A Black Hole Feedback Valve in Massive Galaxies

G. Mark Voit1 ©, Greg L. Bryan2,3 ©, Deovrat Prasad1 ©, Rachel Frisbie1 ©, Yuan Li3,4, Megan Donahue1 ©, Brian W. O'Shea1,5,6 ©, Ming Sun7 ©, and Norbert Werner8 ©

1 Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA
2 Department of Astronomy, Columbia University, New York, NY, USA
3 Center for Computational Astronomy, Flatiron Institute, New York, NY, USA
4 Department of Astronomy, University of California, Berkeley, Berkeley, CA, USA
5 Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, MI, USA
6 National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI, USA
7 Department of Physics and Astronomy, The University of Alabama in Huntsville, Huntsville, AL, USA
8 MTA-Eötvös University Lendület Hot Universe Research Group, Pázmány Péter setány 1/A, Budapest, 1117, Hungary

Received 2019 October 27; revised 2020 June 16; accepted 2020 July 7; published 2020 August 13

Abstract

Star formation in the universe’s most massive galaxies proceeds furiously early in time but then nearly ceases. Plenty of hot gas remains available but does not cool and condense into star-forming clouds. Active galactic nuclei (AGNs) release enough energy to inhibit cooling of the hot gas, but energetic arguments alone do not explain why quenching of star formation is most effective in high-mass galaxies. In fact, optical observations show that quenching is more closely related to a galaxy’s central stellar velocity dispersion (σv) than to any other characteristic. Here we show that high σv is critical to quenching because a deep central potential well maximizes the efficacy of AGN feedback. In order to remain quenched, a galaxy must continually sweep out the gas ejected from its aging stars. Supernova heating can accomplish this task as long as the AGN sufficiently reduces the gas pressure of the surrounding circumgalactic medium (CGM). We find that CGM pressure acts as the control knob on a valve that regulates AGN feedback and suggest that feedback power self-adjusts so that it suffices to lift the CGM out of the galaxy’s potential well. Supernova heating then drives a galactic outflow that remains homogeneous if σv ≥ 240 km s⁻¹. The AGN feedback can effectively quench galaxies with a comparable velocity dispersion, but feedback in galaxies with a much lower velocity dispersion tends to result in convective circulation and accumulation of multiphase gas within the galaxy.

Unified Astronomy Thesaurus concepts: Early-type galaxies (429); Elliptical galaxies (456); Giant elliptical galaxies (651); Quenched galaxies (2016); Red sequence galaxies (1373); X-ray astronomy (1810); Circumgalactic medium (1879)

1. Introduction

The surest way to quench star formation in a galaxy is to rid it of molecular gas. That end can be accomplished either gradually, by turning molecular gas into stars faster than it can accumulate, or abruptly, by destroying or ejecting all of the galaxy’s molecular clouds. However, star formation will eventually resume unless the galaxy can prevent more molecular gas from accumulating. Three different gas sources must therefore be prevented from supplying cold gas.

1. Cold streams. Accretion of cold gas along cosmological dark matter filaments can potentially feed star formation (e.g., Keres et al. 2005, 2009; Dekel et al. 2009). If not disrupted before reaching the bottom of the local potential well, those cold accretion streams will enter the central galaxy. Most current models of galaxy evolution therefore posit that quiescent central galaxies have hot gaseous halos that disrupt cold streams. They also posit that quiescent satellite galaxies orbiting the central one cannot access cold streams because of their displacement from the center.

2. Cooling flows. Even if cold streams are disrupted by a hot halo, radiative cooling of the densest gas in that hot halo can still supply cold gas to the central galaxy (e.g., White & Frenk 1991; Fabian 1994). Most current models of galaxy formation therefore posit that accretion of cooling gas onto a central supermassive black hole releases enough energy to offset most of the radiative losses, thereby limiting the supply of cold gas to the central galaxy (e.g., McNamara & Nulsen 2007, 2012; Werner et al. 2019).

3. Stellar mass loss. The third gas source is the aging stellar population of a quiescent galaxy (e.g., Mathews & Brighenti 2003; Leitner & Kravtsov 2011; Voit & Donahue 2011). As dying stars shed their surplus gas, it accumulates in the galaxy, where it is heated by exploding white dwarfs (SNe Ia). Supernova heating is energetically capable of sweeping ejected stellar gas out of a galaxy, but the pressure of the galaxy’s circumgalactic medium (CGM) limits the rate at which it can do so. That is because the confining CGM pressure determines the outflow’s gas density and radiative losses. If the CGM pressure is too great, supernovae cannot sweep out the ejected stellar gas.

This paper analyzes the three-way coupling that can occur between supernova sweeping of stellar ejecta, the confining CGM pressure, and bipolar kinetic feedback fueled by accretion of cooling gas onto the central black hole. Together, they make a valve that regulates fueling of the active galactic nucleus (AGN). The ideas presented build upon an enormous literature that is impossible to adequately summarize here, so we will restrict ourselves to pointing out some highlights.
Mathews & Baker (1971), in a seminal and insightful paper on galactic winds, laid a foundation that remains strong nearly half a century later. Interestingly, their pursuit of outflow solutions for supernova-driven galactic winds was equally motivated by both the absence of cold interstellar gas and the presence of nonthermal radio sources in massive elliptical galaxies. Mathews & Brighenti (2003) extensively reviewed the line of research that followed.

A couple of decades later, Tabor & Binney (1993) and Binney & Tabor (1995) focused more closely on how cooling-flow accretion onto a central black hole in a massive elliptical can limit the overall condensation rate. They recognized that the radial profiles of pressure, density, and temperature inferred from X-ray observations resemble galactic outflows near the onset of a cooling catastrophe that supernova heating cannot prevent. While supernova heating can drive gas out of a massive elliptical galaxy, it cannot unbind that gas from the galaxy’s halo (e.g., Mathews & Loewenstein 1986; Loewenstein & Mathews 1987; David et al. 1990). Without additional heating, ejected stellar mass therefore collects in the CGM and raises the confining pressure. Eventually, the central density of the confined outflow increases enough for radiative cooling to exceed supernova heating, and cooling gas starts to flow toward the origin. Binney & Tabor proposed that the central cooling flow should then fuel accretion onto a supermassive black hole, producing bipolar jets that heat and expand the surrounding medium, raising its entropy and temporarily alleviating the cooling catastrophe.

Around the same time, Ciotti et al. (1991) launched a series of increasingly sophisticated simulations, at first to explore the evolution of supernova-heated outflows and later to explore the coupling between those outflows and sporadic black hole outbursts (Ciotti & Ostriker 2001, 2007; Ciotti et al. 2010, 2017). Those models demonstrated the importance of both kinetic feedback (Ostriker et al. 2010; Choi et al. 2015) and the circumgalactic environment (Shin et al. 2012) to the evolution of hot gas in massive ellipticals, but they did not directly explore the effects of AGN feedback on the confining CGM pressure. During much of the time, the simulated AGN power is relatively low, allowing radiative cooling to reduce the central entropy and raise the central gas density until a cooling catastrophe begins. The rapid rise in cooling then fuels a strong burst of AGN feedback, lasting several tens of Myr, that boosts the central entropy and cooling time while lowering the central gas density.

More recently, three-dimensional simulations have vividly illustrated how kinetic feedback limits cooling and condensation of ambient galactic gas and why it is more effective than purely thermal feedback (e.g., Gaspari et al. 2011a, 2011b, 2012; Dubois et al. 2012; Li & Bryan 2014a; Li et al. 2015; Prasad et al. 2015, 2017; Yang & Reynolds 2016; Meece et al. 2017; Beckmann et al. 2019). Strong, narrow jets can drill through the ambient medium within a few kpc of the AGN without significantly disturbing much of it, and that enables the jets to transport most of their kinetic energy tens of kpc from the origin before thermalizing it in the CGM (see also Soker 2016). This transport mechanism distributes thermal energy over a large volume without producing excessive convection, which destabilizes the ambient medium and results in overproduction of multiphase gas (Meece et al. 2017; Voit et al. 2017).

Here we build upon those previous efforts by applying recent insights into the thermal stability of the CGM. Our objective is to determine the conditions that allow supernova heating to sweep ejected stellar mass out of a massive elliptical galaxy in an outflow that remains stable to multiphase condensation (see Voit et al. 2015c, hereafter V15, for a preliminary analysis). In order for the ambient medium to remain homogeneous, its ratio of radiative cooling time ($t_{cool}$) to freefall time ($t_f$) must satisfy the condition $t_{cool}/t_f \gtrsim 10$ (see Section 2.7). Also, supernova heating must exceed radiative cooling within most of the galaxy’s volume, requiring the ambient gas density to decline with radius at least as rapidly as $\propto r^{-5}$. In that case, radiative cooling per unit volume declines at least as rapidly as supernova heating, which is proportional to the stellar mass density (approximately $\rho_s \propto r^{-2}$). These conditions cannot be satisfied unless something other than supernova heating lowers the pressure of the CGM surrounding the galaxy. We therefore assume that powerful but sporadic AGN outbursts clear the way for supernova sweeping by lowering the confining CGM pressure. Those outbursts occur whenever cooling at small radii supplies the central black hole with an unusually large amount of fuel. As a consequence, the CGM pressure ultimately governs the maximum black hole fueling rate because it determines the mass inflow rate of the inner cooling flow.

The most far-reaching result of this paper is that this critical gas density profile ($\rho \propto r^{-4}$) corresponds to a critical stellar velocity dispersion ($\sigma_v \approx 240$ km s$^{-1}$), above which coupling of AGN feedback to CGM pressure can keep star formation permanently quenched. Nearly complete quenching can happen because the ambient medium then satisfies the conditions for supernova heating to drive a steady, homogeneous flow of ejected stellar gas out of the galaxy.

Readers interested in the basic physical picture should start with Section 2, which shows that the critical density slope is determined by a critical entropy profile slope that depends almost exclusively on the ratio of the specific energy of stellar ejecta to the square of the gravitational potential’s circular velocity. It therefore depends on the quotient of the specific supernova heating rate and specific stellar mass-loss rate, which both evolve with time. Section 3 presents a more formal treatment, adding rigor to the analytical estimates of Section 2 by comparing them with numerical integrations of the steady flow equations. Observational support for the model can be found in Section 4, which validates the numerical solutions through comparisons with X-ray observations of massive elliptical galaxies and infers the role of intermittent AGN feedback from the discrepancies. Section 5 discusses the resulting implications and predictions for the quenching of star formation, focusing in particular on why quenching should depend primarily on central stellar velocity dispersion and secondarily on halo mass. Section 6 summarizes the paper.

2. The Basic Picture

This section outlines the black hole feedback valve model by presenting a series of simple analytical estimates intended to make the overall physical picture intuitively clear. It begins by modeling the hot ambient medium as a series of concentric spherical shells, each with a constant temperature, so that gas density depends only on specific entropy. It then shows how the structure of the resulting galactic atmosphere depends on the entropy profile slope determined by the average specific energy of stellar ejecta. It explains how the radius $r_{eq}$ at which
local supernova heating equals local radiative cooling is determined by an outer pressure boundary condition, and it demonstrates that supernova heating cannot sweep ejected stellar gas out of the galaxy unless AGN feedback reduces the confining CGM pressure to a small fraction of its cosmological value. The section concludes with an assessment of the conditions under which the outflow is stable to multiphase condensation and some estimates of the AGN feedback power and momentum flux required to make the basic picture self-consistent.

2.1. Piecewise Isothermal Atmosphere

A galaxy’s ambient atmosphere can be approximated with a piecewise isothermal hydrostatic model as long as its velocity field is sufficiently subsonic. Quantifying specific entropy in terms of the entropy index $K \equiv kTn_e^{-2/3}$, where $T$ is the gas temperature and $n_e$ is the electron density, leads to

$$\frac{d \ln P}{d \ln r} = -\frac{3}{2} \frac{\ln K}{\ln r} = -\frac{\mu m_p v_c^2}{kT},$$

(1)

within each hydrostatic, isothermal shell of gas in a gravitational potential with circular velocity $v_c$. The logarithmic entropy slope in each shell is

$$\alpha_K \equiv \frac{d \ln K}{d \ln r} = \frac{2}{3} \frac{\mu m_p v_c^2}{kT},$$

(2)

giving $K \propto r^{-\alpha_K}$ and $n_e \propto r^{-3\alpha_K/2}$ for an isothermal potential in which $v_c$ is constant with radius. Within each shell, the constant of proportionality relating hydrostatic temperature to $v_c^2$ is determined by the entropy slope $\alpha_K$ of the galactic atmosphere in that shell. A galactic atmosphere in which $T$ is a slowly changing function of radius can therefore be approximated with a set of thick concentric isothermal shells in which $T$ and $\alpha_K$ are related through Equation (2).

Cosmological structure formation tends to produce dark matter halos similar to singular isothermal spheres in which gas density is approximately proportional to dark matter density. Hydrostatic gas in an idealized cosmological halo can therefore be approximated by choosing $\alpha_K = 4/3$, $n_e \propto r^{-2}$, and $T = T_\infty$, with $T_\infty \equiv \mu m_p v_c^2/2k \approx \text{const.}$, but radiative cooling generally alters the structure of the inner regions. Without a nongravitational heat source, radiative cooling in an isothermal potential produces a central cooling flow that has $\alpha_K = 1, n_e \propto r^{-3/2}$, and $T = 4T_\infty/3$ (e.g., Voit et al. 2017; Stern et al. 2019). However, cooling of gas near the center of the potential well can also result in star formation followed by supernova explosions that generate heat.

2.2. Supernova Sweeping

Long after star formation has ceased, an aging stellar population continues to supply mass and energy to the ambient medium through normal stellar mass loss, planetary nebulae, and SNe Ia. In order for star formation to remain quenched, the galaxy must sweep out the gas shed by stars as quickly as it accumulates through some combination of supernova heating and AGN feedback (e.g., Mathews & Brighenti 2003; David et al. 2006; Voit & Donahue 2011; V15). This section shows that the entropy slope of the resulting outflow in regions where supernova heating locally exceeds radiative cooling depends primarily on the ratio $\epsilon_s/v_c^2$ relating the specific thermal energy $\epsilon_s$ of ejected stellar gas to the depth of the galactic potential well.

The radial structure of such a steady spherical flow depends on its Bernoulli specific energy $\epsilon$, defined by

$$\epsilon(r) = \frac{v_r^2}{2} + \phi(r) + \frac{5}{2} kT/\mu m_p,$$

(3)

where $v_r$ is the flow’s radial velocity, $\phi(r)$ is the gravitational potential, and the enthalpy term $(5kT/2\mu m_p)$ accounts for both the specific thermal energy and the $PdV$ work done by the flow. If radiative losses are negligible, the steady-state flow of Bernoulli energy through radius $r$ equals the integrated rate of energy input into the gas within the volume bounded by $r$. Three energy sources contribute.

1. Supernova energy. Dividing the rate of stellar heat input by the rate of stellar mass loss gives the mean specific energy $\epsilon_s$ of gas coming from stars.

2. Orbital energy. The equilibrium isotropic velocity dispersion of a singular isothermal sphere of stars is $\sigma_v = v_c/\sqrt{2}$, so thermalization of the orbital kinetic energy associated with stellar mass loss adds a specific energy $3v_c^2/4$.

3. Gravitational potential energy. The integrated stellar mass loss within $r$ adds a mean (mass-weighted) specific potential energy $\bar{\phi}(r)$.

The Bernoulli specific energy of outflowing gas at radius $r$ is therefore

$$\epsilon(r) = \epsilon_s + \frac{3}{4} v_c^2 + \bar{\phi}(r)$$

(4)

if there are no radiative losses.

As long as the outflow remains highly subsonic, its specific kinetic energy can be neglected, and its temperature is

$$kT \approx \frac{2}{5} \mu m_p \left[ \epsilon_s + \frac{3}{4} v_c^2 - (\phi - \bar{\phi}) \right].$$

(5)

In a singular isothermal sphere dominated by stellar mass, one finds $(\phi - \bar{\phi}) = v_c^2$, because $\phi(r) = v_c^2 \ln(r/r_\infty)$ and $\bar{\phi}(r) = v_c^2 \left[ \ln(r/r_\infty) - 1 \right]$. According to Equations (2) and (5), the power-law entropy slope of the outflow in that region is then

$$\alpha_K \approx \frac{5}{3} \left( \frac{\epsilon_s}{v_c^2} - \frac{1}{4} \right)^{-1},$$

(6)

demonstrating that the atmosphere’s structure depends almost entirely on the ratio $\epsilon_s/v_c^2$ (see also Sharma & Nath 2013; V15).

If the gravitational potential is negligibly close compared to the specific energy of stellar mass loss (i.e., $v_c^2 \ll \epsilon_s$), little work is required to drive the outflow. In that case, the region where stars are adding mass and energy is nearly isobaric and isentropic, with $kT \approx (2/5) \mu m_p \epsilon_s$. At the other extreme, as $v_c^2$ approaches $4\epsilon_s$, the outflow’s pressure and entropy gradients become extremely steep because the specific energy of stellar mass loss becomes incapable of driving a steady outflow.
through the isothermal potential. However, the case most relevant to quenching of isolated massive galaxies is \(v^2 \approx \epsilon_s/2\), in which the work necessary to drive the outflow is comparable to the supernova heat input, resulting in \(\alpha_K \sim 1\).

### 2.3. A Critical Entropy Slope

The primary predictions of this paper follow from the fact that galactic outflows with \(\alpha_K > 2/3\) should fundamentally differ from those with \(\alpha_K < 2/3\). At the crossover (i.e., \(\alpha_K = 2/3\)), the electron density profile of isothermal gas is \(n_e \propto r^{-1}\), meaning that radiative cooling per unit volume is \(\propto r^{-2}\). In that configuration, the ratio of radiative cooling per unit volume to stellar heating per unit volume is constant with radius within the region where the stellar mass density is \(\propto r^{-3}\).

The ratio of stellar heating to radiative cooling consequently decreases with radius in galaxies with a shallow ambient entropy slope (\(\alpha_K < 2/3\)) and rises with radius in galaxies with a steep ambient entropy slope (\(\alpha_K > 2/3\)).

Galaxies in which \(\alpha_K > 2/3\) can therefore remain in a steady state consisting of a cooling flow encompassed within a supernova-heated outflow. Those two flows diverge from a stagnation radius near \(r_{eq}\) representing both the outer boundary of the cooling flow and the inner boundary of the outflow. Mass shed by stars at \(\gtrsim r_{eq}\) flows inward, while mass shed by stars at \(\gtrsim r_{eq}\) flows outward.

Another special feature of \(\alpha_K = 2/3\) is that the ratio of cooling time to freefall time is then constant in an isothermal potential. The \(t_{cool}/t_{ff}\) ratio of the ambient medium in a galaxy with \(\alpha_K > 2/3\) consequently increases with radius, implying that the region of the ambient medium most susceptible to multiphase condensation is at small radii. One therefore expects the inner cooling flow to produce cold clouds primarily in the vicinity of the central black hole, potentially supercharging the feedback output fueled by chaotic cold accretion (Gaspari et al. 2013, 2015, 2017), without supplying molecular gas for star formation at larger radii. In contrast, the ambient medium in a galaxy with \(\alpha_K < 2/3\) is more susceptible to multiphase condensation at large radii than at small radii.

Solving Equation (6) for the circular velocity corresponding to the critical entropy slope (\(\alpha_K = 2/3\)) gives

\[
v_c \approx \left(\frac{4 \epsilon_s}{11}\right)^{1/2} \approx 340 \text{ km s}^{-1}\left(\frac{\mu m_p \epsilon_s}{2 \text{ keV}}\right)^{1/2}.
\]

The same result written in terms of an isotropic isothermal velocity dispersion is

\[
\sigma_v \approx \left(\frac{2 \epsilon_s}{11}\right)^{1/2} \approx 240 \text{ km s}^{-1}\left(\frac{\mu m_p \epsilon_s}{2 \text{ keV}}\right)^{1/2}.
\]

Observations of the specific SN Ia rate from a stellar population with an age of \(\sim 10\) Gyr indicate that it is \(3 \times 10^{-14} \text{ SN yr}^{-1} M_\odot^{-1}\) in massive elliptical galaxies (Maoz et al. 2012; Friedmann & Maoz 2018). Multiplying that rate by \(10^{51}\) erg per SN Ia (Shen et al. 2018) and dividing the result by the specific stellar mass-loss rate (\(t_s^{-1}\)) gives

\[
\epsilon_s \approx \frac{2 \text{ keV}}{\mu m_p} \cdot \frac{t_s}{200 \text{ Gyr}}.
\]

The 200 Gyr timescale chosen for scaling \(t_s\) corresponds to 0.5% of the stellar mass per Gyr and is broadly consistent with a stellar population age of \(\sim 10\) Gyr but depends in detail on the stellar initial mass function (IMF; e.g., Leitner & Kravtsov 2011). The specific SN Ia and stellar mass-loss rates are both time-dependent and somewhat uncertain, meaning that the critical velocity dispersion predicted by the model is also time-dependent and somewhat uncertain. However, the ratio of SN Ia heating to stellar mass loss changes relatively slowly with time (e.g., Ciotti et al. 1991), meaning that the time dependence of the critical velocity dispersion should also be rather mild (but see Section 5.3).

### 2.4. Heating/Cooling Equality

Radiative cooling per unit volume equals supernova heating per unit volume at any radius at which

\[
\left(\epsilon_s + \frac{3}{2} \sigma_s^2 \right) \rho_s \approx n_e n_p \Lambda(T),
\]

where \(\rho_s\) is the stellar mass density and \(\Lambda(T)\) is the usual radiative cooling function, defined with respect to the proton density \(n_p\). Solving for \(P\) gives the pressure profile

\[
P_{eq}(r) \equiv \left(\epsilon_s + \frac{3}{2} \sigma_s^2\right) \left(\frac{n_e^2}{n_e n_p} \frac{\rho_s}{T_0 \Lambda(T)}\right)^{1/2},
\]

along which radiative cooling would equal supernova heating, given the temperature \(T\). For \(\sigma_v \approx 240 \text{ km s}^{-1}\), supernova sweeping without radiative cooling gives \(\alpha_K \approx 2/3\) and \(kT \approx 0.75 \text{ keV}\). The critical profiles of pressure, electron density, and entropy in solar metallicity gas near this temperature are

\[
P_{eq}(r) \approx (1.4 \times 10^{-10} \text{ erg cm}^{-3}) \sigma_v^2 \frac{n_e}{n_p},
\]

\[
n_{eq, \text{rn}} \approx (0.06 \text{ cm}^{-3}) \sigma_v \frac{n_e}{n_p},
\]

\[
k_{eq}(r) \approx (5 \text{ keV cm}^2) \sigma_v \frac{n_e}{n_p},
\]

given \(r_{eq} \equiv r/(1 \text{ kpc})\) and \(\sigma_v \approx 240 \text{ km s}^{-1}\), an isothermal stellar mass distribution (\(\rho_s = \sigma_v^2/2\pi G r^2\)), and the fiducial values \(\mu m_p \epsilon_s \approx 2 \text{ keV}\) and \(t_s \approx 200 \text{ Gyr}\) if the weak dependence of \(\Lambda(T)\) on \(\sigma_v\) is ignored.

Figure 1 shows how the ratio of estimated stellar heating to radiative cooling depends on radius for 10 massive ellipticals with high-quality Chandra X-ray observations from Werner et al. (2012, 2014). The interstellar metallicity in each of them is observed to be approximately solar, so we have computed their radiative cooling rates assuming solar metallicity throughout the paper.11 Our estimator for stellar heating per unit volume is the left-hand side of Equation (10), assuming \(\epsilon_s = 2 \text{ keV}/\mu m_p\), \(t_s = 200 \text{ Gyr}\), and \(\rho_s = \sigma_v^2/2 \pi G r^2\), with \(\sigma_v\) for each galaxy from Hyperleda. It overestimates stellar heating.

---

10 The inferred value of \(\epsilon_s\) is more robust to assumptions about the IMF than either the specific stellar mass-loss rate or the specific SN Ia rate, as long as the same IMF is assumed. Here the quoted SN Ia rate is relative to the stellar mass of the initial stellar population, which was assumed to have a Kroupa (2001) IMF. For a similar IMF, the fits of Leitner & Kravtsov (2011) give \(t_s = 217 \text{ Gyr}\), relative to the initial stellar mass, at a stellar population age of 10 Gyr.

11 This assumption is motivated by abundance observations and ignores the fact that the mean iron abundance of the gas ejected from old stars should be several times solar, given the observed SN Ia rate. Consequently, the iron coming from recent SNe Ia must not be well mixed with the gas coming from the rest of the stars, but the reason remains mysterious. (See V15 for a brief discussion.)
at radii beyond where the stellar mass density starts to decline more steeply than $\rho_* \propto r^{-2}$, but for now, we are most interested in the range $0.5 \, \text{kpc} \lesssim r \lesssim 2 \, \text{kpc}$, where all of the estimated heating/cooling ratios are of order unity.

The general proximity of these heating/cooling ratio profiles to unity is remarkable. Naively, one might expect gas below the line of cooling to cool further, becoming increasingly compressed and cooling-dominated. Likewise, heating of gas above the line of equality should cause it to expand and become more strongly dominated by heating. The nearness of these observed profiles to the line of cooling equality therefore indicates that a self-regulation mechanism involving AGN feedback maintains these galaxies near the balance point between stellar heating and radiative cooling. Also notable is the fact that the heating/cooling ratio profiles beyond $\sim 1 \, \text{kpc}$ for galaxies without extended multiphase gas (left panel) tend to rise from $r \sim 1$ to 10 kpc, while the profiles of ellipticals that do contain extended multiphase gas (right panel) tend either to decline or to remain flat.

2.5. Boundary Pressure and Cooling Flows

In steady subsonic solutions for thermally driven, mass-loaded outflows, the normalization of pressure and density depend on the gas pressure at the outer boundary. For simplicity, let $r_b$ be a boundary radius inside of which stellar heating determines the outflow’s entropy profile. Setting the bounding gas pressure at $r_b$ to $P_b$, then leads to the profiles

$$P_{\text{out}}(r) \approx P_b \left( \frac{r}{r_b} \right)^{-3\alpha_K / 2},$$

$$kT_{\text{out}}(r) \approx kT_b,$$

$$n_{e,\text{out}}(r) \approx P_b \left( \frac{\mu}{\mu_e} \right) \left( \frac{r}{r_b} \right)^{-3\alpha_K / 2},$$

where $\alpha_K$ is determined by Equation (6), $\mu, \mu_e$ is the mean mass per electron, and $kT_b = 2\mu m_p v_T^2 / 3\alpha_K$.

In galaxies with $\sigma_v \approx 240 \, \text{km s}^{-1}$, these supernova-sweeping profiles run parallel to the critical profiles for heating–cooling balance. Whether or not the region inside of $r_b$ lies either entirely above or entirely below the critical profile $P_{\text{eq}}(r)$ depends directly on the boundary pressure (see Figure 2). The AGN feedback in a galaxy like this can shut off an galaxy-wide cooling flow simply by heating and lifting its CGM until $P(r_b) < P_{\text{eq}}$. After that happens, the shutdown of the cooling flow may not be permanent, because ejected stellar mass swept out of the galaxy continues to accumulate in the CGM. Unless a nonstellar energy source continually heats, strips, or lifts the accumulating gas, the CGM pressure will gradually increase until $P(r_b) > P_{\text{eq}}$ at which time the resulting cooling flow needs to trigger another feedback episode that lowers $P_b$.

2.6. The Valve

The mechanism depicted in Figure 2 closely links the strength of AGN feedback with the confining CGM pressure. It is essentially a switch that turns on a central cooling flow when the CGM pressure is high and turns the cooling flow off when the CGM pressure is low. The onset of a cooling flow presumably triggers AGN feedback, but the feedback response does not necessarily need to limit cooling by heating gas at small radii. Instead, it can limit cooling at small radii by lowering the CGM pressure, which allows the ambient gas at 1–10 kpc to expand, temporarily making local stellar heating more competitive with local radiative cooling. The AGN feedback in a galaxy with this feature will tend to drive the boundary pressure toward the locus of heating/cooling equality, so that $P_b \sim P_{\text{eq}}(r_b)$.

In galaxies with $\sigma_v \gtrsim 240 \, \text{km s}^{-1}$, this regulation mechanism becomes more like an adjustable valve than a switch. Figure 3 shows how the valve works. An outflow driven by stellar heating in a galaxy with this larger velocity dispersion has a
density profile \(n_{e,\text{eq}}\) that is steeper than the locus of heating/cooling equality \(n_{e,\text{eq}}\). Those profiles intersect at a particular radius \(r_{\text{eq}}\). Inside of that radius is a cooling flow, and outside of it is an outflow driven by stellar heating. The mass inflow rate of the cooling flow \(M_{\text{c}}\) is equivalent to the combined stellar mass-loss rate of all stars within \(r_{\text{eq}}\), which is determined by the confining CGM pressure. Therefore, the time-averaged rate at which the galaxy supplies gas to its central black hole is a continuous function of the CGM pressure, with greater CGM pressure resulting in a greater fueling rate. The outcome is a mechanism capable of tuning itself so that kinetic AGN feedback regulates the confining CGM pressure to remain near a time-averaged level consistent with the required AGN fueling rate.

The mechanism requires kinetic energy output from the AGN to offset three different effects that act to increase CGM pressure.

1. **CGM cooling.** Radiative losses from the CGM around isolated galaxies with \(\sigma_s \gtrsim 240 \text{ km s}^{-1}\) are typically \(10^{41} \text{ erg s}^{-1} \lesssim L_X \lesssim 10^{42} \text{ erg s}^{-1}\). The time-averaged energy supply from the AGN must be at least this large in order to prevent CGM pressure from increasing because of a gradual decline in entropy and a gradual increase in density.

2. **Stellar gas ejection.** As mentioned in Section 2.5, supernova heating alone cannot unbind ejected stellar gas from the dark matter halo around a galaxy with \(\sigma_s \gtrsim 240 \text{ km s}^{-1}\). Continual lifting to large altitude therefore requires an additional power input of \(\sim 4.6M_{\odot}v_e^2/\tau_s\) from the AGN, corresponding to \(\sim 10^{41.5} \text{ erg s}^{-1} (M_{\odot}/10^{11.3} M_{\odot}) \sigma_{240}^2\).

3. **Cosmological gas accretion.** The time-averaged cosmological infall rate of gas into the CGM is \(\sim f_{\text{infl}}M_{\text{halo}}H_0\), which substantially exceeds the mass input from stars. Lifting of that CGM gas within the halo’s gravitational potential \(^{14}\) requires the AGN to supply a time-averaged power of \(\sim 10^{43} \text{ erg s}^{-1} (M_{\text{halo}}/10^{13} M_{\odot}) \sigma_{240}^2\).

The most demanding job for AGN feedback in this halo mass range is therefore to lift the CGM gas that accumulates through cosmological accretion. Integrated over cosmic time, the total amount of energy required to alleviate the galaxy’s boundary pressure by continually lifting accreting CGM gas is \(\sim f_{\text{infl}}M_{\text{halo}}\sigma_{71}^4\), equal to \(\sim 10^{60.3} \text{ erg} (M_{\text{halo}}/10^{13} M_{\odot}) \sigma_{240}^2\). During the same time period, the central black hole grows to have a rest-mass energy \(\sim (10^{43} \text{ erg}) \sigma_{4.4}^{14}\), based on observations indicating that \(M_{\text{BH}} \sim (7 \times 10^8 M_{\odot}) \sigma_{240}^{1.4}\) \((\text{Kormendy \\& Ho 2013})\). Tapping less than 1% of the central black hole’s rest-mass energy as it grows therefore suffices to lift the CGM (see also Section 5.5).

### 2.7. Multiphase Circulation

Star formation in a galaxy resembling the left panel of Figure 3 is effectively quenched because the outflow that carries away stellar ejecta remains homogeneous and the inner cooling flow produces less than 1 \(M_{\odot} \text{ yr}^{-1}\) of condensed gas. However, outflows from galaxies with \(\sigma_s \lesssim 240 \text{ km s}^{-1}\) are unlikely to be stable to multiphase condensation. Figure 4 illustrates the problem. A heated outflow of stellar ejecta through a shallow potential well tends to have a density slope that is shallower than \(n_{e,\text{eq}}\) and therefore passes from a locally heating-dominated state to a locally cooling-dominated state at \(r_{\text{eq}}\). An entropy inversion then develops outside of \(r_{\text{eq}}\) on a timescale of \(\sim t_{\text{cool}}(r_{\text{eq}})\). That inversion is unstable to convection, which allows low-entropy gas to condense out of the outflow and rain back toward the galaxy’s center. Multiphase

\(^{13}\) The factor of 4.6 approximates the work required to lift gas in an isothermal potential to ~100 times its original altitude or, equivalently, to unbind gas originally at the bottom of an NFW potential well.

\(^{14}\) Heating that adds a specific energy \(\sigma_{71}^2\) to CGM gas in an isothermal potential without changing the slope of its entropy profile lifts nearly hydrostatic gas by a factor of \(e^{\sigma_{71}^2} \approx 1.6\) in radius and lowers its density by a factor of ~5.
circulation therefore develops and can resupply the galaxy with cold clouds capable of forming stars.

More generally, the ability of a galactic atmosphere to remain homogeneous depends heavily on its ambient $t_{\text{cool}}/t_{\text{ff}}$ ratio (e.g., Hoyle 1953; Nulsen 1986; Balbus & Soker 1989; Binney et al. 2009; McCourt et al. 2012; Sharma et al. 2012b; Choudhury & Sharma 2016; Thompson et al. 2016; Voit et al. 2017; Choudhury et al. 2019). In the conventional definition of this ratio, the cooling time is $t_{\text{cool}} \equiv 3P/(2\pi\rho n_g \Lambda(T))$, where $\Lambda(T)$ is the usual radiative cooling function, and the freefall time is $t_{\text{ff}} \equiv (2g/r)^{1/2}$, where $g$ is the local gravitational acceleration. A static, thermally balanced medium can, in principle, be stable to multiphase condensation if $t_{\text{cool}}/t_{\text{ff}} \gtrsim 1$, because buoyancy effects enabled by entropy stratification suppress condensation (Cowie et al. 1980; Nulsen 1986; Binney et al. 2009; McCourt et al. 2012). In practice, however, media with $1 \lesssim t_{\text{cool}}/t_{\text{ff}} \lesssim 10$ remain highly susceptible to multiphase condensation because subsonic disturbances can interfere with the condensation-damping effects of buoyancy (Sharma et al. 2012b; Gaspari et al. 2013; Voit et al. 2017), particularly if those disturbances flatten or invert the entropy gradients that are responsible for buoyancy (McNamara et al. 2016; Voit et al. 2017; Choudhury et al. 2019).

The result is a pervasive upper limit on the pressure and density of ambient gas corresponding to $\text{min}(t_{\text{cool}}/t_{\text{ff}}) \approx 10$, which we will call the precipitation limit (Sharma et al. 2012a; Voit et al. 2015b, 2018, 2019; Hogan et al. 2017; Babyk et al. 2018; Voit 2019). Galactic atmospheres tend to have $t_{\text{cool}}/t_{\text{ff}} \gtrsim 10$ because ambient gas with a shorter cooling time cannot persist indefinitely without forming condensates that fuel feedback. As the cooling time rises, so that $t_{\text{cool}}/t_{\text{ff}}$ grows from $\sim 10$ to $\sim 20$, increasingly large hydrodynamical disruptions are required to offset the buoyancy effects that suppress multiphase condensation (Voit et al. 2017; Voit 2018). And if $t_{\text{cool}}/t_{\text{ff}} \gtrsim 20$, then multiphase condensation cannot occur without either large entropy perturbations (Choudhury et al. 2019), strong magnetic fields (Ji et al. 2018), or enough rotational support to offset the stabilizing effects of buoyancy (Gaspari et al. 2015; Sobacchi & Sormani 2019; Sormani & Sobacchi 2019). A subsonic outflow is therefore likely to remain stable to multiphase condensation if it has (1) an entropy gradient with $\alpha_K \gtrsim 2/3$, (2) $t_{\text{cool}}/t_{\text{ff}} \gtrsim 20$ at all radii, and (3) a lack of large entropy perturbations, strong magnetic fields, or significant rotational support.

The right panel of Figure 4 shows how $n_e \text{eq}$ relates to $t_{\text{cool}}/t_{\text{ff}}$ in a galaxy with $\sigma_e \approx 200 \text{ km s}^{-1}$. A green line shows the outflow solution ($n_e \text{eq}$) obtained by requiring heating to equal cooling at $10 \text{ kpc}$. As the outflow approaches that radius, it is already at $t_{\text{cool}}/t_{\text{ff}} < 20$ and approaching $t_{\text{cool}}/t_{\text{ff}} \approx 10$. One therefore expects the outflow to become unstable to multiphase condensation in the neighborhood of $r_{\text{eq}}$. In other words, star formation in lower-mass galaxies is more difficult to quench with feedback, because the outflows that feedback generates tend to destabilize the CGM. Similarly, a cooling flow starting at $r_{\text{eq}}$ becomes increasingly susceptible to multiphase condensation as it moves inward because its density slope ($n_e \text{eq} \propto r^{-1.5}$) is steeper than the lines indicating constant $t_{\text{cool}}/t_{\text{ff}}$. The ambient $t_{\text{cool}}/t_{\text{ff}}$ ratio therefore decreases as the flow moves inward, and the flow cannot remain homogeneous within the radius at which $t_{\text{cool}} \approx t_{\text{ff}}$.

2.8. Jet Propagation

In order for the black hole feedback valve to operate as described, most of the AGN feedback energy produced over cosmic time needs to be thermalized at radii larger than where most of the galaxy’s stars reside. Morphological observations of strong radio jets indicate that they can indeed propagate as narrow streams beyond most of the stars. For example, NGC 4261 (represented by purple stars in Figure 1) has radio jets that are narrow out to $r > 10 \text{ kpc}$ and terminate in lobes at $r \sim 25–45 \text{ kpc}$, consistent with a power output $> 10^{43} \text{ erg s}^{-1}$ (O’Sullivan et al. 2011). Meanwhile, that galaxy’s profiles of gas density, temperature, and entropy at $r < 10 \text{ kpc}$ remain nearly identical to those of the other galaxies with $\sigma_e \sim 300 \text{ km s}^{-1}$, even though its jets are unusually powerful.

Propagation to such a distance requires the momentum flux in the jets to be at least as great as the ambient pressure, which
is \( \approx 10^{-11} \) erg cm\(^{-3}\) at \( \approx 10 \) kpc in NGC 4261 (O'Sullivan et al. 2011). For comparison, the relativistic momentum flux corresponding to a jet power \( \dot{E}_{\text{jets}} \times 10^{42} \) erg s\(^{-1}\) is

\[
3 \times 10^{-12} \text{ erg cm}^{-3} \frac{\Omega_{\text{jets}}}{(100)} \left( \frac{r}{10 \text{ kpc}} \right)^2,
\]

where \( \Omega_{\text{jets}} \) is the solid angle in steradians that the jets subtend at radius \( r \). Observations indicate that the jet opening angle at \( \approx 10 \) kpc in NGC 4261 is \( \approx 12^\circ \) (Nakahara et al. 2018), implying that the combined solid angle of both jets is \( \approx 0.3 \) sr and a jet power \( \approx 10^{44} \) erg s\(^{-1}\) is needed to drill through the gas at \( r < 10 \) kpc.

Jet power in the other galaxies shown in Figure 1 is far smaller, typically \( \approx 10^{41} - 10^{42} \) erg s\(^{-1}\) (Werner et al. 2014). Equation (19) implies that those jets are currently incapable of propagating at relativistic speeds to distances \( \gtrsim 10 \) kpc unless they are extremely narrow, and their observed morphologies indicate that they do not extend to \( \gtrsim 10 \) kpc as narrowly collimated outflows. Instead, the jets thermalize their kinetic energy and inflate bubbles at smaller radii as they decelerate, sometimes driving shocks that propagate into the supernova-heated outflow. According to V15, those shocks should be relatively weak, each imparting an entropy jump

\[
\Delta K_{\text{jets}} \approx (2.8 \text{ keV cm}^2) \left( \frac{\dot{E}_{\text{jets}}}{10^{42}} \right) \left( \frac{r}{10 \text{ kpc}} \right)^{-2} \left( \frac{\Delta_{3e-23}}{3 \times 10^{-23} \text{ erg cm}^{-3} \text{ s}^{-1}} \right)^{4/3} \left( \frac{\sigma_{240}}{10^4} \right)^{4/3} \left( \frac{\sigma_{240}}{10^4} \right)^{-4/3}
\]

with \( \Delta_{3e-23} \equiv \Lambda/(3 \times 10^{-23} \text{ erg cm}^3 \text{ s}^{-1}) \), as they propagate through an atmosphere near the precipitation limit. Shocks driven by \( \approx 10^{42} \) erg s\(^{-1}\) of power become supersonic at \( \approx 10 \) kpc. Beyond that radius, the hot bubbles that drove the shocks are expected to buoyantly rise and eventually mix with the ambient medium, thermalizing much of the original AGN power at greater altitudes (e.g., Churazov et al. 2001; Birzan et al. 2004; Voit & Donahue 2005). As a result, AGN feedback in massive elliptical galaxies does not substantially alter the entropy profile slope established within \( \lesssim 10 \) kpc by a supernova-heated outflow, except at \( \lesssim 1 \) kpc, where the observed core entropy is comparable to \( \Delta K_{\text{jets}} \) (see V15 and Section 4).

3. Steady Flow Solutions

Section 2 considered two energy sources with the potential to offset radiative cooling and limit star formation. Heating by SNe Ia is relatively steady and has a known spatial distribution but is not energetically capable of pushing ejected stellar gas all the way out of the galaxy’s potential well. On the other hand, AGN feedback can provide enough energy to push away all of the gas associated with a massive galaxy but is intermittent in time, with a poorly known spatial distribution. This section therefore considers the steady flow patterns produced by SN Ia heating alone. Section 4 then follows up by attempting to infer the contributions of AGN feedback from the observed discrepancies between the steady flow solutions and observations of the atmospheres of real galaxies.

The fluid equations for steady one-dimensional radial flow are

\[
\frac{1}{r^2} \frac{d}{dr} \left( \rho v_r r^2 \right) = \dot{\rho}(r),
\]

\[
\frac{1}{r^2} \frac{d}{dr} \left( \rho v_r^2 r^2 \right) = -\frac{dP}{dr} - \rho \frac{d\dot{\rho}}{dr},
\]

\[
\frac{1}{r^2} \frac{d}{dr} (e \rho v_r r^2) = \mathcal{H}(r) + \dot{\rho}(r) \dot{\phi}(r).
\]

In a galactic environment, \( \dot{\rho}(r) = \rho_m(r)/t_\star \) is a source term for stellar mass loss, and \( \mathcal{H}(r) \) is the net stellar heating rate per unit volume. The standard integration method for steady galactic flows is to choose a stagnation radius \( r_i \) and integrate away from it in both directions, iteratively seeking a solution that passes smoothly
through a cooling-flow sonic point at small radii (e.g., David et al. 1987; Tabor & Binney 1993). Physically, the steady-state flow configuration and its stagnation radius are determined by the CGM boundary pressure, along with an implicit assumption that some mechanism other than stellar heating keeps the pressure at $r_b$ constant by removing the gas that flows to greater radii.

Figure 5 illustrates how steady flow solutions without AGN heating depend on a galaxy’s central stellar velocity dispersion $\sigma_v$. A teal line in each panel shows the entropy profile produced around the galaxy by pure cosmological structure formation. A thick dashed tan line shows the same profile displaced upward in entropy so that it is continuous with $K_{\text{out}}$. A thin violet line shows the entropy level $K_{\text{eq}}$ at which radiative cooling would balance SN Ia heating. Cooling exceeds SN Ia heating in the blue region below the violet line. A thick solid magenta line shows the precipitation limit at $t_{\text{cool}}/t_\text{ff} \approx 20$. The pink region between those lines corresponds to intermediate values of $t_{\text{cool}}/t_\text{ff}$ that may be susceptible to multiphase condensation. Thin horizontal gray lines indicate where $t_{\text{cool}} = 100$ Myr, 1 Gyr, and 10 Gyr, as labeled. A thick purple line from 0.04 to 4 kpc in each panel shows the entropy jump ($\Delta K_{\text{D}}$) that a spherical shock driven by a $10^{43}$ erg s$^{-1}$ outflow would produce in a medium at the precipitation limit. Vertical black dotted lines in each of the upper two panels show the maximum radius $r_{\text{max}}$ to which stellar heating can drive ejected stellar gas.

Adding an amount of AGN heating comparable to stellar heating and with a similar spatial distribution would tend to decrease the outflow’s entropy slope $\sigma_{K_{\text{out}}}$.15

---

15: The Astrophysical Journal, 899:70 (23pp), 2020 August 10 Voit et al.
profile corresponding to a steady flow solution determined by only two parameters: $\sigma_*$ and either $r_0$ or the boundary pressure $P_b(r_b)$. The figure shows steady flow solutions with $t_{\text{cool}} \sim 1$ Gyr at 10 kpc because X-ray observations of massive elliptical galaxies typically indicate $t_{\text{cool}} \sim 0.5$ to $\sim 2$ Gyr at 10 kpc (Lakhchaura et al. 2018). In the four panels representing galaxies with $\sigma_* \gtrsim 230$ km s$^{-1}$, the steady solution shown consists of a cooling flow inside of $r_0$ and a supernova-heated outflow outside of $r_0$. A kink near $r_{\text{eq}}$ indicates the location of the stagnant region, and the local entropy minimum in that region is at $r_0$. Inside of $r_0$, the steady flow solutions in those four panels resemble pure cooling flows, since they have $K \propto r$, as appropriate for a steady cooling flow in an isothermal potential. At larger radii ($r > r_0$), the steady flow solutions resemble supernova-heated outflow solutions, with $\alpha_K$ close to the prediction from Equation (6), meaning that the outflow’s entropy slope increases as $\sigma_*$ increases.

In galaxies with $\sigma_* \lesssim 240$ km s$^{-1}$, the steady outflow solutions develop unsustainable entropy inversions after passing from the heating-dominated region of the $r$-$K$ plane into the cooling-dominated region. As explained in the analysis of Section 2, the outflow solutions for those galaxies have entropy slopes too shallow to remain in the heating-dominated region. After the outflow solution enters the cooling-dominated region, its entropy profile starts to decline, implying that the flow has become unstable to convection. As the entropy continues to decline with radius, the flow encounters the solid magenta line marking $t_{\text{cool}} / t_{\text{H}} \approx 10$. Near that point, the flow should become unstable to multiphase condensation, enabling low-entropy gas blobs to precipitate out of the flow and start sinking inward. In galaxies with these properties, outflows driven by stellar heating therefore drive multiphase circulation instead of homogeneously expelling the gas being shed by stars.

The remainder of this section interprets the solutions illustrated in Figure 5 in more detail. First, it outlines the galaxy model used to specify the gravitational potential and the source terms for mass and energy input, which depend only on $\sigma_*$. Then, it discusses each panel of the figure. Readers more interested in comparisons with observations may wish to skip to Section 4.

### 3.1. Generic Galaxy Model

The generic galaxy model used in the steady flow integrations assumes that the galaxy’s halo has a Navarro–Frenk–White (NFW) density profile, $\rho_M \propto (r/r_s)^{-1}(1 + r/r_s)^{-2}$, with a maximum circular velocity $v_{\text{c},\text{max}} = \sqrt{2} \sigma_*$ and a scale radius $r_s = 0.1 r_{200}$, where $r_{200}$ is the radius encompassing a mean mass density 200 times the cosmological critical density. It also assumes that the distribution of stellar mass density follows a modified Einasto profile with

$$d \ln \rho_* = -2 \left( \frac{r}{r_{*2}} \right)^{1/\mu_{E}}$$

for $r \geq r_{*2}$ and $\rho_*(r) \propto r^{-2}$ for $r < r_{*2}$. The index describing the outer stellar envelope is set to $n_{E} = 4$, the stellar mass density profile is normalized so that $\rho_s(r_{*2}) = \sigma_*^2 / 2\pi G r_{*2}^2$, and its scale radius $r_{*2}$ is determined by setting the two-dimensional effective radius of the galaxy equal to $0.015 r_{200}$, in approximate agreement with the mean relation observed among galaxies (Kravtsov 2013). This prescription results in an overall potential well with a circular velocity that is nearly constant with radius, remaining within 10% of $v_c = \sqrt{2/3} \sigma_*$ out to 0.75 $r_{200}$, in alignment with the assumptions of Section 2. However, comparable galaxies at the centers of higher-mass groups and clusters sit in potentials with greater maximum circular velocity. Alleviating the boundary pressure around those galaxies by lifting the CGM therefore requires substantially more AGN feedback energy (see Section 5.4).

### 3.2. Description of Solutions

#### 3.2.1. $\sigma_* = 300$ km s$^{-1}$

The most massive generic galaxy shown in Figure 5 has a velocity dispersion $\sigma_* = 300$ km s$^{-1}$. Its total stellar mass is $10^{11.8} M_\odot$, and its total mass within $r_{200}$ is $M_{200} \approx 10^{13.2} M_\odot$. In this potential well, stellar heating alone cannot push ejected stellar gas beyond $v_{\text{max}} \approx 70$ kpc. The steady flow solution depicted for this galaxy is determined by the boundary condition $r_0 = 2.5$ kpc, so that $r_{\text{eq}} \approx 1$ kpc, in alignment with Figure 1. Gas between $r_{\text{eq}}$ and $r_0$ is experiencing net heating but is flowing inward because it rests upon a cooling flow that is moving inward. The cooling-flow rate reaches $M_{\text{cf}} \approx 0.5 M_\odot$ yr$^{-1}$ at small radii because the stellar mass within $r_0$ is $10^{11} M_\odot$, and the specific stellar mass-loss rate is $(200$ Gyr)$^{-1}$. The figure shows that the entropy slope of the inner cooling flow is nearly identical to the $K \propto r$ expectation (thick blue line) for a cooling flow in an isothermal potential.

Outside of $r_0$, Equation (6) predicts an entropy slope $\alpha_K = 1.1$, shown by the thick green line, and the steady outflow solution there has a similar slope. It therefore climbs progressively higher into the heating-dominated region as it moves outward, reaching a cooling time of $\sim 2$ Gyr at $r_b = 10$ kpc, where its pressure is $P_b \approx 3 \times 10^{-12}$ erg cm$^{-3}$. Beyond $r_b$, the thick dashed tan line shows a continuation of $K(r)$ with a cosmological entropy slope $K \propto r^{1.1}$ and a normalization $\sim 20$ times the entropy profile expected from pure cosmological structure formation (thick solid tan line). This steady flow solution therefore requires AGN feedback or some other mechanism to reduce the CGM pressure at $r_b$ by a factor of $\sim 10^{-2}$ (because $P \propto K^{-3/2}$ at a given $T$).

#### 3.2.2. $\sigma_* = 280$ km s$^{-1}$

The second most massive galaxy shown in Figure 5 has a velocity dispersion $\sigma_* = 280$ km s$^{-1}$, a total stellar mass of $10^{11.7} M_\odot$, and $M_{200} \approx 10^{13.1} M_\odot$. A boundary condition $r_0 = 2$ kpc determines the steady flow solution in the figure. Stellar heating cannot push ejected stellar gas beyond $v_{\text{max}} \approx 90$ kpc, and the inner cooling-flow rate reaches $M_{\text{cf}} \approx 0.4 M_\odot$ yr$^{-1}$.

Again, the entropy slope of the outflow beyond $r_{\text{eq}}$ is similar to the prediction of Equation (6), which in this case is $\alpha_K = 0.9$. This slope is steep enough to remain in the heating-dominated region of the $r$-$K$ plane out to tens of kpc. The flow’s cooling time increases through $\sim 2$ Gyr at $r_b = 10$ kpc, where its pressure is $P_b \approx 3 \times 10^{-12}$ erg cm$^{-3}$. As in the $\sigma_* = 300$ km s$^{-1}$ case, the CGM pressure at $r_b$ must be reduced by a factor of $\sim 10^{-2}$ below its cosmological value in order for the solution to be valid.
The galaxy with a critical velocity dispersion ($\sigma_r = 240 \text{ km s}^{-1}$) has a total stellar mass of $10^{11.53} M_\odot$ and a total halo mass $M_{200} \approx 10^{12.89} M_\odot$. Setting $r_0 = 1.5 \text{ kpc}$ determines the steady flow solution shown. Stellar heating cannot push ejected stellar gas beyond $r_{\text{max}} \approx 160 \text{ kpc}$, and the inner cooling-flow rate reaches $M_{\text{cf}} \approx 0.2 M_\odot \text{ yr}^{-1}$. Just beyond $r_0$, the unique requirements of a steady-state solution cause the entropy profile to climb steeply out to $\approx 2 \text{ kpc}$, but the outflow solution from $2$ to $10 \text{ kpc}$ has a slope similar to the prediction $\alpha_K = 2/3$ from Equation (6). In order for this solution to be valid, the CGM pressure at $r_b$ must be reduced by a factor of $\approx 10^{-2}$ below its cosmological value.

Beyond $10 \text{ kpc}$, something interesting happens. The steady outflow solution crosses back into the cooling-dominated region of the $r$–$K$ plane. Its entropy slope therefore becomes progressively shallower until it starts to decline. The resulting entropy inversion is unstable to convection, meaning that a steady homogeneous outflow cannot be sustained. Lower-entropy gas at larger radii cannot be hydrostatically supported and becomes unstable to multiphase condensation. Multiphase circulation near $r_b$ consisting of low-entropy gas clouds descending through a higher-entropy outflow is therefore an inevitable outcome.

### 3.2.4. $\sigma_r = 230 \text{ km s}^{-1}$

One more step down in $\sigma_r$ results in a solution with a more pronounced entropy inversion. This galaxy has a total stellar mass of $10^{11.45} M_\odot$ and a total halo mass $M_{200} \approx 10^{12.83} M_\odot$. The steady flow solution shown has $r_0 = 1 \text{ kpc}$. Stellar heating cannot push ejected stellar gas beyond $r_{\text{max}} \approx 200 \text{ kpc}$, and the inner cooling-flow rate reaches $M_{\text{cf}} \approx 0.1 M_\odot \text{ yr}^{-1}$. Between $1.5$ and $10 \text{ kpc}$, the outflow solution has a slope similar to the prediction $\alpha_K = 0.6$ from Equation (6), but it flattens and inverts after passing into the cooling-dominated region. The entropy profile of the outflow solution then plunges steeply, as the cooling time becomes shorter than the flow time. In the context of the formal solution, the entropy drop leads to a large increase in the density of the smooth outflow. Physically, the outflow becomes convectively unstable, fragmenting into cold, dense clouds that then sink inward. In other words, multiphase circulation is inevitable.

### 3.2.5. $\sigma_r = 220 \text{ km s}^{-1}$

Going down to $\sigma_r = 220 \text{ km s}^{-1}$ yields a steady-state solution without an inner cooling flow. The galaxy in this panel has a total stellar mass of $10^{11.41} M_\odot$ and a total halo mass $M_{200} \approx 10^{12.77} M_\odot$. In this potential well, stellar heating is capable of pushing ejected stellar gas out to $r_{\text{max}} \approx 260 \text{ kpc}$ in the absence of radiative cooling, but not if the CGM pressure is significant. Unlike the other steady solutions in Figure 5, this one has no stagnation point or inner cooling flow. The difficulty is that the characteristic entropy slope of the heated outflow ($\alpha_K \approx 0.55$) is significantly smaller than the slope on the boundary between the heating- and cooling-dominated regions of the $r$–$K$ plane. The most natural steady flows are therefore heated outflows at small radii that change into cooling outflows after crossing that boundary. Out to $\approx 10 \text{ kpc}$, the outflow solution shown has a slope similar to the prediction $\alpha_K = 0.55$ from Equation (6) before it flattens and inverts beyond $10 \text{ kpc}$. At $r_b = 20 \text{ kpc}$, the boundary pressure is $P_b = 8.5 \times 10^{-13} \text{ erg cm}^{-3}$, about $0.02$ times the expected cosmological CGM pressure at that radius. With this boundary condition, the steady outflow solution has $t_{\text{cool}} \approx 1.5 \text{ Gyr}$ at $10 \text{ kpc}$ and crosses the locus of heating–cooling equality at $r_{\text{eq}} \approx 14 \text{ kpc}$. Beyond that point, radiative cooling flattens and inverts the entropy profile, preventing the outflow from going beyond $40 \text{ kpc}$. Multiphase circulation within that radius is therefore unavoidable.

### 4. Comparisons with X-Ray Observations

This section compares X-ray observations of massive elliptical galaxies with the estimates of Section 2 and some steady flow solutions similar to those in Section 3 in order to assess the validity of the overall model and evaluate how AGN feedback alters the steady flows that stellar mass loss would otherwise produce. The following subsection summarizes the general trends, and the next one comments on comparisons with individual galaxies. Readers more interested in the model’s implications for quenching of star formation may wish to skip to Section 5.

#### 4.1. General Trends

Our comparison sample of massive elliptical galaxies consists of the same $10$ galaxies from Werner et al. (2012, 2014) analyzed by V15. Five have extended multiphase gas beyond $1 \text{ kpc}$, and the other five are single phase outside of $1 \text{ kpc}$.

#### 4.1.1. Single-phase Ellipticals

Figure 6 shows that the entropy profiles of the five single-phase galaxies, which all have $\sigma_r > 260 \text{ km s}^{-1}$, are consistent with steady-state outflow solutions having $r_{\text{eq}} \approx 1 \text{ kpc}$ and extending out to $\gtrsim 10 \text{ kpc}$. The observed entropy slopes in this radial interval agree well with the predictions of Equation (6), supporting the hypothesis that supernova sweeping is the primary mechanism for pushing ejected stellar gas out of these galaxies. Section 3 shows that the CGM pressure must be reduced to $\lesssim 1\%$ of the cosmological value in order for supernova sweeping to succeed, and the observed profiles at $\sim 10$–$100 \text{ kpc}$ are consistent with that requirement.

Inside of $1 \text{ kpc}$, the entropy profiles of all but one of the single-phase ellipticals are inconsistent with a pure central cooling flow for two reasons. First, they are flatter than the $K \propto r$ profile expected of pure cooling flows in these potential wells. Second, they remain in the region of the $r$–$K$ plane in which stellar heating exceeds radiative cooling. This configuration cannot persist indefinitely without AGN energy input, because the stellar source terms in regions where the pressure gradient is shallow favor a buildup of gas density over steady expansion. The ejected stellar gas within $1 \text{ kpc}$ is therefore trapped, meaning that its pressure and density are destined to increase and will eventually initiate a central cooling flow unless the AGN can lift the overlying gas and relieve the pressure.

Intermittent bursts of kinetic AGN feedback with a power of $\sim 10^{42} \text{ erg s}^{-1}$ can lift the ejected stellar gas and will also drive shocks capable of producing the observed central entropy plateaus (see also V15). Thick horizontal purple lines in Figures 5 and 6 show the entropy jump $\Delta K_{\Delta/2}$ predicted by Equation (20) for a shock propagating through ambient gas at the precipitation limit and driven by $10^{42} \text{ erg s}^{-1}$ of feedback.
Figure 6. Comparisons of model predictions with data. Symbols represent entropy profile data from Werner et al. (2012, 2014). Thin dashed and dotted-dashed lines represent entropy profile data from other sources described in Section 4. All other figure elements represent the same quantities as in Figure 5. The left column presents the massive ellipticals without multiphase gas beyond \( \sim 1 \) kpc. The right column presents massive ellipticals that do have extended multiphase gas. Within each column, the galaxies are arranged in order of descending stellar velocity dispersion. All of the galaxies with \( s_v > -260 \) km s\(^{-1}\) have entropy profiles that rise from \( \sim 1 \) to \( \sim 5 \) kpc more steeply than the precipitation limit, with slopes in general agreement with Equation (6), implying that stellar heating is driving outflows from those galaxies. All of the galaxies with \( s_v < -240 \) km s\(^{-1}\) have entropy profiles beyond \( \sim 1 \) kpc that track the precipitation limit at \( t_{cool}/t_H \approx 10 \) and remain below \( K_{eq} \), implying that radiative cooling exceeds stellar heating.
The observed entropy flattening is in the vicinity of those lines. The fifth single-phase elliptical, NGC 4261, exhibits no significant central entropy flattening and is consistent with a pure cooling flow at ~0.2–1 kpc.

Accretion of hot gas onto the central black hole at the standard Bondi rate (Bondi 1952) is able to supply ~10^{42} erg s^{-1} of kinetic feedback. The Bondi accretion rate onto a black hole of mass MBH is \( \dot{M}_B \approx 2\pi G^2 M_{BH} (\mu n_p)^{5/2} (5K_0/3)^{-3/2} \) in an atmosphere of constant entropy \( K_0 \) (V15). Observations indicating \( K_0 \approx 2 \) keV cm^{-2} at \( \lesssim 0.5 \) kpc therefore imply

\[
\dot{M}_B \approx 0.005 M_\odot \text{yr}^{-1} \left( \frac{M_{BH}}{10^9 M_\odot} \right)^2 K_0^{-3/2},
\]

where \( K_0 \equiv K_0/(2 \text{ keV cm}^2) \). At this rate, conversion of \( \sim 1\% \) of the accreting rest-mass energy into kinetic feedback energy is sufficient to produce \( \sim 10^{42} \) erg s^{-1}. However, considerably more time-averaged power is required to lift the CGM.

Occasional episodes of chaotic cold accretion can provide the power needed to lift the CGM when the ambient central pressure becomes large enough to lower the central cooling time to \( t_{cool} \approx 10 t_{ff} \). Those episodes tend to boost accretion by a factor as great as \( \sim 10^{2} \) over the ambient Bondi accretion rate (Gaspari et al. 2013), raising the kinetic feedback power to \( \sim 10^{44} \) erg s^{-1}, assuming an \( \sim 1\% \) conversion efficiency. They are ultimately limited by the maximum cooling-flow rate within \( \sim 1 \) kpc, which is \( \sim 0.5 M_\odot \text{yr}^{-1} \) in all of these galaxies. It appears that NGC 4261 is a galaxy that is currently experiencing chaotic cold accretion because it has \( t_{cool/ff} \lesssim 10 \) at \( \lesssim 0.5 \) kpc, a multiphase medium at \( \lesssim 100 \) pc that includes a cold gaseous disk, and is producing \( \sim 10^{44} \) erg s^{-1} of kinetic feedback power.

### 4.2. Comments on Individual Galaxies

#### 4.2.1. NGC 1399

The upper left panel of Figure 6 plots observations of NGC 1399, the central galaxy of the Fornax Cluster, on an entropy profile graph similar to the ones in Figure 5. Red symbols show an entropy profile derived from Chandra entropy observations by Werner et al. (2014), and a red dashed line shows an entropy profile derived from the density and temperature fits of Paolillo et al. (2002) to ROSAT data. The steady outflow model shown is specified by \( \sigma_i = 332 \) km s^{-1} and \( r_0 = 2 \) kpc and is a good description of the data from \( \sim 0.5 \) to at least 8 kpc. At larger radii, the observed \( K(r) \) profile becomes shallower than the steady outflow solution, suggesting that CGM confinement of the outflow begins near \( \sim 10 \) kpc. We have placed the line marking \( r_b \) at 6 kpc, where there is an inflection in the observed entropy profile. Inside of 0.5 kpc, the entropy profile flattens at \( \sim 2 \) keV cm^{-2}, consistent with AGN shock heating by kinetic outflows ranging up to a few times \( 10^{42} \) erg s^{-1}.

#### 4.2.2. NGC 4649

In the panel for NGC 4649, the central galaxy of a Virgo Cluster subgroup, red symbols show an entropy profile derived by Werner et al. (2014) from Chandra observations. A red dashed line from 1 to 25 kpc shows an entropy profile derived from the densities and temperatures observed by Randall et al. (2004), and a dotted–dashed red line from 0.1 to 2 kpc shows the entropy profile fit of Humphrey et al. (2006). The steady outflow model shown is specified by \( \sigma_i = 331 \) km s^{-1} and \( r_0 = 2 \) kpc. It is a good description of the data from \( \sim 0.5 \) to \( \sim 20 \) kpc, but the Werner et al. observations suggest an inflection in the entropy profile near 6 kpc, where we have placed the \( r_b \) marker. Inside of 0.5 kpc, the observed \( K(r) \) profile becomes shallower than the outflow solution, flattening slightly below \( 2 \) keV cm^{-2}, consistent with intermittent shock heating by AGN kinetic outflows with \( \sim 10^{42} \) erg s^{-1}.

#### 4.2.3. NGC 4261

The case of NGC 4261, a central group galaxy, is perhaps the most revealing. Red symbols represent Chandra observations by Werner et al. (2014), and a dashed red line shows a fit to...
observations by Humphrey et al. (2009). From 0.1 to 10 kpc, the data points are consistent with a steady flow solution determined by $\sigma_r = 297 \text{ km s}^{-1}$ and $r_0 = 2.5 \text{ kpc}$. Within this stagnation radius, the inner cooling-flow rate reaches $\approx 0.5 M_\odot \text{ yr}^{-1}$. As the flow moves inward, the ambient $t_{\text{cool}}/t_{\text{ff}}$ ratio declines, ultimately dropping below $t_{\text{cool}}/t_{\text{ff}} \approx 10$ at $r \lesssim 0.5 \text{ kpc}$, indicating that the inflow becomes increasingly prone to multiphase condensation (V15). This multiphase inflow presumably supplies gas to the central dusty disk, which has a radius of $\sim 100 \text{ pc}$ (Jaffe et al. 1996). Accretion of the inflowing gas onto the central black hole is therefore capable of providing the current AGN power, assuming a conversion efficiency $\gtrsim 1\%$ from rest-mass energy to kinetic feedback power.

Comparing stellar heating with radiative cooling in NGC 4261 shows that heating dominates outside of 1 kpc and should therefore drive an outflow at larger radii. The data are consistent with an outflow driven by stellar heating having the entropy profile slope predicted by Equation (6) at radii of $\sim 2\text{--}10 \text{ kpc}$. The energy required to lift the CGM in the potential well of NGC 4261 is $\simeq f_\text{fl}M_{200}\sigma_r^2 \approx 10^{60.7} \text{ erg}$. This amount of energy can plausibly be supplied by the AGN if its kinetic power output has been close to the current $10^{44} \text{ erg s}^{-1}$ for several Gyr of its history.

### 4.2.4. NGC 4472

Observations of NGC 4472, which dominates its own subgroup of the Virgo Cluster, are consistent with a steady flow solution from $\sim 0.5$ to beyond 20 kpc. Red symbols show data from Werner et al., the dashed red line shows a fit from Humphrey et al. (2009), and the teal line represents a steady flow model specified by $\sigma_r = 282 \text{ km s}^{-1}$ and $r_0 = 2.5 \text{ kpc}$. Within $\sim 0.5 \text{ kpc}$, the observed entropy profile departs from the steady flow model, flattening near 2 keV cm$^2$, in the vicinity of the horizontal purple line indicating the entropy jumps of intermittent shocks driven by $\sim 10^{42} \text{ erg s}^{-1}$ of AGN power. Stellar heating exceeds radiative cooling at all radii. Here V15 found $\text{min}(t_{\text{cool}}/t_{\text{ff}}) \approx 20$, indicating that the atmosphere in its current configuration does not condense into cold clouds that fuel the central black hole. Given the observed core entropy level and an observed black hole mass of $\approx 2.5 \times 10^9 M_\odot$ (Kormendy & Ho 2013), the resulting Bondi accretion rate of hot gas onto the black hole is $\sim 0.05 M_\odot \text{ yr}^{-1}$, which yields $\sim 10^{43} \text{ erg s}^{-1}$ of kinetic power and suggests a conversion efficiency of $\sim 1\%$.

### 4.2.5. NGC 1407

Red symbols in the panel for NGC 1407, a central group galaxy, show the Werner et al. observations, and the dashed red line is based on the fits made by Su et al. (2014) to their observations. From $\sim 1$ to 8 kpc, the data generally agree with the steady flow solution having $\sigma_r = 266 \text{ km s}^{-1}$ and $r_0 = 2 \text{ kpc}$, as well as the simple power law predicted by Equation (6) and shown by the thick green line. From 10 to 20 kpc, where the temperature profile peaks, there is a bump in the entropy profile. At larger radii, the entropy profile is more consistent with a cosmological slope $K \propto r^{1.1}$ than with the steady outflow solution, so we have placed the $r_0$ marker at 10 kpc. Inside of $\sim 1 \text{ kpc}$, flattening of the entropy profile suggests intermittent shock heating, although the inner region is not as well resolved as in some of the other galaxies.

### 4.2.6. NGC 6868

Blue symbols in the panel for NGC 6868, one of two large galaxies in the Telescopium group, represent the three data points from Werner et al. (2014). A teal line shows a steady flow solution with $\sigma_r = 252 \text{ km s}^{-1}$ and $r_0 = 2 \text{ kpc}$. The data are sparse but consistent with the predicted entropy slope ($\kappa = 0.7$) from Equation (6). This slope makes the entropy profile of NGC 6868 nearly parallel to the precipitation limit, and it tracks near $t_{\text{cool}}/t_{\text{ff}} \approx 20$ (V15). Juraňová et al. (2019) recently suggested that rotation makes the gas in this galaxy especially prone to precipitation.

### 4.2.7. NGC 5846

The central group galaxy NGC 5846 has $\sigma_r < 240 \text{ km s}^{-1}$, and the blue symbols representing Werner et al. (2014) entropy observations are inconsistent with steady homogeneous flow. Instead, they track the precipitation limit at $t_{\text{cool}}/t_{\text{ff}} \approx 10$ from $\sim 1$ to $26 \text{ kpc}$. Beyond that point, the dashed blue line shows the entropy profile observed by Paggi et al. (2017), which gradually steepens until it matches the cosmological slope near 20 kpc, where we have placed the $r_0$ marker. At that radius, the AGM pressure is $\sim 3\%$ of the cosmological pressure expected without radiative cooling or feedback. Inside of 1 kpc, the entropy profile appears to flatten at a level consistent with $\sim 10^{42} \text{ erg s}^{-1}$ of intermittent kinetic feedback power but is not well resolved.

### 4.2.8. NGC 5813

The largest galaxy in its subgroup of the NGC 5846 group, NGC 5813 has a velocity dispersion nearly identical to NGC 5846 and a nearly identical entropy profile, with blue symbols representing Werner et al. (2014) data and a blue dashed line representing Randall et al. (2015) data. It tracks the precipitation limit from 1 to 10 kpc and is inconsistent with steady homogeneous flow. Beyond 10 kpc, its entropy profile steepens to a slope similar to the $K \propto r^{1.1}$ slope produced by cosmological structure formation, but its entropy normalization at 20 kpc is greater by a factor of $\sim 9$, corresponding to a CGM pressure normalization of $\sim 3\%$ of the cosmological expectation. Inside of 1 kpc, the entropy profile becomes flatter than $K \propto r^{2/3}$, leveling at $\sim 3 \text{ keV cm}^2$, again consistent with $\sim 10^{42} \text{ erg s}^{-1}$ of intermittent kinetic feedback power.

### 4.2.9. NGC 5044

The entropy profile of NGC 5044, a central group galaxy, is similar in many ways to those of NGC 5846 and NGC 5813 but tracks the precipitation limit to even greater radii, as shown by the blue symbols representing Werner et al. (2014) observations. A dotted-dashed line beyond 20 kpc shows a fit to observations by David et al. (2017), which attains a cosmological ($K \propto r^{1.1}$) slope outside of the $r_0$ marker at 25 kpc. Here the pressure is $\sim 7\%$ of the cosmological expectation. A dashed blue line at smaller radii shows a fit to observations by David et al. (2009). Inside of 1 kpc, the profile levels off near $3 \text{ keV cm}^2$, indicating $\sim 10^{42} \text{ erg s}^{-1}$ of intermittent kinetic feedback power.

### 4.2.10. NGC 4636

The central group galaxy NGC 4636 has an entropy profile that tracks the precipitation limit from 0.5 to 10 kpc. The blue
dashed line represents data from Trinchieri et al. (1994) and is consistent with a cosmological slope outside of 10 kpc. The Werner et al. (2014) observations depicted by blue symbols are in the vicinity of the precipitation limit from 0.5 to 8 kpc but are also consistent with a pure cooling flow between 0.5 and 2 kpc. At smaller radii, the entropy profile flattens relative to the \( K \propto r^{2/3} \) precipitation-limited profile but reaches \( \sim 1 \) keV cm\(^2\) inside of 100 pc, considerably below the level expected from \( \sim 10^{42} \text{ erg s}^{-1} \) of intermittent kinetic feedback power. There are several possible explanations for this low central entropy level: (1) the time-averaged kinetic AGN power has been \( \sim 10^{41} \text{ erg s}^{-1} \) for the last \( \sim 100 \) Myr; (2) the AGN power has been highly collimated, as in NGC 4261, and has penetrated to \( \gg 1 \) kpc without dissipating much power; or (3) the AGN power has been too weak to balance cooling for the last \( \sim 100 \) Myr. In this last case, a cooling catastrophe is imminent, as suggested by the entropy profile between 0.5 and 2 kpc, and will soon trigger a strong feedback episode.

### 5. Implications for Quenching

The black hole feedback valve model presented here potentially solves some important puzzles in galaxy evolution. For example, AGN feedback is energetically necessary to quench star formation in massive galaxies, but quenching itself appears to be most closely related to a galaxy’s central stellar velocity dispersion (Wake et al. 2012; Bluck et al. 2016, 2020; Teimoorinia et al. 2016) or, equivalently, the surface density of stars within the central 1 kpc (Bell 2008; Franx et al. 2008; Fang et al. 2013; van Dokkum et al. 2015; Woo et al. 2015; Whitaker et al. 2017). The AGN feedback must somehow be related to galactic structure, but which way does the arrow of causality point? Does galactic structure determine the rate of black hole growth, or does AGN feedback shape galactic structure? Also, why is AGN feedback more effective at quenching star formation in massive galaxies than in smaller ones, even though far more energy is required to offset radiative cooling, alleviate CGM pressure confinement, and prevent accumulation of ejected stellar gas?

According to the model, AGN feedback is responding to galactic structure. A galaxy’s ability to remain quenched for long time periods depends most critically on the entropy gradient of its ambient gas. A steep entropy slope (\( \alpha_K > 2/3 \)) strongly inhibits widespread star formation for two reasons.

1. The corresponding gas density slope is steeper than \( n_e \propto r^{-1} \), minimizing the local ratio of stellar heating to radiative cooling at small radii (\( \lesssim 1 \) kpc) and allowing local stellar heating to exceed local radiative cooling at larger radii (\( \approx 1-10 \) kpc).
2. The \( t_{\text{cool}}/t_{\text{H}} \) ratio rises with radius. Multiphase condensation therefore happens primarily near the central black hole, where \( t_{\text{cool}}/t_{\text{H}} \) is minimized, and is suppressed by buoyancy effects at larger radii.

A fundamentally important consequence is that the central stellar mass density of a galaxy should be self-limiting, because Equation (6) predicts that growth in \( \alpha_K \) should accompany growth in \( \sigma_r \). When \( \sigma_r \) becomes large enough, only the region within \( \sim 1 \) kpc of the central black hole can persist in a state with \( \min(t_{\text{cool}}/t_{\text{H}}) \approx 10 \), resulting in episodes of chaotic cold accretion that intermittently supercharge AGN feedback while strongly limiting star formation elsewhere.

Figure 6 indicates that present-day elliptical galaxies with \( \sigma_r > 240 \text{ km s}^{-1} \) have attained such a state, which requires AGN feedback to have lowered the confining CGM pressure by a factor of \( \sim 10^2 \) relative to expectations from cosmological structure formation. However, the other galaxies in Figure 6 are also quenched and have \( \sigma_r = 200-237 \text{ km s}^{-1} \), suggesting that the dependence of quenching on \( \sigma_r \) is a continuous transition rather than a step function at the critical value. Also, the model predicts that the critical value of \( \sigma_r \) should be time-dependent, because the ratio of SN Ia heating to stellar mass loss does not remain constant with time. The remainder of this section discusses these issues, along with some other implications of the black hole feedback valve model.

#### 5.1. Maximum Stellar Velocity Dispersion

Equation (6) predicts that cooling and condensation of ambient gas should become increasingly concentrated toward the center of a galaxy as its stellar velocity dispersion \( \sigma_v \) increases. The result was derived for quiescent stellar populations but has more general applications. For example, consider an actively star-forming galaxy in which explosions of massive stars (SNe II) are driving an outflow at a rate \( \eta \) times the star formation rate \( M_* \). If star formation results in \( 10^{51} \) erg of supernova energy per 100 \( M_\odot \) of star formation, then the specific thermal energy of a supernova-heated outflow is no greater than \( E_\nu \approx 3 \eta^{-1} \text{ keV/} \mu \text{m}_p \), and possibly much less if radiative losses are substantial. The critical velocity dispersion at which \( \alpha_K \approx 2/3 \) in the outflow is then no greater than

\[
\sigma_v \approx 300 \eta^{-1/2} \text{ km s}^{-1}.
\] (26)

Steady supernova-driven flows in galaxies that exceed this limit are cooling-dominated at small radii and become increasingly focused on the central black hole as \( \sigma_v \) rises. Once \( \alpha_K \) exceeds 2/3, the black hole feedback valve described in Section 2.6 responds by lowering the confining CGM pressure until \( t_{\text{cool}}/t_{\text{H}} > 10 \) outside of the central kiloparsec.

A galaxy in this state remains quenched indefinitely because buoyancy prevents multiphase condensation of the ambient medium, except near the central black hole or during eruptions of AGN feedback that lift large quantities of low-entropy gas to greater altitudes. This limiting effect of the black hole feedback valve on \( \sigma_v \) should also be present in numerical simulations of galaxy evolution that implement AGN feedback in the form of bipolar jets capable of thermalizing their energy in the CGM after drilling through the ambient medium out to \( \sim 10 \) kpc. High spatial resolution is necessary because the jets need to be much narrower than a kiloparsec at the base in order to pass through the central few kiloparsecs (as in NGC 4261) without completely disrupting the ambient gas there.

A dramatic demonstration of what happens without such a feedback mechanism can be found in Keller et al. (2016). Figure 4 of that paper presents rotation curves for a set of simulated massive galaxies with efficient superbubble feedback but no AGN feedback. Those galaxies fall into two distinct subsets. One has flat rotation curves with \( \max(v_c) < 250 \) km s\(^{-1}\), while the other has rotation curves with a sharp peak at \( \lesssim 1 \) kpc at which 450 km s\(^{-1}\) \( \lesssim v_c \lesssim 700 \) km s\(^{-1}\). Apparently, the galaxies with centrally peaked rotation curves experienced episodes of centrally focused cooling and runaway star formation after \( \max(v_c) \) exceeded a critical value between 250 and 450 km s\(^{-1}\).
that corresponds to \(180 \text{ km s}^{-1} \leq \sigma_v \leq 320 \text{ km s}^{-1}\). If kinetic AGN feedback had been enabled in these simulations, centrally focused cooling would instead have shut down star formation shortly following the onset of the runaway, thereby preventing the central stellar velocity dispersion from greatly exceeding \(300 \text{ km s}^{-1}\).

This limiting effect of AGN feedback on \(\sigma_v\) does indeed show up in cosmological simulations of massive galaxies by Choi et al. (2018). Without AGN feedback, the central concentration of star formation in their galaxies causes the central stellar mass density to grow to a level corresponding to \(\sigma_v \geq 500 \text{ km s}^{-1}\) by \(z \approx 0\). In those same galaxies, implementation of a kinetic AGN feedback mechanism limits the central stellar mass density to an equivalent stellar velocity dispersion in the range \(250 \text{ km s}^{-1} \leq \sigma_v \leq 400 \text{ km s}^{-1}\) (e.g., de la Rosa et al. 2016), consistent with Equation (26). High-resolution X-ray observations of red nuggets at \(z \geq 1\) are currently not feasible, but examples of analogous galaxies in the low-redshift universe now have \(\sigma_v \geq 240 \text{ km s}^{-1}\) and \(t_{\text{cool}} / t_{\text{ff}} > 10\) (Werner et al. 2018; Buote & Barth 2019), also as expected from the model.

5.2. Dependence of Quenching on \(\sigma_v\)

A velocity dispersion \(\sigma_v > 240 \text{ km s}^{-1}\) appears to be a sufficient condition for the quenching of present-day ellipticals through the black hole feedback valve mechanism, but it is not a necessary condition. For example, Figure 6 shows several elliptical galaxies with \(\sigma_v < 240 \text{ km s}^{-1}\) and specific star formation rates of \(\sim10^{-12} \text{ yr}^{-1}\). The presence of extended multiphase gas in those galaxies suggests that AGN feedback cannot completely suppress multiphase condensation but is still sufficiently well coupled with the CGM to strongly suppress star formation. However, the fraction of galaxies that are quenched is observed to decline with decreasing \(\sigma_v\), implying that AGN feedback is less well coupled to the CGM in smaller galaxies.

Analyses of large galaxy samples from the Sloan Digital Sky Survey show that quenching correlates more closely with \(\sigma_v\) than with any other galactic property (Wake et al. 2012; Bluck et al. 2016, 2020; Teimoorinia et al. 2016). When a quenched state is defined to be a specific star formation rate less than 10% of the average among star-forming galaxies of similar mass (e.g., Bluck et al. 2014), the fraction of both central and satellite galaxies that qualify as quenched is \(\geq90\%\) for \(\sigma_v > 240 \text{ km s}^{-1}\), with no apparent dependence on \(\sigma_v\). Among central galaxies with \(\sigma_v < 240 \text{ km s}^{-1}\), the quenched fraction continuously declines to \(\sim25\%\) at \(\sigma_v = 100 \text{ km s}^{-1}\) (Bluck et al. 2016). Among satellite galaxies, the quenched fraction also declines below \(\sigma_v = 240 \text{ km s}^{-1}\), but not as steeply, down to \(\sim50\%\) at \(\sigma_v = 100 \text{ km s}^{-1}\).

The black hole feedback valve model outlined in this paper is not sophisticated enough to make quantitative predictions for this dependence of the quenched fraction on \(\sigma_v\) or to model environmental effects. It will therefore need to be tested with numerical simulations employing feedback algorithms that produce galaxies similar to those shown in Figure 6. Wang et al. (2019) already made progress by simulating galaxies resembling NGC 4472 and NGC 5044. In their simulations, AGN feedback keeps the galaxy resembling NGC 4472 quenched for several Gyr without producing extended multiphase gas, while the galaxy resembling NGC 5044 develops a persistent multiphase medium but still remains quenched. It will be intriguing to see how the same algorithms play out in simulated galaxies with \(\sigma_v < 200 \text{ km s}^{-1}\).

Generically, we expect this AGN feedback mechanism to be less effective at quenching smaller galaxies because the CGM entropy gradient in those galaxies is less able to inhibit multiphase condensation during feedback bursts, for the reasons shown in Figure 7. Solid magenta lines in that figure indicate the precipitation limit at \(t_{\text{cool}} / t_{\text{ff}} \approx 10\), along which ambient gas is marginally susceptible to multiphase condensation. The X-ray observations show that the central entropy profiles of early-type galaxies rarely, if ever, fall below that limit. Data of sufficient quality to derive resolved entropy profiles like those in Figure 6 are available for only a few galaxies in this range of \(\sigma_v\) (e.g., Babyk et al. 2018). However, larger samples show that X-ray luminosity from within the effective radius of an early-type galaxy does not exceed the limit imposed by the condition \(\min(t_{\text{cool}} / t_{\text{ff}}) \geq 10\) and is more often consistent with \(\min(t_{\text{cool}} / t_{\text{ff}}) \approx 20-30\) (Goulding et al. 2016; Voit et al. 2018).

Given that condition, Figure 7 shows that bursts of AGN feedback tend to raise the central entropy level above the cosmological CGM entropy level in potential wells with \(\sigma_v \lesssim 150 \text{ km s}^{-1}\). Horizontal purple lines in the figure indicate the entropy jump produced by \(10^{42} \text{ erg s}^{-1}\) of kinetic feedback, as given by Equation (20). Horizontal tan lines indicate the CGM entropy scale produced by cosmological structure formation.

Those latter lines mark

\[
K_{200} \equiv kT_0 \left( \frac{200 \text{ km s}^{-1}}{\mu_e m_p} \right)^{-2/3}
\]

(27)

and represent the typical CGM entropy level resulting from accretion shocks (Voit et al. 2003, Voit 2005). If the tan line is below the purple line, then feedback near the center of the potential well produces bubbles of high-entropy gas that buoyantly rise through lower-entropy CGM gas. This configuration is unstable and promotes multiphase condensation of the lower-entropy gas (e.g., McNamara et al. 2016; Voit et al. 2017). Furthermore, even if the tan line is slightly above the purple line, the corresponding entropy gradient remains shallow, meaning that small CGM disturbances are able to promote multiphase condensation. In Figure 7, the tan line falls below the purple line at \(\sigma_v \leq 140 \text{ km s}^{-1}\), for which the generic galaxy model gives \(M_1 \leq 10^{10.8}\) and \(M_{200} \leq 10^{12.2} M_\odot\). This is indeed the boundary below which most of the universe’s star formation now occurs, but the role of entropy gradients in establishing it needs to be explored more quantitatively with high-resolution simulations of cosmological galaxy evolution.

Bower et al. (2017) proposed a similar explanation for how AGN feedback can cause star formation quenching in halos with \(\gtrsim10^{12} M_\odot\). However, they emphasized the role of halo mass instead of \(\sigma_v\). Both proposed explanations recognize that
hot bubbles produced at small radii by feedback will buoyantly rise to large radii in halos with \( \gtrsim 10^{12} M_\odot \). That happens because the entropy produced by accretion shocks then tends to be smaller than the entropy produced closer to the galaxy by feedback heating. Bower et al. (2017) argued that quenching happens in halos with \( \gtrsim 10^{12} M_\odot \) because the higher cosmological CGM entropy prevents supernova-heated bubbles from rising. The resulting buildup of galactic gas then triggers AGN feedback, causing an eruption of energy that heats the CGM and quenches star formation.

Motivated by the fact that quenching is observed to correlate much more directly with \( \sigma_v \) than with halo mass, we argue here that the role of buoyancy is more subtle. In the black hole feedback valve model, both supernova and AGN feedback can inflate high-entropy bubbles that add heat to the CGM and regulate star formation in halos with \( \lesssim 10^{12} M_\odot \). However,
those bubbles fail to quench star formation in lower-mass galaxies because they promote multiphase condensation by producing large-scale entropy inversions.

In order to be effective at long-term quenching of star formation, AGN feedback must maintain both $t_{\text{cool}}/t_{\text{ff}} > 10$ and a positive entropy gradient sufficient for buoyancy to suppress multiphase condensation. We have shown that AGN feedback is able to do so when the depth of the galactic potential well is comparable to the specific energy of supernova heating. As the entropy gradient of a supernova-heated outflow rises with increasing $\sigma_v$ (see the green and teal lines in Figure 7), the ability of buoyancy to suppress condensation increases. And the entropy slope becomes great enough for quenching to be inevitable when $\sigma_v \gtrsim 240 \text{ km s}^{-1}$. Halo mass also plays a role because it determines the CGM entropy level produced by accretion shocks, which enhances buoyancy if the cosmological entropy is great enough, but that role is secondary.

5.3. Redshift Evolution of the Critical $\sigma_v$

One observationally testable prediction of the black hole feedback valve model is that the $\sigma_v$ scale for quenching should be greater earlier in time because the specific energy of stellar ejecta ($\epsilon_{\text{ej}}$) from a younger stellar population is greater. Section 5.1 showed that the critical velocity dispersion is likely to be greater than $240 \text{ km s}^{-1}$ in an actively star-forming population, but even in a quiescent stellar population, $\epsilon_{\text{ej}}$ should be greater earlier in time. That is because the specific SN Ia rate is $\propto t^{-1.3}$ in massive ellipticals (e.g., Friedmann & Maoz 2018), while the specific stellar mass-loss rate at late times is $\propto t^{-1}$ (Leitner & Kravtsov 2011), resulting in $\epsilon_{\text{ej}} \propto t^{-0.3}$.

Predictions for how the critical $\sigma_v$ for quenching should depend on redshift need to account for the fact that the zero-point for $t$ should coincide with the end of rapid star formation, which may not be the same for all galaxies in a given sample. This paper will therefore not attempt a detailed analysis. Instead, we present an illustrative example for a population of galaxies that formed most of their stars by $z \approx 3$ Gyr after the big bang ($z \approx 2$). One billion yr later ($z \approx 1.5$), most of the heat generated by the aging stellar population was from SNe Ia. The value of $\epsilon_{\text{ej}}$ then declined by a factor of $\approx 2$ (because $\epsilon_{\text{ej}} \propto t^{-0.3}$) during the ensuing 9 Gyr, implying that the $\sigma_v$ scale at which $\epsilon_{\text{ej}} \approx 2/3$ was a factor of $\approx 1.4$ greater at $z \approx 1.5$ than at $z \approx 0$.

Optical observations suggest that the actual quenching scale has declined by a similar factor, but quantitative comparisons await greater consistency among the definitions of quenching as a function of redshift. For example, Franx et al. (2008) defined the stellar surface density threshold for quenching to be the level at which the specific star formation rate drops more than a factor of 3 below the prevailing rate at low stellar surface density. They found that this surface density threshold was a factor of $\sim 2.5$ smaller at $z \approx 0$ than at $z \approx 1.5$, corresponding to a factor of $\sim 1.6$ in $\sigma_v$. A similar study by van Dokkum et al. (2015) focused more narrowly on the properties of massive compact galaxies at $1.5 < z < 2.25$, finding that their properties are consistent with a quenching probability that rises from zero at $\sigma_v = 220 \text{ km s}^{-1}$ to unity at $\sigma_v = 320 \text{ km s}^{-1}$. When translated to low redshift through division by the predicted factor of 1.4, this range shifts to $160 \text{ km s}^{-1} \lesssim \sigma_v \lesssim 230 \text{ km s}^{-1}$, consistent with the low-redshift model outlined in this paper. Whitaker et al. (2017) focused instead on the interval $0.5 < z < 2.5$ and examined multiple definitions of quenching. Their work aligns with van Dokkum et al. (2015) at $z = 2$ and is consistent with the central density threshold decline seen by Franx et al. (2008) down to $z = 0.5$.

5.4. Quenching in Galaxy Cluster Cores

We do not expect star formation quenching in the central galaxies of galaxy clusters to depend as directly on $\sigma_v$ because their potential wells are much deeper than the stellar velocity dispersion of the central galaxy would indicate. In the generic galaxy model of Section 3.1, we assumed that the maximum circular velocity of the halo around a massive galaxy was similar to that of the galaxy itself. That assumption applies to galaxy groups with a velocity dispersion of $\sim 300 \text{ km s}^{-1}$ and an X-ray temperature of $\sim 1 \text{ keV}$ but not to galaxy groups and clusters with a velocity dispersion $\gtrsim 450 \text{ km s}^{-1}$ and an X-ray temperature $\gtrsim 2 \text{ keV}$.

Because of the deeper halo potential well, the gas pressure in the central galaxy of a galaxy cluster is often considerably greater than in the galaxies this paper has analyzed, so radiative cooling can greatly exceed SN Ia heating. Cool-core clusters, with a central entropy level $< 30 \text{ keV cm}^2$ and $t_{\text{cool}} \lesssim 1 \text{ Gyr}$ at 10 kpc, have the largest central gas pressure. Observations show that extended multiphase gas is nearly always present near the center of a cool-core cluster, independent of $\sigma_v$. The AGN feedback limits the cooling flows in those central galaxies but allows multiphase gas to collect and star formation to proceed at $\sim 10^{-5}$ of the uncompensated cooling-flow rate (McDonald et al. 2018). In that regard, central galaxies in cool-core clusters are similar to the galaxies in Figure 6 with 200 km s$^{-1} < \sigma_v < 240 \text{ km s}^{-1}$.

Some of those central cluster galaxies have $\sigma_v > 240 \text{ km s}^{-1}$, meaning that there is a halo mass above which the mechanism outlined in Section 2.6 does not work as described. In order for the mechanism to operate, the central AGN must be capable of lowering the CGM pressure by pushing much of the ambient halo gas to greater altitudes. For that to happen, the total amount of kinetic feedback energy must be at least as great as the CGM binding energy, which is several times $10^{45} \text{ erg}$ in a $10^{14} \text{ M}_\odot$ halo. Significant lifting of that CGM requires $\sim 10^{45} \text{ erg s}^{-1}$ of energy input, sustained over a cosmological timescale. The analogous requirement for a $10^{15} \text{ M}_\odot$ halo is several times $10^{46} \text{ erg s}^{-1}$. Observations of feedback in galaxy clusters show that the AGN’s kinetic power is rarely greater than a few times $10^{45} \text{ erg s}^{-1}$ (e.g., McNamara & Nulsen 2012), accounting for why essentially all cool-core clusters with AGN feedback also have extended multiphase gas, regardless of $\sigma_v$ in the central galaxy. However, why kinetic AGN power in the centers of galaxy clusters is not greater remains an open question.

5.5. Quenching and the $M_{\text{BH}}-\sigma_v$ Relation

The black hole feedback valve mechanism for quenching implies that $M_{\text{BH}}$ should depend on $\sigma_v$ because it requires the black hole’s integrated kinetic energy output to be at least comparable to the binding energy of the CGM. The resulting lower limit can be expressed as

$$M_{\text{BH}} \gtrsim \frac{f_{\text{f}} M_{\text{200}} \sigma_v^2}{\epsilon_{\text{BH}} c^2} \approx 10^8 \text{ } M_\odot \left(\frac{\epsilon_{\text{BH}}}{10^{-8}}\right)^{-1} \sigma_v^{-5},$$

(28)

Voit et al.
where $\epsilon_{\text{BH}}$ is the proportion of the black hole’s rest-mass energy that becomes thermalized in the CGM. A similar limit follows from setting $\epsilon_{\text{BH}} M_{\text{BH}} c^2 \gtrsim L_x H_{\text{vir}}^{-1}$ (Voit et al. 2015a). These limits are generic to any quenching model that relies primarily on AGN feedback to reduce the CGM pressure around a central galaxy (e.g., Davies et al. 2019, 2020; Oppenheimer et al. 2020), but the efficiency factor $\epsilon_{\text{BH}}$ is difficult to infer from simple models. Observations showing that $M_{\text{BH}} \approx 7 \times 10^8 M_\odot$ (e.g., Kormendy & Ho 2013) suggest that the scaling of $M_{\text{BH}}$ with $\sigma_v$ may result from the CGM lifting requirement of AGN feedback. And the observations still imply $\epsilon_{\text{BH}} \gtrsim 1.4 \times 10^{-3} \sigma_v^{-0.6}$ even if the connection is not directly causal.

In cosmological simulations, the conversion efficiency $\epsilon_{\text{kin}}$ of accreted rest-mass energy to kinetic AGN feedback output is generally a free parameter. Oppenheimer et al. (2020) recently analyzed an EAGLE simulation in which that conversion efficiency was set to $\epsilon_{\text{kin}} = 1.67 \times 10^{-2}$. They found that star formation quenching was strongly correlated with both $M_{\text{BH}}$ and lifting of the CGM. They also found that CGM lifting is important role of angular momentum. In general, rotation is expected to promote development of a multiphase medium by suppressing the stabilizing effects of buoyancy on condensation (Gaspari et al. 2015; Sobacchi & Sormani 2019; Sormani & Sobacchi 2019). A rotating CGM is therefore more likely to condense at greater levels of $\min(t_{\text{cool}}/t_{\text{ff}})$, making quenching less likely at a given $\sigma_v$. We plan to explore the effects of angular momentum in future work. This limitation currently precludes us from applying the model directly to late-type galaxies, in which angular momentum will be more important.

5.7. Angular Momentum

So far, our model completely ignores the undoubtedly important role of angular momentum. In general, rotation is expected to promote development of a multiphase medium by suppressing the stabilizing effects of buoyancy on condensation. Galactic outflows driven by quiescent stellar populations. We demonstrate that an outflow’s profiles of pressure, density, and gas entropy depend directly on $\epsilon_{\text{ADR}}/v_{\text{vir}}^2$, the ratio of the specific energy of ejected stellar gas to the square of the galaxy’s circular velocity. Galaxies with $\sigma_v \gtrsim 240$ km s$^{-1}$ can remain in a steady state consisting of an inner cooling flow surrounded by a supernova-heated outflow, while galaxies with $\sigma_v \lesssim 240$ km s$^{-1}$ are prone to multiphase circulation. The AGN feedback in the subset with $\sigma_v \gtrsim 240$ km s$^{-1}$ is therefore able to tune itself to lift the CGM, reduce the cooling pressure it exerts, and quench star formation through a mechanism we have called the black hole feedback valve. In galaxies with $\sigma_v \lesssim 240$ km s$^{-1}$, this quenching mechanism becomes less reliable because feedback is more likely to produce multiphase precipitation away from the galaxy’s center. Figure 8 graphically summarizes the model.

The most notable features of the model are as follows.

1. It predicts that a radial supernova-heated outflow through a massive galaxy should have a power-law entropy profile slope ($\alpha_K$) that depends primarily on $\epsilon_{\text{ADR}}/v_{\text{vir}}^2$ (Section 2.2).

2. The entropy slope $\alpha_K \approx 2/3$ is special because it corresponds to an electron density profile $n_e \propto r^{-1}$ along which the ratio of SN Ia heating to radiative cooling remains approximately constant with radius. It is also special because the $t_{\text{cool}}/t_{\text{ff}}$ ratio remains approximately constant with radius. Galactic outflows that have $\alpha_K > 2/3$ therefore tend to promote cooling and condensation of ambient gas near the origin, in the vicinity of the central black hole, and become less prone to multiphase condensation as $r$ increases. Conversely, outflows with $\alpha_K < 2/3$ are more prone...
Figure 8. Graphical summary of the black hole feedback valve model. At the top is a graph showing relationships between the power-law entropy slope $\alpha_K$ of a galaxy’s ambient gas and the circular velocity $v_c$ of its potential well. The lower horizontal axis gives $\sigma_v = \mu_v / \sqrt{2}$, along with arrows indicating how galactic star formation correlates with $\sigma_v$ (Section 2.3). A salmon line shows the predictions of Equation (6) for supernova-heated outflows driven by a specific stellar heat input, $\epsilon_s = 2 \text{ keV}/\mu m_p$. An orange line shows predictions for $\epsilon_s = 3 \text{ keV}/\mu m_p$. A dashed purple line shows the critical entropy slope $\alpha_K = 2/3$ (Section 2.3). Dotted salmon and orange lines indicate the values of $\sigma_v$ that are critical for each value of $\epsilon_s$. A dotted teal line illustrates how the family of steady flow solutions shown in Figure 5 departs from the predictions of Equation (6) above $\sigma_v \approx 300 \text{ km s}^{-1}$ and saturates near $\alpha_K \approx 1.2$. Symbols correspond to the galaxies shown in Figure 1 and represent the best-fitting values of $\alpha_K$ from 1 to 10 kpc. Within galaxies having $200 \text{ km s}^{-1} < \sigma_v < 240 \text{ km s}^{-1}$, radiative cooling exceeds supernova heating in this radial interval (Section 2.4), so $\alpha_K$ is not expected to follow the predictions of Equation (6). Most have entropy slopes similar to the precipitation limit (magenta line), but one has a slope closer to that of a pure cooling flow (blue line). Below the graph are three schematic illustrations of the model’s qualitative predictions for how the flow pattern in a galaxy’s ambient medium should depend on $\epsilon_s$. On the left is a lower-mass galaxy around which feedback blows bubbles with greater specific entropy than the CGM. Those buoyant bubbles drive multiphase circulation and fail to prevent cold, star-forming gas from collecting in the galaxy’s disk (Section 5.2). In the middle is a more massive galaxy in which CGM entropy exceeds what supernova heating can produce. The CGM pressure therefore confines ejected stellar gas and causes some of it to accrete onto the central black hole. Then, AGN feedback suspend the ambient medium in a marginally precipitating state, driving multiphase circulation but preventing significant star formation. In the high-mass galaxy on the right, supernova heating beyond the stagnation radius $r_\text{stagn}$ can drive an outflow with $\alpha_K > 2/3$ because $\sigma_v > 240 \text{ km s}^{-1}$. Cooling and condensation of ambient gas is therefore focused on the central black hole, which responds by producing strong jets that heat the CGM. In this configuration, CGM pressure at the boundary radius $r_r$ determines the strength of the cooling flow inside of $r_r$ and therefore acts like a knob on a valve that governs AGN feedback power. The valve adjusts itself so that time-integrated AGN feedback power suffices to lift much of the CGM out of the galaxy’s potential well, leading to complete quenching of star formation.

to condensation at large radii and therefore promote multiphase circulation. For a stellar population age of $\approx 10 \text{ Gyr}$, the critical slope $\alpha_K \approx 2/3$ corresponds to $\sigma_v \approx 240 \text{ km s}^{-1}$ (Section 2.3).

3. The X-ray observations of massive elliptical galaxies agree with the heated outflow analysis, in that galaxies with $\sigma_v > 240 \text{ km s}^{-1}$ tend to have SN Ia heating rates that exceed radiative cooling at $r \approx 1$–10 kpc, while radiative cooling tends to exceed SN Ia heating over the same radial range among galaxies with $\sigma_v < 240 \text{ km s}^{-1}$ (Section 2.4).

4. The normalization of the density, pressure, and entropy profiles of a subsonic heated outflow are determined by an outer pressure boundary condition set by the CGM. If the pressure is too large, radiative cooling will exceed SN Ia heating, causing a cooling flow that triggers AGN feedback. Coupling between AGN feedback and CGM pressure should therefore keep the ambient gas near the
AGN close to the point of heating/cooling balance (Section 2.5).

5. In galaxies with $\sigma_v \gtrsim 240$ km s$^{-1}$, this coupling between AGN feedback and the CGM should form a self-regulating valve that links AGN fueling to CGM pressure. Feedback in those galaxies adds heat to the CGM, causing it to expand until the reduction in its pressure brings time-averaged heating into balance with the energy needed to lift the CGM (Section 2.6). This black hole feedback valve mechanism inevitably quenches star formation because the ambient value of $t_{\text{cool}}/t_{\text{ff}}$ rises beyond $\approx 20$ at $r \gtrsim 1$ kpc, making the ambient medium stable to multiphase condensation (Section 2.7).

6. In galaxies with $\sigma_v \lesssim 240$ km s$^{-1}$, outflows heated by SNe Ia and confined by a significant CGM pressure cannot remain heating-dominated as they propagate to large radii. They are destined to become cooling-dominated, producing entropy inversions that are unstable to multiphase condensation. Those galaxies are consequently prone to precipitation and evolve toward a precipitation-limited state with $t_{\text{cool}}/t_{\text{ff}} \approx 10$–20 over a broad range of radii (Section 2.7).

7. Simulations designed to test the black hole feedback valve mechanism require high spatial resolution ($<100$ pc) because the mechanism calls for high-powered jets ($\gtrsim 10^{44}$ erg s$^{-1}$) to drill through the galaxy’s ambient medium at 1–10 kpc without significantly disrupting it before thermalizing their energy in the CGM at 10–100 kpc (Section 2.8).

8. Numerical steady flow solutions corroborate the analytical estimates on which the black hole feedback valve model is based. In particular, steady flows in galaxies with $\sigma_v \gtrsim 240$ km s$^{-1}$ are cooling flows at small radii and SN Ia–heated outflows at larger radii, with $\alpha_K$ in alignment with the predictions of Equation (6). Steady flows in galaxies with $\sigma_v \lesssim 240$ km s$^{-1}$, on the other hand, are cooling-dominated at large radii and prone to developing entropy inversions (Section 3).

Comparisons with high-quality Chandra X-ray observations of 10 massive elliptical galaxies support the model.

1. In the galaxies with $\sigma_v \gtrsim 240$ km s$^{-1}$, the entropy slope $\alpha_K$ from $\sim 1$ to 10 kpc generally agrees with the predictions of Equation (6), and those galaxies are single phase in that radial interval, with one exception (NGC 6868, possibly a borderline case with $\sigma_v = 252$ km s$^{-1}$). The galaxies with $\sigma_v \lesssim 240$ km s$^{-1}$, in contrast, are multiphase in that radial interval and track the precipitation limit at $t_{\text{cool}}/t_{\text{ff}} \approx 10$ (Section 4). Outside of 10 kpc, ambient gas around the multiphase galaxies tends to have lower entropy, greater pressure, and greater density, resulting in greater X-ray luminosities relative to the single-phase galaxies. Larger samples of elliptical galaxies should therefore be checked to see if there is an inflection of the $L_X - \sigma_v$ relation above 240 km s$^{-1}$ once the central galaxies of groups and clusters with $kT \gtrsim 2$ keV have been excluded.

2. Inside of $\sim 0.5$ kpc, the entropy profiles of most of the ellipticals analyzed depart from the power laws observed at larger radii and flatten near $K_0 \approx 2$ keV cm$^2$. That entropy level is consistent with intermittent shock heating of a precipitation-limited atmosphere by $\sim 10^{42}$ erg s$^{-1}$ of kinetic feedback. Bondi accretion of ambient gas at that entropy level onto a central black hole with $M_{\text{BH}} \sim 10^8 M_\odot$ is capable of supplying the currently observed feedback power, but it is insufficient to lift the CGM. However, the one galaxy with an unbroken power-law entropy distribution inside of 0.5 kpc (NGC 4261) has a kinetic power output 2 orders of magnitude greater. It is also the only galaxy in the sample with $t_{\text{cool}}/t_{\text{ff}} \lesssim 10$ at <200 pc, implying that chaotic cold accretion onto the central black hole is temporarily supercharging AGN feedback and enabling it to lift the CGM (Section 4.1.1; see also V15).

The model was inspired by X-ray observations but has broad implications for optical/IR studies of galaxy evolution (Section 5).

1. It links the ability of AGN feedback to quench star formation directly to $\sigma_v$, which optical/IR observations have shown to be the galaxy attribute most closely correlated with quenching of central galaxies (Wake et al. 2012; Bluck et al. 2016, 2020; Teimoorinia et al. 2016). At $\sigma_v \gtrsim 240$ km s$^{-1}$, observations show that $\gtrsim 90\%$ of both central and satellite galaxies are quenched, with no apparent dependence on $\sigma_v$. This finding aligns well with the model’s prediction that supernova heating outflows should be homogeneous in galaxies with $\sigma_v \gtrsim 240$ km s$^{-1}$. In contrast, the fraction of galaxies with $\sigma_v \lesssim 240$ km s$^{-1}$ that are quenched appears to be a continuous function of $\sigma_v$ that drops below 25% for central galaxies with $\sigma_v \lesssim 100$ km s$^{-1}$. According to the model, the quenched fraction should be smaller at smaller $\sigma_v$ because energetic central feedback more easily produces entropy inversions that result in multiphase condensation (Section 5.2). However, this qualitative conclusion needs to be investigated more quantitatively with numerical simulations.

2. Earlier in time, the predicted critical value of $\sigma_v$ for quenching is greater, because $\epsilon_e$ is greater in younger stellar populations. While star formation is still active, the critical value for effective quenching by AGN feedback is $\sigma_v \approx 300$ km s$^{-1}$, where $\eta$ is the ratio of gas outflow rate to star formation rate. It implies that galaxies cannot greatly exceed $\sigma_v \approx 300$ km s$^{-1}$ without supercharging AGN feedback. The properties of “red nugget” galaxies at $z \sim 2$ support this implication because they are compact, quenched galaxies with 250 km s$^{-1} \lesssim \sigma_v \lesssim 400$ km s$^{-1}$ (Section 5.1).

3. As a galaxy’s stellar population ages, the critical value of $\sigma_v$ declines because the specific SN Ia rate drops more rapidly than the specific stellar mass-loss rate. This decline plausibly accounts for observations showing that the central stellar surface density associated with quenching declines with time (Section 5.3).

4. In the central galaxies of massive galaxy clusters, $\alpha_K$ is not expected to correlate as closely with $\sigma_v$ because the greater central CGM pressure in cool-core clusters causes radiative cooling to greatly exceed SN Ia heating. Those...
systems are instead observed to have roughly constant $\tau_{\text{cool}}/\tau_{\text{ff}}$ profiles that track the precipitation limit, with $\sigma_2 \approx 2/3$ at 5–20 kpc (Section 5.4).

5. Any model for quenching that requires AGN feedback to alleviate CGM pressure confinement, including the one presented in this paper, predicts a lower limit on $M_{\text{BH}}$ that depends on $\sigma_2$. The resulting scaling ($M_{\text{BH}} \propto \sigma_2^5$) is interestingly close to the observed scaling. If central black hole mass is indeed linked to the binding energy of the CGM, then observations imply that AGN feedback thermalizes a fraction $\geq 10^{-3}$ of the black hole’s rest-mass energy in the surrounding CGM (Section 5.5).

6. Galaxies in which rotation helps to support the ambient medium might not conform as precisely to the model’s predictions because of how angular momentum alters the buoyancy effects that limit condensation. Werner et al. (2014) suggested that NGC 6868 may be one such example (see also Juraño et al. 2019). Future modeling will therefore need to account for how rotation affects the critical value of $\sigma_2$ (Section 5.7).

G.M.V. acknowledges helpful conversations with Iu. Babyk, A. Bluck, L. Ciotti, A. Evrard, D. Maoz, B. McNamara, B. Oppenheimer, and B. Terrazas and support from Chandra Science Center grant TM8-19006X. G.B. acknowledges support from the NSF (grant AST-1615955, OAC-1835509) and NASA (grant NNX15AB20G) and computational support from NSF XSEDE. B.W.O. acknowledges support from the NSF (AST-1517908), NASA ATP (15AP396), and 80NSSC18K1105.

**ORCID iDs**

G. Mark Voit https://orcid.org/0000-0002-3514-0383
Greg L. Bryan https://orcid.org/0000-0003-2630-9228
Deovrat Prasad https://orcid.org/0000-0003-1255-6375
Rachel Frisbie https://orcid.org/0000-0002-3662-8591
Megan Donahue https://orcid.org/0000-0002-2808-0853
Brian W. O’Shea https://orcid.org/0000-0001-5880-0703
Ming Sun https://orcid.org/0000-0001-5880-0703

**References**

Babyk, I. V., McNamara, B. R., Nulsen, P. E. J., et al. 2018, *ApJ*, 862, 39
Balbus, S. A., & Soker, N. 1989, *ApJ*, 341, 611
Beckmann, R. S., Dubois, Y., Guillard, P., et al. 2019, *A&A*, 631, A60
Bell, E. F. 2008, *ApJ*, 682, 355
Binney, J., Nipoti, C., & Fraternali, F. 2009, *MNRAS*, 397, 1804
Binney, J., & Tabor, G. 1995, *MNRAS*, 276, 663
Birzan, L., Rafferty, D. A., McNamara, B. R., Wise, M. W., & Nulsen, P. E. J. 2004, *ApJ*, 607, 800
Black, A. F. L., Maiolino, R., Sánchez, S. F., et al. 2020, *MNRAS*, 492, 96
Black, A. F. L., Mendel, J. T., Ellison, S. L., et al. 2014, *MNRAS*, 441, 599
Black, A. F. L., Mendel, J. T., Ellison, S. L., et al. 2016, *MNRAS*, 462, 2559
Bondi, H. 1952, *MNRAS*, 112, 195
Bower, R. G., Schaye, J., Frenk, C. S., et al. 2017, *MNRAS*, 465, 32
Buote, D. A., & Barth, A. J. 2019, *ApJ*, 887, 91
Choi, E., Ostriker, J. P., Naab, T., Oser, L., & Moster, B. P. 2015, *MNRAS*, 449, 4105
Choi, E., Somerville, R. S., Ostriker, J. P., Naab, T., & Hirschmann, M. 2018, *ApJ*, 866, 91
Choudhury, P. P., & Sharma, P. 2016, *MNRAS*, 457, 2554
Choudhury, P. P., Sharma, P., & Quataert, E. 2019, *MNRAS*, 488, 3195
Churazov, E., Brüggen, M., Kaiser, C. R., Böhringer, H., & Forman, W. 2001, *ApJ*, 554, 261
Ciotti, L., & Ostriker, J. P. 2001, *ApJ*, 551, 131

Voit et al.
Pizzolato, F., & Soker, N. 2005, ApJ, 632, 821
Prasad, D., Sharma, P., & Babul, A. 2015, ApJ, 811, 108
Prasad, D., Sharma, P., & Babul, A. 2017, MNRAS, 471, 1531
Randall, S. W., Sarazin, C. L., & Irwin, J. A. 2004, ApJ, 600, 729
Randall, S. W., Nulsen, P. E. J., Jones, C., et al. 2015, ApJ, 805, 112
Revaz, Y., Combes, F., & Salomé, P. 2008, A&A, 477, L33
Sharma, M., & Nath, B. B. 2013, ApJ, 763, 17
Sharma, P., McCourt, M., Parrish, I. J., & Quataert, E. 2012a, MNRAS, 427, 1219
Sharma, P., McCourt, M., Quataert, E., & Parrish, I. J. 2012b, MNRAS, 420, 3174
Shen, K. J., Kasen, D., Miles, B. J., & Townsley, D. M. 2018, ApJ, 854, 52
Shin, M.-S., Ostriker, J. P., & Ciotti, L. 2012, ApJ, 745, 13
Sobacchi, E., & Sormani, M. C. 2019, MNRAS, 486, 205
Soker, N. 2016, NewAR, 75, 1
Sormani, M. C., & Sobacchi, E. 2019, MNRAS, 486, 215
Stern, J., Fielding, D., Faucher-Giguère, C.-A., & Quataert, E. 2019, MNRAS, 488, 2549
Su, Y., Gu, L., White, R. E., III, & Irwin, J. 2014, ApJ, 786, 152
Tabor, G., & Binney, J. 1993, MNRAS, 263, 323
Teimoorinia, H., Black, A. F. L., & Ellison, S. L. 2016, MNRAS, 457, 2086
Terrazas, B. A., Bell, E. F., Henriques, B. M. B., et al. 2016, ApJL, 830, L12
Terrazas, B. A., Bell, E. F., Woo, J., & Henriques, B. M. B. 2017, ApJ, 844, 170
Thompson, T. A., Quataert, E., Zhang, D., & Weinberg, D. H. 2016, MNRAS, 455, 1830
Trinchieri, G., Kim, D. W., Fabiano, G., & Canizares, C. R. C. 1994, ApJ, 428, 555
van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, ApJ, 813, 23
Voit, G. M. 2018, ApJ, 868, 102
Voit, G. M. 2019, ApJ, 880, 139
Voit, G. M., Bryan, G. L., O’Shea, B. W., & Donahue, M. 2015a, ApJL, 808, L30
Voit, G. M., & Donahue, M. 2005, ApJ, 634, 955
Voit, G. M., & Donahue, M. 2011, ApJL, 738, L24
Voit, G. M., Donahue, M., Bryan, G. L., & McDonald, M. 2015b, Nat, 519, 203
Voit, G. M., Donahue, M., O’Shea, B. W., et al. 2015c, ApJL, 803, L21,(V15)
Voit, G. M., Donahue, M., Zahedy, F., et al. 2019, ApJL, 879, L1
Voit, G. M., Ma, C. P., Greene, J., et al. 2018, ApJ, 853, 78
Voit, G. M., Meece, G., Li, Y., et al. 2017, ApJ, 845, 80
Wake, D. A., van Dokkum, P. G., & Franx, M. 2012, ApJL, 751, L44
Wang, C., Li, Y., & Ruszkowski, M. 2019, MNRAS, 482, 3576
Werner, N., Allen, S. W., & Simionescu, A. 2012, MNRAS, 425, 2731
Werner, N., Lakhchaura, K., Canning, R. E. A., Gaspari, M., & Simionescu, A. 2018, MNRAS, 477, 3868
Werner, N., McNamara, B. R., Churazov, E., & Scannapieco, E. 2019, SSRv, 215, 5
Werner, N., Oonk, J. B. R., Sun, M., et al. 2014, MNRAS, 439, 2291
Whitaker, K. E., Bezanson, R., van Dokkum, P. G., et al. 2017, ApJ, 838, 19
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
Woo, J., Dekel, A., Faber, S. M., & Koo, D. C. 2015, MNRAS, 448, 237
Yang, H. Y. K., & Reynolds, C. S. 2016, ApJ, 829, 90