiently.

ACKNOWLEDGEMENTS

MR acknowledges NSF grant AST 1715140. HYKY acknowledges support from NASA ATP (NNX17AK70G) and NSF (AST 1713722). We thank Yuan Li for useful discussions. MR thanks Eric Blackman, Jim Drake, Peng Oh, Christoph Pfrommer, and Ellen Zweibel for very enlightening discussions at KITP. This work used the Extreme Science and Engineering Discovery Environment (XSEDE, Towns et al. 2014), which is supported by National Science Foundation grant number ACI-1548562. We acknowledge the XSEDE computational grant TG-AST180063. MR thanks the Aspen Center for Physics and the Kavli Institute for Theoretical Physics at the University of California at Santa Barbara, where part of this work was performed. The Aspen Center for Physics is supported by National Science Foundation grant PHY-1607611. This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1748958.

REFERENCES

Barai P., Murante G., Borgani S., Gaspari M., Granato G. L., Monaco P., Ragone-Figueroa C., 2016, MNRAS, 461, 1548
Brüggen M., Kaiser C. R., 2002, Nature, 418, 301
Buote D. A., Lewis A. D., Brighenti F., Mathews W. G., 2003, ApJ, 595, 151
Churazov E., Ruszkowski M., Schekochihin A., 2013, MNRAS, 436, 526
De Pontieu B., Martens P. C. H., Hudson H. S., 2001, ApJ, 558, 859
De Young D. S., 2010, ApJ, 710, 743
Dubey A., Reid L., Fisher R., 2008, Physica Scripta, 2008, 014046
Dunn R. J. H., Fabian A. C., 2004, MNRAS, 355, 862
Ehlert K., Weinberger R., Pfrommer C., Pakmor R., Springel V., 2018, MNRAS, 481, 2878
Eisenreich M., Naab T., Choi E., Ostriker J. P., Emsellem E., 2017, MNRAS, 468, 751
Enßlin T., Pfrommer C., Miniati F., Subramanian K., 2011, A&A, 527, A99
Fabian A. C., 1994, ARA&A, 32, 277

Wang et al.
