

MISTRAL: a model for AGN winds from radiatively efficient accretion in cosmological simulations

Marion Farcy¹,¹★ Michaela Hirschmann¹, Rachel S. Somerville², Ena Choi³,
 Sophie Koudmani^{4,5,6}, Thorsten Naab⁷, Rainer Weinberger⁸, Jake S. Bennett⁹,
 Aklant K. Bhowmick¹⁰, Hyunseop Choi^{11,12}, Lars Hernquist⁹, Julie Hlavacek-Larrondo¹¹,
 Bryan A. Terrazas¹³ and Francesco Valentino^{14,15}

¹*Institute of Physics, Laboratory for Galaxy Evolution, EPFL, Observatoire de Sauverny, Chemin Pegasi 51, CH-1290 Versoix, Switzerland*

²*Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA*

³*Department of Physics, University of Seoul, 163 Seoulsiripdaero, Dongdaemun-gu, Seoul 02504, Republic of Korea*

⁴*St Catharine's College, University of Cambridge, Trumpington Street, Cambridge CB2 1RL, UK*

⁵*Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

⁶*Centre for Astrophysics Research, Department of Physics, Astronomy and Mathematics, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK*

⁷*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, D-85741 Garching, Germany*

⁸*Leibniz Institute for Astrophysics Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*

⁹*Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA*

¹⁰*Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA*

¹¹*Département de Physique, Université de Montréal, Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada*

¹²*Mila – Quebec Artificial Intelligence Institute, Montréal, QC H2S 3H1, Canada*

¹³*Department of Physics & Astronomy, Oberlin College, Oberlin, OH 44074, USA*

¹⁴*Cosmic Dawn Center (DAWN), Denmark*

¹⁵*DTU Space, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark*

Accepted 2025 September 1. Received 2025 September 1; in original form 2025 April 10

ABSTRACT

Feedback from active galactic nuclei (AGNs) is crucial for regulating galaxy evolution. Motivated by observations of broad absorption line winds from rapidly accreting supermassive black holes (SMBHs), we introduce the MISTRAL AGN feedback model, implemented in the AREPO code. MISTRAL comes in two versions: continuous radial (MISTRAL-CONTINUOUS) and stochastic bipolar momentum deposition (MISTRAL-STOCHASTIC). Using the framework of the IllustrisTNG simulations, we explore the effect of MISTRAL on BH and galaxy properties, through an idealized Milky Way-mass galaxy and cosmological zoom simulations run down to $z = 2$. Unlike standard thermal AGN feedback prescriptions, MISTRAL generates galaxy-scale winds that mimic outflows driven by BH accretion. MISTRAL-CONTINUOUS produces short-lived galactic fountains, and is inefficient at regulating the growth of massive galaxies at $z = 2$. In contrast, MISTRAL-STOCHASTIC efficiently suppresses star formation in massive galaxies, reproduces the empirical stellar-to-halo mass relation, and yields a consistent trend of BH-stellar mass evolution. By supporting large-scale outflows while simultaneously preventing gas inflows, MISTRAL-STOCHASTIC additionally regulates the cold and hot gas fractions at both galaxy and halo scales. MISTRAL-STOCHASTIC therefore works self-consistently across the halo mass range explored (10^{12} – $3 \times 10^{13} M_{\odot}$), without adopting an SMBH-mass-dependent AGN feedback scheme such as the one used in IllustrisTNG. Our model is a promising tool for predicting the impact of AGN winds on galaxy evolution, and interpreting the growing population of high-redshift galaxies and quasars observed by *James Webb Space Telescope*. This work is part of the ‘Learning the Universe’ collaboration, which aims to infer the physical processes governing the evolution of the Universe.

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

1 INTRODUCTION

For more than two decades, most (if not all) massive galaxies are thought to host a supermassive black hole (SMBH) in their centres, with masses ranging from a million to a few billion solar masses

(Magorrian et al. 1998; Schödel et al. 2002). As material accretes on to SMBHs, vast amounts of energy are released in the form of energetic radiation, winds, and jets (Fabian 2012; Tombesi et al. 2013; Ciccone et al. 2014; Genzel et al. 2014), powering so-called active galactic nuclei (AGNs). Since the energy released by an SMBH can easily exceed the binding energy of its host galaxy (e.g. King & Pounds 2015), AGNs are thought to play a significant role in galaxy

* E-mail: marion.farcy@epfl.ch

evolution (Silk & Rees 1998; Kauffmann & Haehnelt 2000; Bower et al. 2006; Croton et al. 2006; Somerville et al. 2008; Somerville & Davé 2015; Naab & Ostriker 2017). Depending on the efficiency with which the BH accretion energy released couples to the surrounding gas, the back reaction of an AGN – referred to as AGN feedback – can affect both BH growth and the evolution of the host galaxy by heating and ejecting gas, and (or) preventing gaseous inflows, and thus depriving both star formation and BH accretion of their source of fuel (Fabian 2012; Gaspari, Ruszkowski & Sharma 2012; Harrison 2017). This BH–galaxy coevolution scenario is further supported by the tight correlations observed between the BH mass and the properties of their host galaxies, such as stellar bulge luminosity, mass, and stellar velocity dispersion (Tremaine et al. 2002; Kormendy & Ho 2013; Reines & Volonteri 2015, even if also non-causal origins of these relations have been discussed by e.g. Hirschmann et al. 2010 and Jahnke & Macciò 2011).

In the past decade, observations have found increasingly strong and abundant evidence for AGN feedback, in the form of jets and winds. Collimated radio-emitting jets are thought to be powered by the extraction of the SMBH spin, through the Blandford–Znajek process (Blandford & Znajek 1977). Jets are associated with radiatively inefficient AGN events, when SMBHs are accreting at low rates (e.g. McNamara et al. 2000; Fabian et al. 2011; Hlavacek-Larrondo, Li & Churazov 2022). By inflating hot X-ray cavities in the intra-cluster medium of galaxies, jets are commonly invoked as an efficient mechanism for reducing cooling flows on to galaxies, and thus, subsequent star formation, which is also referred to as the radio or jet feedback mode (e.g. Fabian 2012). However, in the radiatively inefficient regime, other forms of feedback can exist, such as outflows, without the presence of strong relativistic jets. Therefore, throughout this paper, we refer to AGN feedback in this low-accretion state as radiatively inefficient feedback.

Conversely, in high-luminosity AGNs (with bolometric luminosities of 10^{46} – 10^{48} erg s $^{-1}$), the accreting mass is efficiently converted into radiation, driving winds from radiatively efficient accretion discs. These winds are powered through magnetic or radiation pressure on lines, free electrons, or dust. They can originate on small scales within the SMBH torus or accretion disc, in the broad-line region, or at larger nuclear scales. AGN winds are observed via broad absorption lines (BALs) in the rest-frame UV, as well as broad optical, UV, and infrared emission lines in the spectra of quasars (e.g. King & Pounds 2015), which are characteristic of strong outflows of molecular and ionized gas (e.g. Richards et al. 2011; Ciccone et al. 2014; Harrison et al. 2014; Hamann et al. 2018; Herrera-Camus et al. 2020; Wylezalek et al. 2020). In particular, BAL winds, found in ~ 20 – 50 percent of the quasar population across all redshifts (e.g. Hewett & Foltz 2003; Dai, Shankar & Sivakoff 2008; Bischetti et al. 2023; Maiolino et al. 2024), are considered a key driver of AGN feedback; they are often highly energetic, with outflow velocities ranging from a few thousand to a few tens of thousands of km s $^{-1}$ (e.g. Fiore et al. 2017; Choi et al. 2020; Rodríguez Hidalgo et al. 2020), and can lead to outflows that extend from parsec to galactic scales, where the latter can reach several kpc from the central SMBH or even further (e.g. Arav et al. 2018; Choi et al. 2022).

By transferring mass, momentum, and energy from the nuclear region to the interstellar medium (ISM) and the circumgalactic medium (CGM) of galaxies, winds from radiatively efficient accretion expel and heat large amounts of gas, and therefore have the ability to impact star formation (e.g. Veilleux et al. 2020, for a review). This AGN feedback mode is thought to be particularly relevant at cosmic noon ($z \simeq 2$), when both cosmic star formation density and quasar activity peak (Madau & Dickinson 2014). This scenario is further supported

by the prevalence of AGN-driven circum-nuclear outflows observed at this epoch (e.g. Förster Schreiber et al. 2014; Genzel et al. 2014), indicating the importance of radiatively efficient AGN feedback (also referred to as quasar-mode feedback in the literature) in the early assembly of galaxies.

The close connection between AGN feedback and galaxy evolution has been explored extensively in semi-analytic models and numerical simulations (for reviews see Somerville & Davé 2015; Naab & Ostriker 2017). AGN feedback is a key ingredient of current galaxy formation models, which is necessary to reproduce observable properties of galaxy populations in the nearby to intermediate redshift Universe. This includes the formation of a population of red and dead massive galaxies, the regulation of the massive end of the galaxy stellar mass function, realistically low stellar content at a given halo mass, and the regulation of the cooling rate and X-ray temperature of gas at galaxy cluster scales (Somerville et al. 2008; Ciotti, Ostriker & Proga 2010; Gaspari et al. 2011; Dubois et al. 2013; Harrison 2017; Henden et al. 2018; Eckert et al. 2021). However, the details of how these processes actually work remain uncertain, due to the impossibility for simulations to capture the dynamic range from BH to galaxy scales. For instance, fully resolving BH accretion would require reaching physical scales of the order of the Schwarzschild radius. For a BH mass of $10^7 M_{\odot}$, this corresponds to 10^{-6} pc, which is computationally prohibitive in cosmological simulations that aim at modelling hundreds of Mpc.

To smooth over the small-scale complexity, the common approach to modelling AGN feedback in cosmological simulations is to rely on subgrid prescriptions, meant to reproduce the consequences of processes that happen at unresolved scales. Initially, AGN feedback was implemented by simply assuming that a fraction of the rest-mass SMBH accreted energy couples thermally to the surrounding gas (Di Matteo, Springel & Hernquist 2005; Springel, Di Matteo & Hernquist 2005), which was shown to be insufficient in reducing cooling and star formation in most massive systems with stellar masses beyond $10^{11} M_{\odot}$ (at the typical resolution achieved in cosmological simulations). To overcome this, more complex prescriptions have been progressively developed (Bourne & Sijacki 2017; Choi et al. 2017; Talbot, Bourne & Sijacki 2021; Huško et al. 2022, 2024, 2025; Rennehan et al. 2024), though some of these are still prohibitively expensive to include in large cosmological simulations. In such coarse resolution simulations, isotropic thermal BH energy deposition remains a widely adopted approach to implementing feedback in the radiatively efficient BH regime (e.g. Di Matteo et al. 2012; Dubois et al. 2014; Hirschmann et al. 2014; Vogelsberger et al. 2014; Springel et al. 2018, but see Choi et al. 2015; Schaye et al. 2015; Davé et al. 2019).

To better regulate gas fractions and star formation in the most massive galaxies, some simulations transition to another AGN feedback mode when the Eddington ratio drops below a certain threshold ($\sim 10^{-2}$), injecting kinetic energy (Dubois et al. 2012; Weinberger et al. 2017), or thermal energy with a higher efficiency (Di Matteo et al. 2012; Steinborn et al. 2015), or in the form of AGN-inflated bubbles of hot gas (Sijacki et al. 2007). Other simulations, which do not distinguish between radiatively efficient and inefficient regimes (Schaye et al. 2015, 2023), distribute the thermal AGN energy from a reservoir of feedback energy by stochastically heating gas elements to a high temperature (Booth & Schaye 2009). To reproduce galaxy observables at $z = 0$, cosmological simulations often enhance the effectiveness of their AGN thermal mode, for instance by stochastically heating a selected number of cells to a large enough temperature to overcome excessive radiative losses (Schaye et al. 2015), or implement a two-mode feedback scheme

with an accretion rate-based switch (Dubois et al. 2012; Weinberger et al. 2017; Davé et al. 2019). The latter approach is adopted for instance in the Illustris-TNG simulations (TNG hereafter; Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019a; Pillepich et al. 2018a; Springel et al. 2018).

Despite adopting a range of different AGN feedback models, cosmological simulations have been successful in reproducing basic observations of galaxy, BH, and AGN populations, predominantly for the low-redshift Universe. Such simulations include MassiveBlack (Di Matteo et al. 2012), Illustris (Vogelsberger et al. 2014), Magneticum (Hirschmann et al. 2014), Horizon-AGN (Dubois et al. 2014), MassiveBlack-II (Khandai et al. 2015), Eagle (Schaye et al. 2015), Romulus (Tremmel et al. 2017), Illustris-TNG (Springel et al. 2018), Simba (Davé et al. 2019), and Flamingo (Schaye et al. 2023). However, their galaxy formation models are usually calibrated to reproduce selected galaxy and BH properties at $z = 0$, which limits their predictive power for the interpretation of galaxy evolution through cosmic time. In particular, Habouzit et al. (2021, 2022) showed that predictions for high-redshift BH and AGN populations from cosmological simulations tend to strongly diverge with increasing redshift. Moreover, most current cosmological simulations underestimate the observed number density of high-redshift massive quiescent galaxies (e.g. Valentino et al. 2023; Weller et al. 2025).

In addition, it remains unclear whether radiatively efficient AGN feedback in the form of thermal energy deposition can generate realistic galactic-scale gaseous outflows, as observed out to $z = 7.5$ (Liu et al. 2024, and references therein). Thus, motivated by observations of BAL winds, Ostriker et al. (2010) and Choi et al. (2012, 2015) first implemented a momentum/kinetic energy based AGN feedback model for generating winds driven by radiatively efficient BH accretion. Using similarly motivated prescriptions (Anglés-Alcázar et al. 2017; Davé et al. 2019; Costa, Pakmor & Springel 2020), recent simulations suggest that quasar-driven outflows can play an important role in galaxies. Notably, they contribute to efficiently regulating star formation and X-ray luminosities (Choi et al. 2018; Costa et al. 2018; Davé et al. 2019), bringing the $z = 0$ galaxy stellar mass function and cold and hot gas fractions in simulations closer to observations, compared to incorporating only continuous thermal quasar energy injection (Springel et al. 2005; Sijacki et al. 2007).

Motivated by this success, it is now within reach to investigate how AGN winds from radiatively efficient accretion affect galaxy and BH evolution out to early cosmic epochs. In this paper, we build on the AGN feedback model of Ostriker et al. (2010) and Choi et al. (2012) and introduce the MISTRAL¹ subgrid model for cosmological simulations, implemented into the AREPO code (Springel 2010). MISTRAL bridges the gap between radiatively efficient AGN wind prescriptions in high-resolution idealized simulations (e.g. Costa et al. 2020; Sivasankaran et al. 2025), and traditional subgrid models implemented in large volume cosmological simulations. As such, MISTRAL allows us to investigate the role of radiatively efficient AGN-driven winds in galaxy evolution. In particular, in this paper, we use MISTRAL to explore to what extent AGN-driven winds affect BH and their host galaxy properties around cosmic noon.

For this purpose, in this study, we couple MISTRAL with the well-tested galaxy formation physics from the TNG cosmological simulations. For reproducing realistic galaxy populations at $z = 0$, the TNG simulations distinguish between a thermal isotropic AGN energy input at high Eddington ratios (referred to as the TNG thermal model), and a kinetic prescription in the low-accretion regime

(introduced by Weinberger et al. 2017). Our aim with MISTRAL is to incorporate a physically motivated prescription for AGN winds from radiatively efficient accretion, in order to understand their impact on the BH–galaxy co-evolution compared to previous, standard AGN feedback models. This is also one of the key goals of the ‘Learning the Universe’ collaboration,² of which this work is a part. With this intention in mind, we have explored two variants of our model implementation: a continuous radial deposition of momentum (MISTRAL-CONTINUOUS), and stochastic bipolar momentum injection (MISTRAL-STOCHASTIC) – the AREPO analogue of the model implemented by Choi et al. (2012) in the GADGET code (Springel 2005). In this study, both versions are used independently, and at all Eddington ratios.

The paper is organized as followed. Section 2 gives an overview of the galaxy formation and BH physics used in the TNG simulations and in this study (Section 2.1), before presenting the details of the MISTRAL model (Section 2.2). Throughout the paper, we study how AGN winds generated by our two versions of MISTRAL impact BH growth and galaxy evolution, compared to the AGN feedback models used in the standard TNG simulations. We split this analysis into two sections. We first explore the effect of MISTRAL within an idealized simulation of a Milky Way-mass galaxy (Section 3). To evaluate the ability of MISTRAL to realistically regulate BHs and galaxies, we analyse galaxy and outflow properties at $z = 2$ in a suite of 15 cosmological zoom simulations (Section 4), whose initial conditions are extracted from the TNG100 simulation. We finally comment on our results in the context of other studies in Section 5, and summarize our conclusions in Section 6.

2 GALAXY FORMATION AND BLACK HOLE PHYSICS MODELS

In this section, we first focus on the models for galaxy and BH physical processes used in the TNG project, which we adopt in the simulations in this paper. Then, we present our new MISTRAL feedback model and its implementation into the astrophysical code AREPO (Springel 2010).

2.1 Galaxy formation physics in TNG

Throughout this paper (aside from our new AGN feedback model), we use the same galaxy formation model as in the published TNG simulations, which reproduce several key galaxy properties and scaling relations at $z = 0$ (see e.g. Pillepich et al. 2018b). These models are implemented in the magnetohydrodynamical moving-mesh code AREPO (Springel 2010; Weinberger, Springel & Pakmor 2020), and explained in detail in Weinberger et al. (2017) and Pillepich et al. (2018a). The ISM is modelled with a two-phase effective equation of state, based on the model developed by Springel & Hernquist (2003). Radiative gas cooling is computed from pre-tabulated cooling functions, depending on gas density, temperature, and metallicity. Stars form stochastically from gas denser than $n_H \simeq 0.13 \text{ cm}^{-3}$, following the Kennicutt–Schmidt relation and the Chabrier (2003) initial mass function. As they evolve, stars chemically enrich the surrounding gas and release mass and energy as Type Ia and Type II supernova and asymptotic giant branch stars. Stellar feedback is modelled as supernova-driven isotropic winds, releasing 10 per cent of the energy thermally, following Marinacci, Pakmor & Springel (2014), and the rest in a kinetic isotropic fashion. Unlike the model

¹Mistral is the name of a cold and strong wind that blows in South of France.

²<https://learning-the-universe.org>

used in the original Illustris simulations (Vogelsberger et al. 2014), supernova-driven winds in TNG have an energy that depends on the star-forming gas cell metallicity, such that the total wind energy decreases with metallicity (equation 3 from Pillepich et al. 2018a). The stellar-driven outflows are mediated by wind particles, launched at a velocity that scales with the dark matter velocity dispersion. These wind particles are decoupled from the dense local ISM, and propagate until they encounter a gas cell with a low density (0.05 times lower than the density threshold for star formation), into which they transfer their mass, momentum, metal, and energy.

In the TNG simulations, SMBHs are modelled as collisionless sink particles. They are seeded with an initial mass $M_{\text{BH}} = 1.2 \times 10^6 M_\odot$ in haloes with a dark matter mass above $7.4 \times 10^{10} M_\odot$ that have not been previously seeded with a BH. Throughout the simulation, the SMBHs are (re)located at the position of the halo potential minimum every time-step. They can grow by merging with other BHs and by accreting gas, as we explain in more detail in the next section. AGN feedback operates in a dual mode: a thermal mode at high Eddington ratios (the radiatively efficient BH regime, Section 2.1.2), and a kinetic mode at low Eddington ratios (radiatively inefficient regime, Section 2.1.3).

2.1.1 Black hole accretion

Black hole accretion in the TNG simulations follows the Eddington-limited Bondi–Hoyle–Lyttleton prescription (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944; Bondi 1952), where the Bondi–Hoyle–Lyttleton accretion rate \dot{M}_{Bondi} and the Eddington rate \dot{M}_{Edd} are defined as follows:

$$\dot{M}_{\text{Bondi}} = \frac{4\pi G^2 M_{\text{BH}}^2 \rho}{c_s^3} \quad (1)$$

$$\dot{M}_{\text{Edd}} = \frac{4\pi G M_{\text{BH}} m_p}{\epsilon_r \sigma_T c} \quad (2)$$

Here, G is the universal gravitational constant, m_p is the proton mass, ϵ_r is the radiative accretion efficiency, σ_T is the Thompson cross-section, and c is the vacuum speed of light. M_{BH} is the black hole mass, and ρ and c_s are the average density and effective sound speed³ of the gas near the black hole, respectively. All the gas properties are calculated within the BH smoothing length h_{BH} , by performing a kernel-weighted average over a prescribed number of gas cells $n_{\text{BH,ngb}}$ that are the nearest BH neighbours, such that

$$n_{\text{BH,ngb}} = \sum_i \frac{4\pi h_{\text{BH}}^3}{3} \frac{m_i}{m_{\text{target}}} w_k(r_i), \quad (3)$$

where m_i/m_{target} is the mass of the gas cell i compared to the target gas mass resolution enforced by the refinement scheme of the AREPO code. $w_k(r_i)$ is the cubic spline smoothing kernel weight associated with this gas cell i (in units of inverse volume), which scales with its distance r_i to the BH. Essentially, this means that the BH smoothing length may vary over time, in order to enclose a similar number of cells $n_{\text{BH,ngb}} \pm 4$. In TNG50, TNG100, and TNG300 (the TNG simulations with box side lengths of $\sim 50, 100$, and 300 cMpc, respectively), $n_{\text{BH,ngb}}$ is set to 512, 256, and 128 (see also appendix B in Weinberger et al. 2017). In what follows, we refer to the sphere of radius h_{BH} as the BH smoothing volume. In this work, the value of h_{BH} varies from a few hundred pc to a few kpc. This parameter increases when the gas density near the BH decreases (as a result

of the gas cell width expanding), which preferentially occurs after efficient AGN feedback episodes.

To first order, the rate of inflowing gas mass on to the BH corresponds to the Bondi accretion rate capped at the Eddington limit, that is,

$$\dot{M}_{\text{BH,inf}} = \min(\dot{M}_{\text{Bondi}}, \dot{M}_{\text{Edd}}) \quad (4)$$

In addition, Vogelsberger et al. (2013) introduced a modification to the BH accretion rate to prevent the formation of a hot bubble of low-density gas around the BH, whenever there is no star-forming gas in its immediate vicinity. As such a configuration would overestimate the BH accretion rate, $\dot{M}_{\text{BH,inf}}$ is lowered by a factor $(P_{\text{ext}}/P_{\text{ref}})^2$ whenever P_{ext} (defined as the kernel-weighted pressure of gas in the BH smoothing volume) is lower than a reference pressure $P_{\text{ref}} = (\gamma - 1)\rho_{\text{sfr}}u_{\text{eq}}$. In this equation, γ is the adiabatic index of the gas, ρ_{sfr} is the gas density threshold beyond which star formation happens, and u_{eq} is the equilibrium thermal energy when the cooling losses in the BH smoothing volume balance the AGN feedback energy injected.

When flowing towards the SMBH, gas forms an accretion disc where some fraction of the inflowing mass (with $\epsilon_r = 0.2$ in TNG) is converted into radiation, such that the actual BH accretion rate \dot{M}_{BH} is

$$\dot{M}_{\text{BH}} = (1 - \epsilon_r) \dot{M}_{\text{BH,inf}} \quad (5)$$

Numerically, the SMBH grows continuously according to equation (5), by draining gas mass from the BH smoothing volume cells, in a kernel and volume weighted manner.

2.1.2 AGN feedback in the radiatively efficient regime

As explained in detail by Weinberger et al. (2017), AGN feedback in TNG operates in a dual mode, with different AGN energy injection schemes depending on the BH accretion regime. This prescription is motivated by the physical properties of BH accretion discs: when the SMBH accretes at or above a few 10^{-2} of \dot{M}_{Edd} , the AGN is in a radio-quiet radiative mode, releasing energy from an optically thick and geometrically thin accretion disc (Shakura & Sunyaev 1973). Because this system radiates its heat very efficiently, this is also referred to as the SMBH radiatively efficient regime. Otherwise, the AGN energy is deposited in an optically thin and geometrically thick, radiatively inefficient accretion structure, as described by the Advection-Dominated Accretion Flow model (Narayan & Yi 1994) or the Adiabatic Inflow–Outflow Solutions (Blandford & Begelman 1999). As noted above, super-Eddington accretion is not accounted for in TNG, and we also do not allow it throughout the work presented here.

To determine whether the SMBH is in a radiatively efficient or inefficient state, the common procedure is to compare the Eddington ratio $f_{\text{Edd}} = \dot{M}_{\text{Bondi}}/\dot{M}_{\text{Edd}}$ to a certain value $\chi = 10^{-3}$ – 10^{-1} (e.g. Sijacki et al. 2007; Merloni & Heinz 2008; Dubois et al. 2012), as determined from observations of soft-to-hard state transitions of X-ray binaries (Maccarone 2003). Instead of adopting a constant value for χ , the TNG simulations adopt a mass-dependent parametrization, in order to promote (prevent) the triggering of the radiatively inefficient feedback mode for high-mass (low-mass) BHs, where

$$\chi = \min \left[0.002 \times \left(\frac{M_{\text{BH}}}{10^8 M_\odot} \right)^2, 0.1 \right] \quad (6)$$

Whenever $f_{\text{Edd}} > \chi$, energy is deposited following the TNG thermal model that we also refer to as the ISOTROPIC THERMAL

³The effective gas sound speed includes the contribution from the thermal gas speed $c_{s,\text{th}}$ and the speed of the Alfvén waves c_A , such that $c_s^2 = c_{s,\text{th}}^2 + c_A^2$.

model. In this radiatively efficient regime, a fraction ϵ_f of the luminosity L_r radiated during the accretion process thermally couples to the gas. In the TNG simulations, $\epsilon_r = 0.2$, and $\epsilon_f = 0.1$. Per injection event (during a time-step Δt), a total energy rate \dot{E}_{BH} is distributed thermally, spherically, and isotropically, in a kernel-weighted fashion, to the $n_{\text{BH,ngb}}$ gas cells enclosed in the BH smoothing volume, such that

$$\dot{E}_{\text{BH}} = \epsilon_f L_r = \epsilon_f \epsilon_r \dot{M}_{\text{BH,inf}} c^2 \quad (7)$$

Therefore, each gas cell i is given a thermal energy increment ΔE_i , where m_{tot} is the total kernel-weighted gas mass enclosed in the BH smoothing volume:

$$\Delta E_i = \dot{E}_{\text{BH}} \Delta t w_k(r_i) \frac{m_i}{m_{\text{tot}}} \quad (8)$$

Particularly relevant in the high-accretion regime, when the bolometric luminosity is the highest, Vogelsberger et al. (2013) introduced what they refer to as a radiative electro-magnetic AGN feedback channel. This channel consists of depositing thermal energy, and is included in the TNG simulations. More specifically, this phenomenological model aims at capturing the effect of the ionizing radiation emitted by SMBHs on the thermal state of the gas. For this purpose, the fraction of AGN bolometric luminosity that does not couple to the gas is combined with a redshift-dependent UV background (Faucher-Giguère et al. 2009), which is then used to infer metal line cooling and heating rates from pre-tabulated tables. To determine their net cooling rate, each gas cell is assigned a total bolometric intensity, derived from the obscured bolometric luminosity of all the BHs from the simulation, depending on their distance to the gas cell. This radiative feedback channel renders the TNG thermal feedback more efficient, as star formation is slightly more suppressed in the high-accretion regime compared to when no AGN radiative feedback is included (see fig. 15 from Vogelsberger et al. 2013).

2.1.3 AGN feedback in the radiatively inefficient regime

When $f_{\text{Edd}} < \chi$ (defined in equation 6), AGN feedback in TNG consists of momentum kicks to the BH's neighbouring gas cells, meant to mimic the effects of kinetic winds and jets. In this low accretion state, ϵ_f^4 is a variable feedback coupling efficiency defined as

$$\epsilon_f = \min \left(\frac{\rho}{0.01 \rho_{\text{SF}}}, 1 \right) \quad (9)$$

Here, $\rho_{\text{SF}} \simeq 10^{-23} \text{ g cm}^{-3}$ is the density threshold for star formation. This non-constant parametrization of ϵ_f in the radiatively inefficient regime weakens the coupling of the BH energy to the surrounding gas at low density.

After several BH accretion events, when a minimum AGN energy of $10 \sigma_{\text{DM}}^2 m_{\text{tot}} \text{ erg}$ has accumulated (where σ_{DM} is the 1D dark matter velocity dispersion), a feedback event takes place. The total kinetic feedback energy available ΔE_{BH} (which corresponds to the sum of $\epsilon_f L_r$ accumulated over several time-steps) is distributed to all gas cells from the BH smoothing volume, such that each cell i receives a momentum kick Δp_i in a random direction:

$$\Delta p_i = m_i \sqrt{2 \frac{\Delta E_{\text{BH}} w_k(r_i)}{\rho}} \quad (10)$$

⁴We use a slightly different parametrization of ϵ_f than the one in Weinberger et al. (2017), who combine ϵ_r and ϵ_f into a single efficiency parameter $\epsilon_{f,\text{kin}}$.

The wind direction is the same for all gas cells per injection event, so that momentum is only conserved on a time-averaged basis. We also refer to this TNG kinetic AGN feedback as the RANDOM KINETIC model.

2.2 The MISTRAL model

When gas flows from galaxies down to their SMBH, radiation is emitted, predominantly in dense regions close to the BH. There, in the so-called broad-line region (and, to a lesser extent, in the narrow-line region), winds can be driven by line, radiation, and magnetic pressure, carrying mass, momentum, and energy that affect their surroundings. To model such AGN-driven winds, Ostriker et al. (2010) and Choi et al. (2012) developed a subgrid AGN feedback scheme that they implemented into the smoothed-particle hydrodynamics (SPH) code GADGET (Springel 2005). Based on their methodology, we implement the MISTRAL model into the moving-mesh code AREPO (Springel 2010), as explained in this section. We note that, in this work, MISTRAL operates at all Eddington ratios, in both the radiatively efficient and inefficient regimes (i.e. we do not use any of the TNG AGN feedback models with MISTRAL). In order to focus on a set-up as close as possible to the one adopted by Choi et al. (2012), we do not include the modification in BH accretion introduced by Vogelsberger et al. (2013), nor their radiative AGN feedback channel (both of which are used in the TNG simulations). In this work, unlike Choi et al. (2012), we do not consider the radiative feedback from the X-ray radiation associated with BH accretion.

On its journey towards an SMBH, not all the inflowing gas mass reaches its final destination: instead, some of the gas will be entrained into a wind, outflowing at a rate $\dot{M}_{\text{BH,wind}}$. Therefore, the BH accretion rate can be rewritten as

$$\dot{M}_{\text{BH}} = \dot{M}_{\text{BH,inf}} - \dot{M}_{\text{BH,wind}}, \quad (11)$$

where $\dot{M}_{\text{BH,inf}}$ is defined in equation (4) as the usual Eddington-limited Bondi accretion rate. Equation (12) defines the energy released by the SMBH as \dot{E}_{BH} , where v_w is the average wind velocity. In this equation, we relate the kinetic energy deposition of the outflowing material to the accretion rate via the AGN feedback efficiency parameter ϵ_w , which translates how efficiently the accreted matter turns into wind energy.

$$\dot{E}_{\text{BH}} = \epsilon_w \dot{M}_{\text{BH}} c^2 = \frac{1}{2} \dot{M}_{\text{BH,wind}} v_w^2 \quad (12)$$

Correspondingly, the BH wind momentum rate \dot{p}_{BH} can be defined as

$$\dot{p}_{\text{BH}} = \dot{M}_{\text{BH,wind}} v_w \quad (13)$$

Using equation (12), we can introduce, as done by Ostriker et al. (2010) and Choi et al. (2012), a dimensionless parameter ψ such that

$$\psi = \frac{\dot{M}_{\text{BH,wind}}}{\dot{M}_{\text{BH}}} = \frac{2 \epsilon_w c^2}{v_w^2} \quad (14)$$

Using this parameter, mass, energy, and momentum conversion equations can be written as follows:

$$\dot{M}_{\text{BH}} = \frac{1}{1 + \psi} \dot{M}_{\text{BH,inf}}, \quad (15)$$

$$\dot{M}_{\text{BH,wind}} = \frac{\psi}{1 + \psi} \dot{M}_{\text{BH,inf}}, \quad (16)$$

$$\dot{E}_{\text{BH}} = \epsilon_w \frac{1}{1 + \psi} \dot{M}_{\text{BH,inf}} c^2, \quad (17)$$

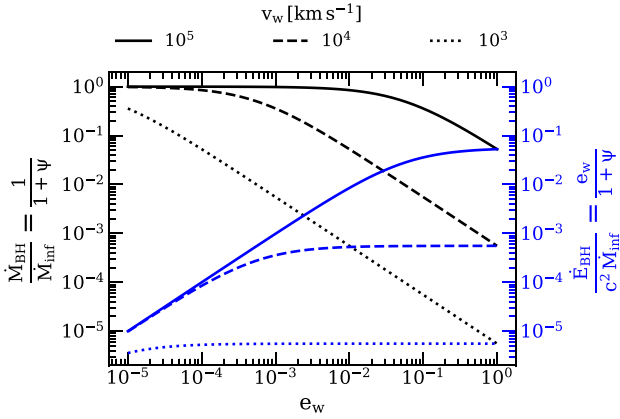


Figure 1. Fraction of the inflowing gas mass rate \dot{M}_{inf} that contributes to the BH mass accretion rate \dot{M}_{BH} (black lines, left y-axis) and to the BH energy rate \dot{E}_{BH} (blue lines, right y-axis) as a function of the AGN wind feedback coupling efficiency ϵ_w . The solid, dashed, and dotted lines, respectively, correspond to an average wind velocity of 10^5 , 10^4 , and 10^3 km s^{-1} .

$$\dot{p}_{\text{BH}} = \frac{\psi}{1 + \psi} \dot{M}_{\text{BH,inf}} v_w \quad (18)$$

Using X-ray and UV absorption line spectroscopy, studies of BAL quasars and ultra-fast outflows suggest that AGN winds velocities can span from thousands to tens of thousands of km s^{-1} , with inferred kinetic efficiencies estimated to lie in the range 10^{-4} – 10^{-3} (Tombesi et al. 2011, 2013; Matzeu et al. 2023). However, these observational constraints remain uncertain, and may vary depending on BH masses and accretion rates. For simplicity, and given the observational uncertainties, this work treats these two parameters as constants, with the aim of capturing the average impact of AGN winds on galaxy evolution at resolved scales. For typical values $\epsilon_w = 5 \times 10^{-3}$ and $v_w = 10^4 \text{ km s}^{-1}$, $\psi = 9$, such that a momentum flux $\dot{p}_{\text{BH}} = 0.3 L_{\text{r}}/c$ is injected. This also means that 10 per cent of the inflowing gas mass rate will be actually accreted on to the BH, driving the remaining 90 per cent as a (subgrid) wind. More generally, Fig. 1 shows the fraction of inflowing gas mass accreted by the BH (in black, left y-axis) and the fraction of inflowing rest-mass energy released as a wind (in blue, right y-axis) as a function of ϵ_w . The three sets of curves correspond to an average wind velocity of 10^5 , 10^4 , and 10^3 km s^{-1} (solid, dashed, and dotted lines). By definition, the energy released is lower for low ϵ_w values, allowing the amount of BH mass accreted to be correspondingly higher. In addition, at a fixed ϵ_w , a higher wind velocity v_w will drive a smaller fraction of the inflowing gas into the outflowing wind, and a higher fraction into actual BH accretion (see equations 12 and 14). Thus, these two parameters control the amount of mass and energy loading in our model, and both the fractions of mass accreted and energy released are more sensitive to v_w than to the coupling efficiency ϵ_w .

In this paper, the BH inflowing gas mass rate is determined following the Bondi prescription. The total gas mass accreted by the BH during a given time-step Δt follows equation (15), and this gas mass is continuously swallowed from gas cells that belong to the BH smoothing volume, as described in Section 2.1.2. Once gas is drained, the BH energy resulting from the accretion process is distributed over the remaining gas mass, at the same time-step. This is done in a kinetic fashion, either in a continuous way with MISTRAL-CONTINUOUS, or in a stochastic manner with MISTRAL-STOCHASTIC. MISTRAL-CONTINUOUS provides a kinetic alternative to the TNG

thermal model, as both models release AGN feedback continuously in all the cells from the BH smoothing volume. Conversely, MISTRAL-STOCHASTIC inputs momentum only in a subset of stochastically selected cells. While MISTRAL-STOCHASTIC is a closer analogue to the feedback model developed by Ostriker et al. (2010) and Choi et al. (2012), MISTRAL-CONTINUOUS will be used to illustrate the transition from (1) a purely thermal to a kinetic, radiatively efficient AGN feedback model and (2) a continuous to a stochastic feedback injection scheme. We now review the details of these two flavours of the MISTRAL model, which are tested and used separately in this study. Both versions of MISTRAL are schematically illustrated and compared to the TNG AGN feedback models in Fig. 2.

2.2.1 MISTRAL-CONTINUOUS: continuous radial momentum injection

In MISTRAL-CONTINUOUS, gas cells are kicked radially away from the BH, at a velocity v_{kick} , such that both momentum and energy are conserved. For this purpose, we adopt the following procedure. We first adjust the direction of the kicks by subtracting the centre of mass to each radial vector \mathbf{r}_i , so that each cell i is kicked following the unit vector $\mathbf{n}_i = \mathbf{r}_i / \|\mathbf{r}_i\| - \sum_i m_i w_k(r_i) \mathbf{r}_i / \sum_i m_i w_k(r_i)$. To favour the emergence of a bipolar outflow out of this spherically symmetric injection scheme, we define v_{kick} such that a higher velocity increment is imparted to gas cells that are aligned with the mass-weighted angular momentum \mathbf{j}_{tot} of the smoothing volume (close to the BH):

$$\Delta p_i = m_i v_{i,\text{kick}} = m_i w_k(r_i) w_{\text{AM}}(r_i) v_{\text{kick}} \quad (19)$$

Here, $w_{\text{AM}}(r_i) = \|\mathbf{r}_i \cdot \mathbf{j}_{\text{tot}}\|$ corresponds to an angular momentum weight, and v_{kick} is defined to ensure energy conservation such that

$$E_{\text{BH}} = \sum_i \frac{(\mathbf{p}_i + \Delta \mathbf{p}_i)^2}{2m_i} - \frac{\mathbf{p}_i^2}{2m_i}, \quad (20)$$

where $\mathbf{p}_i = m_i \mathbf{v}_i$ is the current momentum of cell i (before the injection of AGN feedback). Solving equation (20) and determining the value of v_{kick} is equivalent to solving the quadratic equation $a_{\text{tot}} v_{\text{kick}}^2 + b_{\text{tot}} v_{\text{kick}} - E_{\text{BH}} = 0$, which gives

$$v_{\text{kick}} = \frac{b_{\text{tot}} + \sqrt{b_{\text{tot}}^2 + 4E_{\text{BH}}a_{\text{tot}}}}{2a_{\text{tot}}}, \quad (21)$$

$$a_{\text{tot}} = \sum_i \frac{m_i [w_k(r_i) w_{\text{AM}}(r_i) \mathbf{n}_i]^2}{2}, \quad (22)$$

$$b_{\text{tot}} = \sum_i \frac{m_i w_k(r_i) w_{\text{AM}}(r_i) \mathbf{n}_i \cdot \mathbf{v}_i}{2} \quad (23)$$

Once determined, v_{kick} is then used to update each cell velocity, momentum, and energy. While this prescription ensures conservation of energy, it does not strictly distribute the BH energy to the gas mass expected from equation (16). This is because we implicitly assume that all gas mass in the BH vicinity which is not accreted will be kicked, at a certain velocity determined such that we distribute the total AGN energy. As a consequence, if there is a large number of cells over which we must distribute a low BH energy, the effective velocity $v_{i,\text{kick}}$ imparted to a cell (corresponding to the the wind velocity at the scale of h_{BH}) may be orders of magnitude lower than the input wind velocity v_w (corresponding to the wind velocity at the accretion scale). In order to inject the correct amount of momentum, as determined by equations (13) and (18), when equation (16) is not fulfilled, one solution might be to change the amount of gas mass over which momentum is injected (e.g. by changing the size of the BH

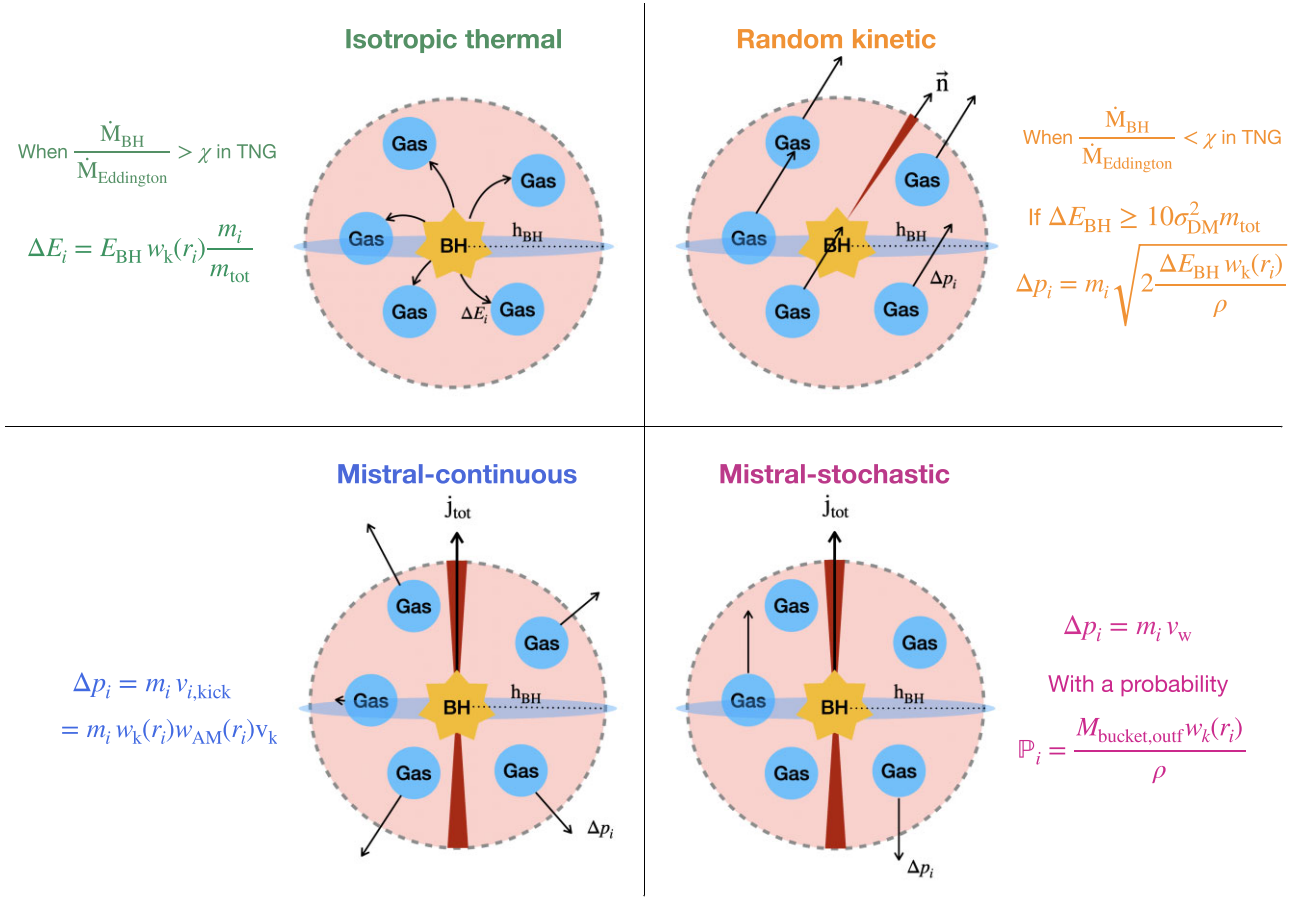


Figure 2. Clockwise: schematic illustration of the ISOTROPIC THERMAL and RANDOM KINETIC models used in TNG, MISTRAL-STOCHASTIC, and MISTRAL-CONTINUOUS. For each model, we represent the BH sink particle as a star, surrounded by gas cells depicted by circles. The dashed circle shows the extent of the spherical smoothing volume with radius h_{BH} . With the ISOTROPIC THERMAL model, each gas cell receives a mass-weighted fraction of the BH energy in a thermal form (equation 8). With the RANDOM KINETIC model, gas cells receive a momentum kick in the same random direction, provided that enough BH energy has accumulated (equation 10). In the TNG simulations, this model acts when the Eddington ratio is below a certain threshold χ (equation 6), and the ISOTROPIC THERMAL model operates otherwise. With MISTRAL-CONTINUOUS, cells are kicked radially away from the BH, with a velocity weighted by the alignment of the cell to the angular momentum of the gas surrounding the BH (equation 19). With MISTRAL-STOCHASTIC, cells are kicked at a velocity v_w with a probability \mathbb{P} (equation 25), in a direction parallel or antiparallel to the gas angular momentum.

smoothing volume), which we defer for future work.⁵ In this paper, to ensure a meaningful comparison with the TNG feedback models, we instead choose to distribute the BH energy to all gas cells within the BH smoothing volume, even though the momentum injection then differs from what is specified by equations (13) and (18).

⁵An alternative approach to conserve both the total energy and momentum exactly (regardless of the gas mass involved in the feedback event) would be to inject momentum via kinetic kicks, and then distribute any residual energy as thermal energy. However, in cases where the gas mass receiving feedback is significantly larger than the target wind mass, this hybrid scheme would result in feedback being dominated by thermal energy. Since our goal with MISTRAL-CONTINUOUS is to explicitly investigate the impact of kinetic feedback, in contrast to more common thermal implementations in the radiatively efficient regime, we deliberately choose to preserve energy at the expense of exact momentum injection.

2.2.2 MISTRAL-STOCHASTIC: stochastic bipolar momentum injection

MISTRAL-STOCHASTIC aims at being the AREPO equivalent of the AGN wind model from Ostriker et al. (2010) and Choi et al. (2012). Instead of continuously injecting momentum, we now switch to a stochastic approach: gas cells from the BH smoothing volume are kicked at the chosen input velocity v_w , with a probability \mathbb{P}_i for each cell i :

$$\mathbb{P}_i = \frac{\dot{M}_{\text{BH,wind}} \Delta t w_k(r_i)}{\rho} \quad (24)$$

This ensures that we will impart momentum only to the fraction of $\dot{M}_{\text{BH,inf}}$ that should be launched in a wind, as defined in equation (16). Each cell eligible for receiving momentum is then kicked with a velocity v_w in a direction parallel (antiparallel) to the mass-weighted angular momentum of the gas j_{tot} of the BH smoothing volume, if the gas cell is above (below) the mid-plane perpendicular to j_{tot} .⁶

⁶We checked that randomly kicking the cells parallel or antiparallel to j_{tot} (as opposed to imposing the direction of the velocity kick based on the position of the cell relative to the galaxy disc) barely changes the impact of the MISTRAL-STOCHASTIC model on galaxy and BH properties.

Table 1. Initial conditions and properties of the idealized galaxy simulations. Left: M_{halo} : dark matter halo mass, M_{disc} : baryonic disc mass (gas + stars), $M_{\text{BH,seed}}$: BH seed mass, R_{vir} : virial radius, f_{gas} : gas disc fraction, Z_{disc} : disc metallicity. Right: L_{box} : length of the simulated box, m_{DM} : dark matter mass resolution, m_{baryons} : mean baryon mass resolution, $\epsilon_{\text{DM,stars}}$: Plummer equivalent gravitational softening of the collisionless component, ϵ_{gas} : minimum value of the gravitational softening of the gas, $n_{\text{BH,ngb}}$: weighted number of gas cells in the BH smoothing volume.

Galaxy properties		Simulation properties	
M_{halo}	$1.5 \times 10^{12} M_{\odot}$	L_{box}	3000 kpc
M_{disc}	$6.1 \times 10^{10} M_{\odot}$	m_{DM}	$5 \times 10^5 M_{\odot}$
$M_{\text{BH,seed}}$	$5 \times 10^7 M_{\odot}$	m_{baryons}	$8 \times 10^4 M_{\odot}$
R_{vir}	230 kpc	$\epsilon_{\text{DM,stars}}$	195 pc
f_{gas}	0.07	ϵ_{gas}	50 pc
Z	0.05–1 Z_{\odot}	$n_{\text{BH,ngb}}$	512

Because we cannot kick a fraction of the gas mass contained in a cell (this would require changing the refinement strategy), we allow an AGN feedback event only if $\dot{M}_{\text{BH,wind}} \Delta t$ is larger than the gas mass resolution. Therefore, at a given time-step, the probability \mathbb{P}_i requires a sufficiently high BH accretion rate to have a chance to be satisfied. To circumvent this issue, we redefine the effective probability for gas cell i to be outflowing as

$$\mathbb{P}_i = \frac{M_{\text{bucket,outf}} w_k(r_i)}{\rho}, \quad (25)$$

where $M_{\text{bucket,outf}} = \sum_j (\dot{M}_{\text{BH,outf},j} \Delta t_j)$ is a bucket variable used to keep track of the outflowing gas mass expected after each time-step Δt_j . When a gas cell is kicked, we remove its mass from $M_{\text{bucket,outf}}$ at the end of the time-step. Therefore, the probability of receiving a velocity increment at a given time-step remains the same for all the eligible gas cells. In order to ensure that we effectively kick the outflowing gas mass expected on a time-average basis, gas cells are eligible for being kicked only if their mass is lower than the $M_{\text{bucket,outf}}$. We eventually update the velocity, momentum, and internal energy of the cells that have been kicked. Given the stochastic nature of this model, the injected energy and momentum correspond, on average over several AGN feedback events, to the values from equations (17) and (18), and typically differ by no more than 20 per cent per injection event.

3 IDEALIZED MILKY WAY-MASS SIMULATIONS

A major goal of this work is comparing the impact of MISTRAL on galaxy properties, compared to the AGN feedback models used in the TNG simulations. For this purpose, we first use idealized Milky Way-mass simulations in this section, before applying a similar analysis on $z = 2$ cosmological zoom simulations (Section 4).

3.1 Set-up of the idealized simulations

To examine the impact of MISTRAL winds on galaxies, we first use simulations of an idealized Milky Way-mass disc-dominated galaxy as a controlled experiment. Specifically, the initial conditions correspond to those of galaxy m12 from Su et al. (2019), converted from GIZMO (Hopkins 2015) to AREPO-readable data. We summarize some of the galaxy properties below and in Table 1, and refer the reader to section 2.2 of Su et al. (2019) for a more complete description.

The galaxy disc is embedded in a box of 3000 kpc in width, and is hosted in a $1.5 \times 10^{12} M_{\odot}$ dark matter halo, which follows a spherical isotropic NFW density profile (Navarro, Frenk & White 1996). The dark matter halo has a virial radius⁷ of $R_{\text{vir}} = 230$ kpc, an NFW scale length of 20.4 kpc, a concentration parameter $c = 12$, and a maximum circular velocity of 174 km s^{-1} . The galaxy contains a stellar bulge that follows a Hernquist (1990) profile, and exponential rotation-supported gas, and stellar discs, in which the initial temperature is at pressure equilibrium. In order to roughly reproduce the observed Milky Way profile (Miller & Bregman 2013, 2015), the initial conditions include a hydrostatic spherical gaseous halo with a mass of $1.5 \times 10^{11} M_{\odot}$, following a beta profile with a scale radius of 20.4 kpc and $\beta = 0.5$.

Dark matter is modelled by collisionless particles with a mass of $m_{\text{DM}} = 5 \times 10^5 M_{\odot}$. The baryon mass resolution (for both gas and stellar particles) has a mean value of $m_{\text{baryons}} = 8 \times 10^4 M_{\odot}$, and can vary by a factor of 2 for gas upon cell (de)refinement or by mass-losses for the stellar particles. The Plummer equivalent gravitational softening of the collisionless component (dark matter, stellar, and wind particles) is fixed to $\epsilon_{\text{DM,stars}} = 195$ pc, and the gravitational softening of the gas is set to 2.5 times the cell radius, with a minimum value of $\epsilon_{\text{gas}} = 50$ pc.

The galaxy includes exponential gas and stellar discs, with a scale length of 6 and 3 kpc, respectively, and a scale height of 0.3 kpc for both. The galaxy also features an extended spherical gas halo with a scale radius of 20 kpc. In what follows, we define the galaxy disc as a cylinder of $R_{\text{disk}} = 20$ kpc in radius and $h_{\text{disk}} = 2$ kpc in height. Initially, this disc has a gas fraction $f_{\text{gas}} = M_{\text{gas}}/(M_{\text{gas}} + M_{\text{star}}) \simeq 7$ per cent. It contains a gas mass $M_{\text{gas}} = 4.2 \times 10^9 M_{\odot}$, hosts a stellar mass $M_{\text{star}} = 5.7 \times 10^{10} M_{\odot}$, and has a metallicity $Z = Z_{\odot} = 0.02$ at the core. Gas metallicity decreases as a function of radius r as $Z_{\odot} \times (0.05 + 0.95/(1 + (r/R_{\text{disk}})^{1.5}))$, until it reaches $Z = 0.001$ outside the disc. We first run the initial conditions adiabatically, without feedback and cooling processes, for 100 Myr, in order to relax the AREPO mesh before turning on star formation and stellar feedback for another 500 Myr. These two steps, respectively, allow the AREPO grid and the galaxy disc to stabilize, so that the galaxy reaches an equilibrium state, before we allow the central SMBH to accrete gas and release energy. Throughout this section, we refer to $t = 0$ Myr as the time at which BH physics is enabled. During the 500 Myr pre-BH phase, a stellar mass of $9 \times 10^9 M_{\odot}$ forms, so that at $t = 0$ Myr, the galaxy disc has a gas mass $M_{\text{gas}} = 3.3 \times 10^9 M_{\odot}$, a total stellar mass $M_{\text{star}} = 6.6 \times 10^{10} M_{\odot}$, and a gas fraction $f_{\text{gas}} = 0.05$.

The SMBH is initially positioned at the centre of the simulation box, and at the centre of the galaxy disc. The SMBH has an initial mass of $M_{\text{BH}} = 5 \times 10^7 M_{\odot}$, lying between the local BH mass–stellar mass scaling relations from Kormendy & Ho (2013) and Reines & Volonteri (2015). Although this BH seed mass is an order of magnitude greater than the SMBH in the Milky Way, it enables the BH to grow without requiring an accretion boost parameter, which would otherwise be necessary to examine the effects of different feedback prescriptions, given the short duration and idealized nature of these simulations. The BH grows via gas accretion and releases energy as described in Section 2, depending on the AGN feedback model used. Specifically, in this section, we study the following set of simulations:

- (i) NO BH: without BH accretion and feedback.

⁷We define the virial radius as the radius of the sphere that encloses a mean density that is 200 times the critical density of the Universe.

(ii) **ISOTROPIC THERMAL**: isotropic thermal BH energy release, corresponding to the thermal feedback used in TNG at high Eddington ratios (in the radiatively efficient regime). We set $\epsilon_r = 0.1$ and $\epsilon_f = 5 \times 10^{-3}$, and use this feedback model at any Eddington ratio. Note that here, $\epsilon_r \times \epsilon_f = 5 \times 10^{-4}$, unlike 0.02 in the TNG simulations.

(iii) **RANDOM KINETIC**: random AGN winds, corresponding to the AGN kinetic feedback used in TNG at low Eddington ratios (in the radiatively inefficient regime). Similarly as for the **ISOTROPIC THERMAL** simulation, we use $\epsilon_r = 0.1$ and $\epsilon_f = 5 \times 10^{-3}$, and we use the **RANDOM KINETIC** model at all Eddington ratios (unlike what is done in the fiducial TNG simulations).

(iv) **FIDUCIAL TNG**: default TNG feedback models and parameters. In the radiatively efficient and inefficient regimes, $\epsilon_r = 0.1$ or follows equation (9), respectively, and $\epsilon_f = 0.2$. In practice, the SMBH never enters the radiatively inefficient regime during the simulation runtime, such that $\epsilon_r \times \epsilon_f = 0.02$ at all times. This simulation is the only one of the set that includes the radiative AGN feedback developed by Vogelsberger et al. (2013).

(v) **MISTRAL-CONTINUOUS**: continuous radial momentum injection mode of MISTRAL. We set $\epsilon_w = 5 \times 10^{-4}$ and $v_w = 10^4 \text{ km s}^{-1}$.

(vi) **MISTRAL-STOCHASTIC**: stochastic bipolar momentum injection mode of MISTRAL. We set $\epsilon_w = 5 \times 10^{-4}$ and $v_w = 10^4 \text{ km s}^{-1}$.

In all simulations that include BH physics, we adopt $n_{\text{BH,ngb}} = 512$ as the weighted number of gas cells in the BH smoothing volume. Except for the simulation with the **FIDUCIAL TNG** model, we use the same feedback coupling efficiencies in order to better determine how injecting a given BH rest-mass accreted energy using different schemes impacts galaxy properties. We also checked (but do not show) that adopting the same net efficiency for all simulations (i.e. $\dot{E}_{\text{BH}}/\dot{M}_{\text{BH,inf}}c^2 = 2.6 \times 10^{-4}$) does not impact the conclusions drawn in this section. Given our choice of parameters, 52 per cent of the inflowing gas mass is accreted in the simulations with MISTRAL, which goes up to 90 per cent in the simulations with the **ISOTROPIC THERMAL** and **RANDOM KINETIC** models, and to 80 per cent in the simulation with the **FIDUCIAL TNG** set-up. We note that our choice of $v_w = 10^4 \text{ km s}^{-1}$ is consistent with the velocity of AGN-driven winds measured at small scales, such as the broad-line region (e.g. Tombesi et al. 2011; Matzeu et al. 2023) and the inner edge of the dusty torus (Choi et al. 2020). This velocity however differs from that of the outflow measured at kpc scales, which may reach lower values (e.g. Choi et al. 2022). We also emphasize that in this section, our aim is to illustrate how the continuous or stochastic injection of thermal or kinetic energy affects galaxy properties. We note that the results from the **FIDUCIAL TNG** idealized simulation that use the same coupling efficiency parameters as TNG are presented solely for reference. We do not apply these parameters to MISTRAL since doing so would result in 97 per cent of the inflowing mass being expelled as a wind, leaving only a small amount available for black hole accretion. Instead, we adopt a similar parametrization as that used by Choi et al. (2012) in a similar idealized context, with a wind efficiency within the range of observational estimates.

3.2 Comparison of MISTRAL to the TNG AGN feedback models

Throughout this section, we explore the impact of MISTRAL on gas properties (Section 3.2.1), star formation and BH growth (Section 3.2.2), and outflow properties (Section 3.2.3), compared to the **ISOTROPIC THERMAL** and **RANDOM KINETIC** models used in TNG.

3.2.1 Effect of AGN feedback on gas properties

We begin with a qualitative comparison of the effects of AGN feedback on gas properties. Fig. 3 shows edge-on maps of gas density and temperature, for the simulations without BH physics (**NO BH**), and with the **FIDUCIAL TNG**, **RANDOM KINETIC**, **ISOTROPIC THERMAL**, **MISTRAL-CONTINUOUS**, and **MISTRAL-STOCHASTIC** AGN feedback models. The maps are plotted at a time $t \simeq 586 \text{ Myr}$, except in the simulation with the **RANDOM KINETIC** model, for which we show them just after an AGN feedback event occurs, at $t \simeq 630 \text{ Myr}$.

In the absence of AGN feedback and with the **ISOTROPIC THERMAL** model, gas cycles around the galaxy and the BH without producing a galaxy-scale outflow. Although the **ISOTROPIC THERMAL** model is also at play in the simulation with the **FIDUCIAL TNG** physics, some gas escapes the galaxy perpendicular to its disc, following the path of least resistance, which creates a thin bipolar wind. In this simulation, we find that the **RANDOM KINETIC** model is never activated, as the Eddington ratio is always between 2 and 10 times higher than the thermal-to-kinetic mode threshold χ (equation 6). Therefore, the difference in gas distribution between the **FIDUCIAL TNG** and the **ISOTROPIC THERMAL** simulations is the consequence of the TNG thermal feedback being stronger in the former, for which the fraction of BH accreted rest-mass energy coupling to the gas ϵ_f is 20 times higher. With the **RANDOM KINETIC** model, a large-scale one-sided wind is launched in a random direction. Conversely, **MISTRAL-CONTINUOUS** generates a bipolar wind whose shock front, composed of dense cold gas, extends up to 10 kpc before flowing back to the ISM at later times. On the other hand, the episodic BH energy injection with **MISTRAL-STOCHASTIC** efficiently carves low-density hot bipolar outflows into the CGM. In these idealized simulations, the angular momentum of gas in the BH vicinity aligns with the z -axis. For this reason, and by construction, gas cells with **MISTRAL-STOCHASTIC** are always kicked vertically. As a result, this creates a collimated region depleted of gas after successive AGN feedback episodes. The biconical winds produced with the two versions of MISTRAL are also visible from the velocity field overlaid in grey arrows in the gas density maps.

The gas velocity is further quantified in Fig. 4, which shows the mass-weighted probability density function of radial velocities stacked between $t = 500$ and 600 Myr (which encloses the time for which the maps in Fig. 3 are displayed). We plot gas within half the virial radius of the galaxy (i.e. within 115 kpc from the galaxy centre) and that is outside of the galaxy disc (as defined in Section 3.1), in order to exclude turbulent gas from the ISM and better focus on inflowing and outflowing gas in the CGM. The leftmost part of Fig. 4 shows similar velocities of $v_r \simeq -200 \text{ km s}^{-1}$ for the inflowing gas (i.e. negative radial velocities) with all models, with the exception of the **RANDOM KINETIC** model, which produces inflows as fast as $v_r \simeq -300 \text{ km s}^{-1}$. However, some models clearly lead to different positive radial velocities. Outside of the galaxy disc in the **NO BH**, **FIDUCIAL TNG**, **ISOTROPIC THERMAL**, and **MISTRAL-CONTINUOUS** simulations, gas reaches a maximum radial velocity of $200\text{--}300 \text{ km s}^{-1}$. On the other hand, the **RANDOM KINETIC** model and **MISTRAL-STOCHASTIC** are more efficient at generating winds as fast as 1000 km s^{-1} , and **MISTRAL-STOCHASTIC** produces more gas at $v_r > 300 \text{ km s}^{-1}$ than the **RANDOM KINETIC** model. Therefore, this shows that injecting kinetic energy stochastically, rather than in a continuous manner or in thermal form, is more efficient at producing fast AGN-driven winds out of the ISM with our set-up. This result is similar to what is found for instance by Choi et al. (2012), using the mechanical feedback model that inspired **MISTRAL-STOCHASTIC**.

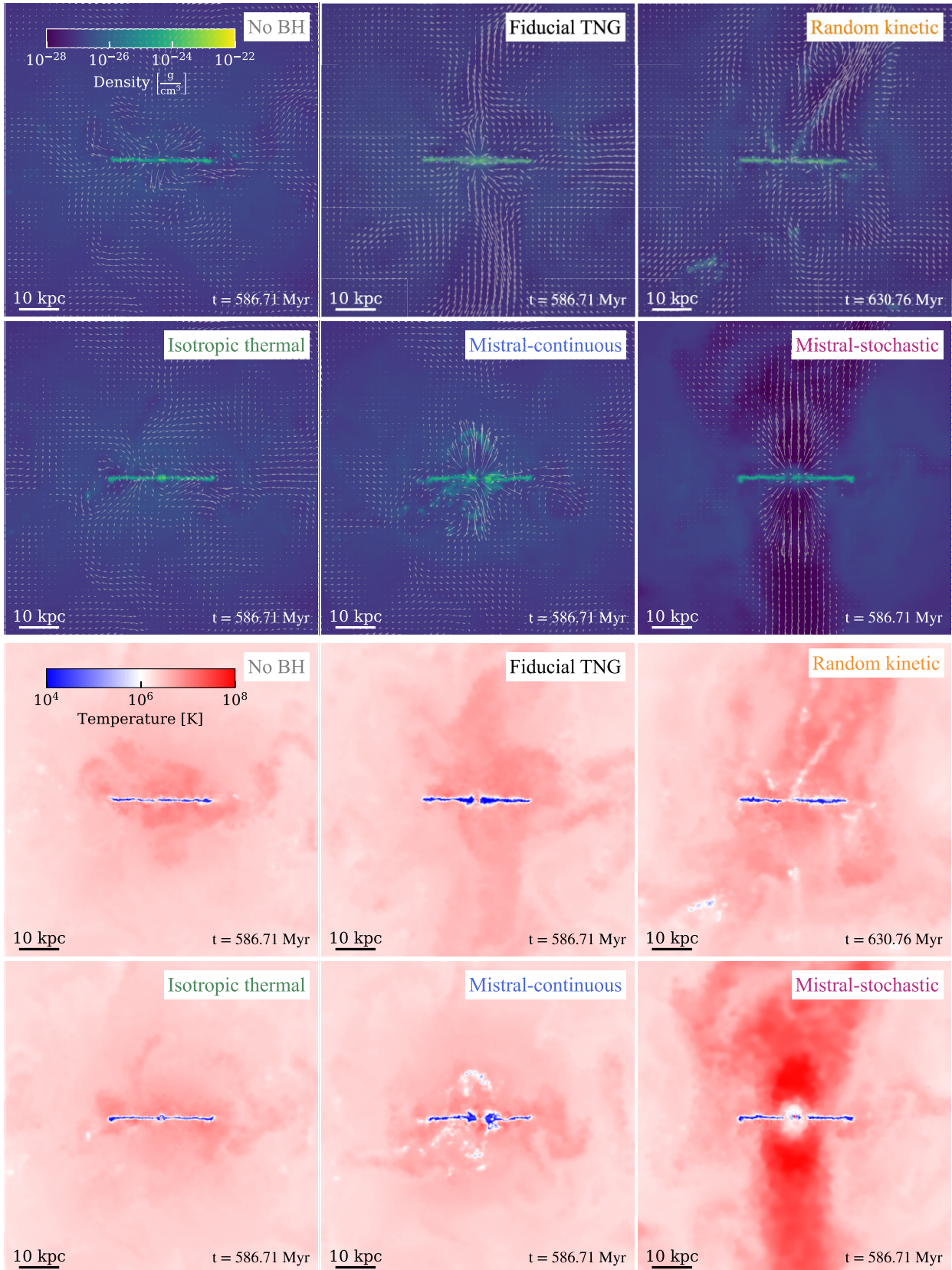


Figure 3. Edge-on 80 kpc wide slices of the gas density (top rows) and temperature (bottom rows). From left to right and top to bottom, we show the NO BH, FIDUCIAL TNG, RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC simulations. A 10 kpc width scale bar is plotted in the lower left corner, and the time at which these maps are plotted is indicated in the lower right corner. In the density maps, the arrows show the gas velocity field overlaid.

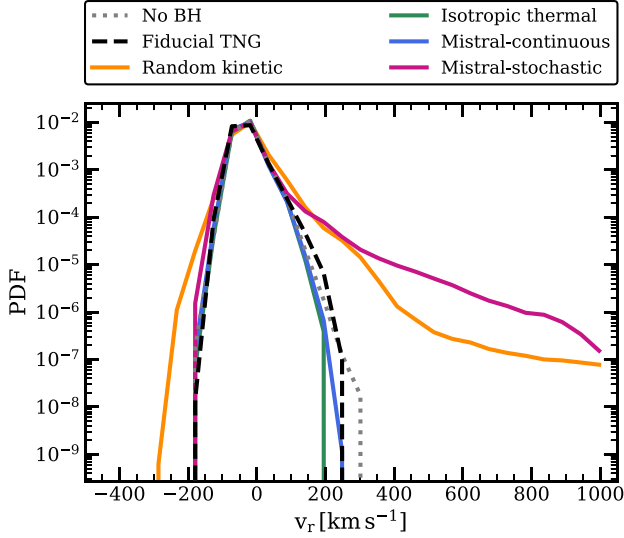


Figure 4. Probability density function of the gas radial velocity with data stacked from 500 to 600 Myr. We restrict the material used in the analysis to gas within half the galaxy virial radius and outside the galaxy disc. We show simulations without BH physics (NO BH) in grey dotted, and with the FIDUCIAL TNG, RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC feedback models in dashed black, solid orange, green, blue, and purple, respectively.

Fig. 5 quantitatively compares the effect of the different AGN feedback models on gas properties. The figure shows phase diagrams of temperature as a function of radial velocity for gas within half the virial radius of the galaxy. We stack data between 500 and 600 Myr to smooth any transient effects, which may arise from the burstiness of star formation and from the stochasticity of feedback processes. We also highlight the 1σ contours (black lines) where the gas hydrogen density is $n_H = 10^{-2}$, 10^{-1} , and 1 cm^{-3} , to guide the comparison between the simulations. These contours globally reveal two regimes: dense relatively cold gas with $T \leq 10^5 \text{ K}$, which corresponds to ISM gas, and diffuse gas at $T > 10^5 \text{ K}$, which is predominantly located in the CGM. Because of the two-phase ISM model adopted in these simulations, the temperature of the dense star-forming gas is limited to 10^4 K , and can reach higher temperatures as a consequence of overpressurization (Springel & Hernquist 2003). Overall the NO BH, ISOTROPIC THERMAL, and FIDUCIAL TNG simulations have similar phase diagrams. While the six galaxies all have dense gas at $T \simeq 10^4 \text{ K}$, only the simulations with the RANDOM KINETIC model and MISTRAL-STOCHASTIC lack a second reservoir of dense gas at $n_H = 1 \text{ cm}^{-3}$ and $T \simeq 10^5 \text{ K}$. This gas phase, present in the four other simulations, corresponds to heated and ejected gas, which has formed a small-scale galactic fountain that quickly cycles back to the ISM. With the RANDOM KINETIC model and MISTRAL-STOCHASTIC, this gas is sufficiently accelerated that it is ejected out of the ISM. With MISTRAL-STOCHASTIC in particular, this produces an extended tail of fast, hot, and diffuse gas in the upper right corner of the plot. The phase diagrams also show that MISTRAL accelerates gas to higher velocities than the ISOTROPIC THERMAL model. The higher velocity gas is cold and dense with MISTRAL-CONTINUOUS, and mostly belongs to the galaxy disc (as it is not visible in Fig. 4), while it is predominantly hot and diffuse with MISTRAL-STOCHASTIC, supporting the qualitative results from Fig. 3.

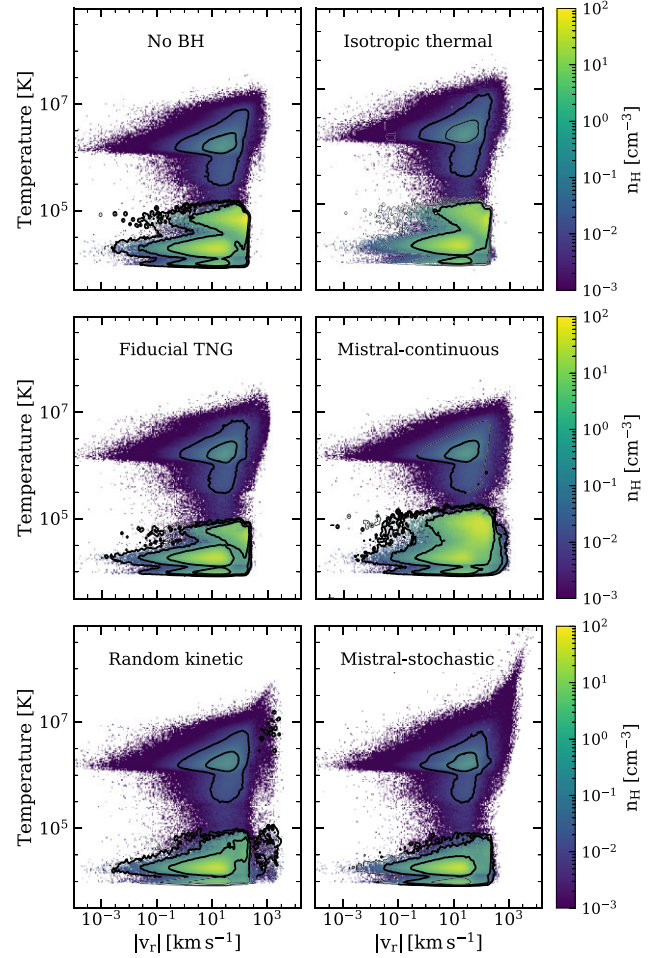


Figure 5. Phase diagrams of gas temperature against the gas radial velocity, colour-coded by the mean hydrogen number density. From top to bottom and left to right, we show the NO BH, ISOTROPIC THERMAL FIDUCIAL TNG MISTRAL-CONTINUOUS, RANDOM KINETIC, and MISTRAL-STOCHASTIC simulations, with data stacked from 500 to 600 Myr (21 snapshots). The black lines show the 1σ contours for $n_H = 10^{-2}$, 10^{-1} , and 1 cm^{-3} .

3.2.2 Regulation of star formation and BH growth

We now quantify the impact of the different AGN feedback models on regulating star formation and black hole growth. From top to bottom, Fig. 6 shows the star formation rate, cumulative stellar mass formed, and gas fraction in the galaxy disc as a function of time. Here, we define gas fraction as $f_{\text{gas,disk}} = M_{\text{gas,disk}} / (M_{\text{gas,disk}} + M_{\text{stellar,disk}})$, where $M_{\text{stellar,disk}}$ includes both newly formed stars and the stellar mass formed during the pre-BH phase of the simulations.

Focusing on stellar properties first, we find that MISTRAL-STOCHASTIC suppresses star formation more significantly than other models, while MISTRAL-CONTINUOUS, in contrast, globally leads to the highest SFR after the first 200 Myr. This can be directly linked to the gaseous content in the disc, which has a time evolution similar to that of the SFR. Despite having the highest stellar mass, the simulation with MISTRAL-CONTINUOUS also has the highest gas fraction of the six runs. With this feedback model, the variation of $f_{\text{gas,disk}}$ with time reflects the fact that the dense outflowing gas accelerated by AGN feedback rapidly falls back to the galaxy, episodically providing more fuel for star formation. By construction, MISTRAL-CONTINUOUS injects momentum radially away, which

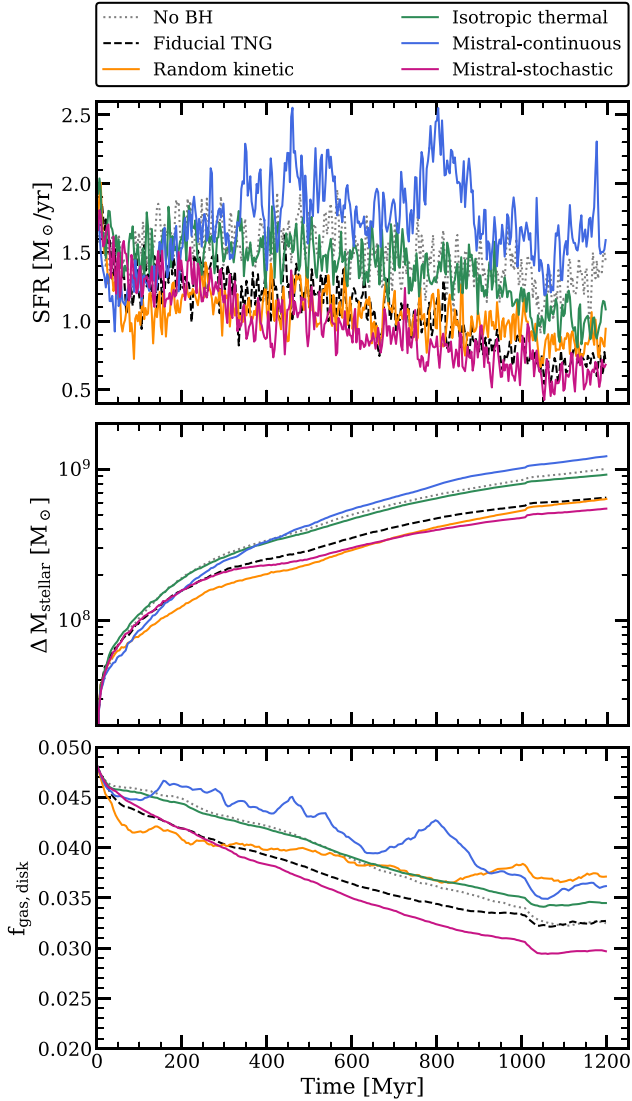


Figure 6. From top to bottom: SFR, cumulative stellar mass formed, and gas fraction in the galaxy disc, as a function of time. We show simulations without BH physics in dotted grey, with the FIDUCIAL TNG model in dashed black, and with RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC shown by orange, green, blue, and purple solid lines, respectively.

creates density contrasts in the ISM and leads to gas compression that triggers star formation compared to the other models. In particular, the galaxy with MISTRAL-CONTINUOUS has higher SFRs and gas fractions than the simulation without BH physics. We note that our results would likely differ if using different star formation and resolved multiphase ISM prescriptions (which we will investigate in a forthcoming study). With this set-up, MISTRAL-CONTINUOUS acts as positive feedback, leading to a stellar mass formed 1.2 times higher than in the absence of AGN feedback. Given our parametrization, the simulation with the ISOTROPIC THERMAL model produces very similar SFR and gas fraction as the NO BH simulation. With its isotropic release of thermal energy that gently heats up the ISM gas, the ISOTROPIC THERMAL simulation produces lower disc gas fractions than MISTRAL-CONTINUOUS, but forms almost as many stars as the NO BH simulation. These quantities are reduced in the

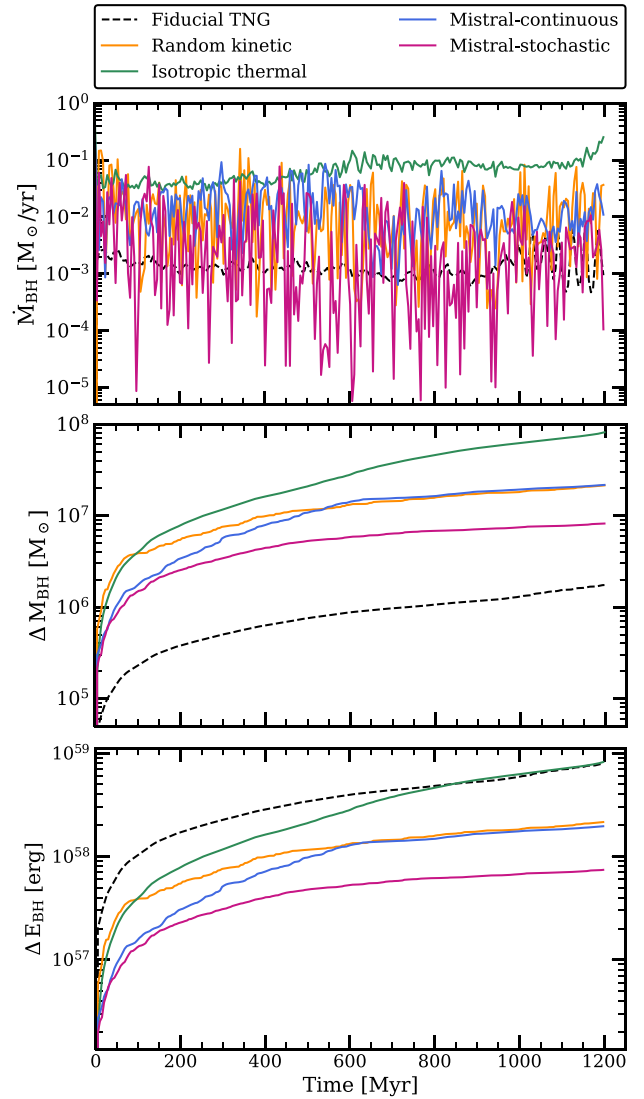


Figure 7. From top to bottom: BH accretion rate, cumulative BH mass formed, and cumulative BH energy released, as a function of time. We show simulations with the FIDUCIAL TNG model in dashed black, and with RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC shown by orange, green, blue, and purple solid lines, respectively.

FIDUCIAL TNG simulation, as a result of the higher BH energy released (see also Fig. 7), reducing the stellar mass formed by a factor of 1.6 compared to the simulation without BH physics. Because a similar star formation suppression is achieved with the RANDOM KINETIC model with less BH energy released, this shows that a larger amount of thermal energy is needed to achieve the same effect on star formation as kinetic winds (but the impact on gas properties and BH growth differs with the BH energy deposition scheme). Finally, being more efficient at ejecting gas out of the disc, MISTRAL-STOCHASTIC has even lower $f_{\text{gas,disk}}$ and SFR, forming 1.8 times less stellar mass than the counterpart NO BH run. We note however that these differences in stellar mass formed remain negligible compared to the total stellar mass of the galaxy, which varies by no more than 2 per cent during the 1200 Myr of the simulations.

AGN feedback does not only impact star formation but also BH accretion. Fig. 7 shows the black hole accretion rate, cumulative black hole mass accreted, and cumulative BH energy released. Overall, the RANDOM KINETIC and the two MISTRAL models produce a more bursty BH accretion history than the ISOTROPIC THERMAL and FIDUCIAL TNG models. Given our choice of efficiency parameters (see Section 3.1), we remind the reader that the fraction of inflowing gas mass accreted by the BH corresponds to 90 per cent in the ISOTROPIC THERMAL and RANDOM KINETIC simulations, to 50 per cent with MISTRAL and, by default, to 80 per cent with the FIDUCIAL TNG model. This explains why the BH accretion rates are consistently higher with the ISOTROPIC THERMAL model than in the other simulations, despite releasing more BH energy than MISTRAL, for instance. Unlike with the deposition of kinetic energy, the thermal energy released by the BH with the ISOTROPIC THERMAL model is more prone to being radiated away in dense gas, which reduces the actual feedback efficiency (Weinberger et al. 2018). Therefore, with the low feedback coupling efficiency adopted in our simulation, the ISOTROPIC THERMAL model is inefficient at regulating both star formation and BH growth. When this feedback model is stronger, with a higher feedback coupling efficiency in the FIDUCIAL TNG simulation, the higher amount of BH energy initially released efficiently regulates BH growth (Terrazas et al. 2020), and the total BH mass accreted is almost 50 times lower. Just as it is more efficient at regulating star formation, the RANDOM KINETIC model leads to lower BH accretion rates and releases less energy than the ISOTROPIC THERMAL model, despite having the same feedback coupling efficiency.

Similarly as for the SFR, the simulation with MISTRAL-CONTINUOUS has higher BH accretion rates than the one with MISTRAL-STOCHASTIC, as the former retains more gas near the BH. This is also because MISTRAL-STOCHASTIC is more efficient at expelling and heating the gas surrounding the BH, leading to lower BH accretion rates. Compared to the simulation with the ISOTROPIC THERMAL model, the integrated BH mass accreted differs by a factor of 3.8 and 10 for MISTRAL-CONTINUOUS and MISTRAL-STOCHASTIC, respectively, such that the BH eventually reaches a mass 1.8 lower with MISTRAL-CONTINUOUS, and 2.3 times lower with MISTRAL-STOCHASTIC. By the end of the simulation, the BH reaches a mass of 5.1×10^7 , 7.1×10^7 , 1.3×10^8 , 7.1×10^7 , and $5.8 \times 10^7 M_\odot$ with the FIDUCIAL TNG, RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC models, respectively. Interestingly, despite releasing less BH energy over the duration of the simulation, MISTRAL-STOCHASTIC efficiently regulates star formation, resulting in the lowest overall SFR among all the models we investigated.

3.2.3 Properties of the outflowing gas

The main goal of MISTRAL is to capture the consequences of the winds that are accelerated in the vicinity of the SMBH, as a result of the energy released from the accretion disc. To quantify the efficiency of AGN feedback in generating galactic winds, this section now focuses on the properties of the outflowing gas. We define gas as outflowing when it is flowing away from the galaxy centre, in a radial direction (i.e. gas cells with a positive radial velocity). To compute the mass and kinetic energy outflow rates \dot{M}_{outf} and \dot{E}_{outf} , we define spherical shells of width $\Delta R_{\text{shell}} = 2$ kpc, located at a distance R_{shell} between 4 and 100 kpc from the galaxy centre. \dot{M}_{outf} and \dot{E}_{outf} are computed as follows, by summing the contribution of all the outflowing cells i , of gas mass m_i , and radial velocity $v_{r,i} > 0$, which are intersected by the shells:

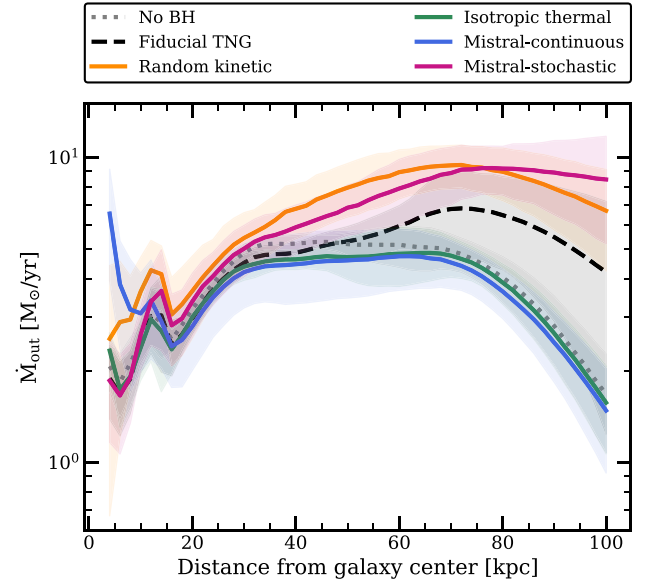


Figure 8. Mass outflow rate as a function of distance from the galaxy centre, with data stacked from 200 to 1000 Myr. The coloured lines represent average quantities, and the limits of the standard deviation from the mean are shown by the shaded area. The dotted grey, dashed black, solid orange, green, blue, and purple lines show simulations with NO BH, the FIDUCIAL TNG, RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC feedback models.

$$\dot{M}_{\text{outf}} = \sum_i \frac{m_i v_{r,i}}{\Delta R_{\text{shell}}} \quad (26)$$

$$\dot{E}_{\text{outf}} = \sum_i \frac{1}{2} \frac{m_i v_{r,i}}{\Delta R_{\text{shell}}} v_{r,i}^2 \quad (27)$$

Fig. 8 first focuses on the mass outflow rate, shown as a function of distance from the galaxy centre for our six simulations. In order to reduce the noise due to the burstiness of supernova and AGN feedback, we stack 161 outputs between 200 and 1000 Myr. We show the average mass outflow rate (coloured lines) together with the standard deviation (shaded areas). Because the outflow properties are calculated in spherical shells, distances below 20 kpc include gas that resides in the ISM. There, MISTRAL-CONTINUOUS and the RANDOM KINETIC model produce the highest mass outflow rates, by accelerating large quantities of dense gas. This is also visible in the bottom right part of the corresponding phase diagram (Fig. 5), where we see more cold high-velocity gas for these two models. At larger distances and up to 50 kpc, MISTRAL-CONTINUOUS, the ISOTROPIC THERMAL, and the FIDUCIAL TNG model all produce mass outflow rates similar to the case where no BH physics is included. The outflow rates with the ISOTROPIC THERMAL model have a smaller scatter than with MISTRAL-CONTINUOUS, due to BH accretion being less bursty. Interestingly, the mass outflow rates with the FIDUCIAL TNG model start to diverge at 50 kpc and beyond, where enough BH energy remains to support outflowing gas (see also Fig. 9), and where the gas density is sufficiently low to avoid significant radiative losses. Beyond 20 kpc, kinetic winds with MISTRAL-STOCHASTIC and the RANDOM KINETIC produce the highest outflow rates, reaching values close to $10 M_\odot \text{ yr}^{-1}$ 100 kpc away from the galaxy, compared to $1\text{--}2 M_\odot \text{ yr}^{-1}$ for MISTRAL-CONTINUOUS and the ISOTROPIC THERMAL model. While the mass outflow rates start to drop when $R_{\text{shell}} \geq 70\text{--}80$ kpc with all the other models, this is not the case with

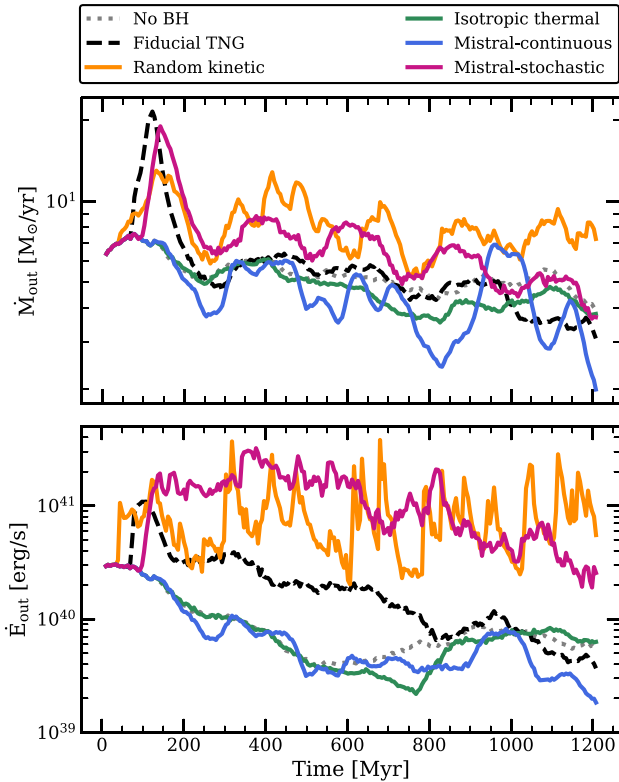


Figure 9. Mass outflow rate and kinetic energy outflow rate as a function of time, measured in a shell located at 50 kpc from the galaxy centre. The dotted grey, dashed black, solid orange, green, blue, and purple lines show simulations with NO BH, the FIDUCIAL TNG, RANDOM KINETIC, ISOTROPIC THERMAL, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC feedback models.

MISTRAL-STOCHASTIC. This is a consequence of the episodic acceleration of gas along the same direction, maintaining the highest gas outflow rates at radii between 80 and 100 kpc from the galaxy centre. At $R_{\text{shell}} = 100$ kpc, MISTRAL-STOCHASTIC therefore leads to mass outflow rates almost one order of magnitude higher than the ISOTROPIC THERMAL model and MISTRAL-CONTINUOUS.

To better quantify the efficiency of different AGN feedback models in driving large-scale winds, Fig. 9 shows the mass outflow rate and the kinetic energy outflow rate as a function of time.⁸ Both quantities are measured at $R_{\text{shell}} = 50$ kpc, a distance that probes the CGM, and at which the mass outflow rates with MISTRAL-CONTINUOUS and the ISOTROPIC THERMAL model reach a plateau (as visible in Fig. 8).

Overall, the mass outflow rates at that distance are the highest with MISTRAL-STOCHASTIC and the RANDOM KINETIC model. Conversely, MISTRAL-CONTINUOUS and the ISOTROPIC THERMAL model produce the lowest mass and kinetic energy outflow rates, despite the total amount of BH energy released being up to five times lower with MISTRAL-CONTINUOUS. Compared to the simulation with the ISOTROPIC THERMAL model, the one with the FIDUCIAL TNG physics globally has higher mass outflow rates. Because of the large amount

⁸We refrain from showing energy outflow rates normalized by the instantaneous AGN luminosity given the high variability of the BH accretion rates with MISTRAL. This also allows better visualization of the BH energy injection events, which happen occasionally with MISTRAL-STOCHASTIC and the RANDOM KINETIC model, unlike their continuous BH accretion.

of BH energy released during the first 200 Myr, we also note the presence of an initial peak in mass and energy outflow rates in this simulation. We report a similar feature with MISTRAL-STOCHASTIC and the RANDOM KINETIC model after the first AGN feedback event, despite more than 10 times less BH energy being released with MISTRAL-STOCHASTIC.

Unlike the ISOTROPIC THERMAL and FIDUCIAL TNG models, the two versions of MISTRAL both lead to outflowing properties that fluctuate episodically. This can be interpreted as follows. Depending on BH accretion, MISTRAL-CONTINUOUS continuously injects momentum radially away, and this occasionally launches small-scale winds, as visible in Fig. 3. The time needed for these galactic fountains to fall back to the galaxy leads to ~ 100 – 200 Myr fluctuations in outflow properties, as gas has to reach the ISM and be accreted by the SMBH before AGN feedback has a chance to expel gas again. Fluctuations with a similar time delay can also be observed in galaxy properties, such as SFR, BH accretion rate, and gas fraction (Figs 6 and 7). This illustrates that there is a cycle of gas, which affects BH accretion, and consequently the injection of AGN feedback. With the RANDOM KINETIC model and MISTRAL-STOCHASTIC, the episodic launching of winds directly correlates with the stochastic injection of kinetic AGN feedback, by construction. The first AGN feedback episodes are sufficiently efficient that they remove gas from the ISM, which can also be seen in Fig. 6 and the subsequent reduction of star formation and BH growth. The following AGN feedback events then contribute to maintaining this trend. Despite the total BH energy released being the lowest with MISTRAL-STOCHASTIC, this model leads to the highest mass and energy outflow rates (except for the last 300 Myr dominated by the RANDOM KINETIC model at this distance from the galaxy), which is the signature of the powerful AGN-driven winds generated by this stochastic kinetic AGN feedback model.

To conclude, in these idealized simulations, the RANDOM KINETIC model and MISTRAL-STOCHASTIC happen to produce similar mass and energy outflow rates, and lead to a similar star formation suppression. However, the AGN winds they produce have different morphologies, which differently impacts BH growth and gas properties, both in the galaxy disc and in the CGM. Because idealized simulations exclude the complexity of the CGM and cosmic inflows, the differences we see between AGN feedback models in these idealized galaxies are amplified in a more realistic cosmic environment, which we examine in the next section.

4 ZOOM SIMULATIONS OF MASSIVE GALAXIES AT COSMIC NOON

In this section we show how the two versions of our model affect the evolution of massive galaxies in a more realistic cosmological context. To take into account the impact of galaxy mergers and of gas inflows from the cosmic web, we turn to cosmological simulations. For this purpose, we conduct a suite of cosmological zoom-in simulations of massive $z = 2$ galaxies from the TNG100 simulation (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018b; Springel et al. 2018), whose galaxy formation model has been already detailed in Section 2.1. This simulation corresponds to the intermediate cosmological volume of the IllustrisTNG project, with a box size of 110 cMpc in width. TNG100 has a dark matter mass resolution of $7.5 \times 10^6 M_{\odot}$, a baryonic mass resolution of $1.6 \times 10^6 M_{\odot}$, an adaptive gravitational softening length of 185 comoving pc at minimum for gas, and a gravitational softening length of 740 pc at $z = 0$ for stars and dark matter. It adopts a Planck Collaboration XIII (2016) cosmology with $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, $\Omega_{b,0} = 0.0486$, $\sigma_8 = 0.8159$,

Table 2. Properties at $z = 2$ of the 15 galaxies targeted for re-simulation, as obtained from the snapshot 33 ($z = 2$) of the TNG100 simulation. From left to right: GrNr: halo group number of the galaxy, M_{halo} : total mass of the halo hosting the galaxy, M_{stellar} : stellar mass within twice the stellar half-mass radius, sSFR: specific star formation within twice the stellar half-mass radius, M_{BH} : BH mass. Starred halo numbers correspond to objects that are evolved down to $z = 0$.

GrNr at $z = 2$	M_{halo} [$10^{12} M_{\odot}$]	M_{stellar} [$10^{10} M_{\odot}$]	sSFR [yr^{-1}]	M_{BH} [$10^8 M_{\odot}$]
631*	1.1	0.6	1.6×10^{-9}	0.03
511	2.3	4.8	3.5×10^{-10}	0.9
449	1.6	4.2	3.5×10^{-12}	4.1
306	2.7	5.5	5×10^{-10}	2.5
200	3.2	5.5	2×10^{-11}	5.1
190	4.3	12.7	0	5.7
123*	3.5	2.3	1.2×10^{-9}	1.8
113	6.6	7.96	1.3×10^{-13}	8.9
61	8.8	12.1	3.3×10^{-14}	6.2
57*	8.4	11.2	10^{-10}	6.3
55	6.1	12.8	9.4×10^{-15}	7.5
30	10.0	15.8	3×10^{-11}	11.4
18	17.6	32.9	4.7×10^{-11}	15
10	26.3	20.2	1.5×10^{-12}	8.7
6	31.5	30.1	1.7×10^{-12}	18.2

$n_s = 0.9667$, and $h = 0.6774$. Unless stated otherwise, we use the same galaxy formation models and parameters as in the TNG100 simulation. Below, we expand on the galaxies targeted for our sample of zoom simulations.

4.1 Initial conditions and sample of zoom simulations

We randomly select 12 galaxies among the most massive from TNG100 at $z = 2$ (three of them are also simulated at higher resolution, see Appendix A). We additionally target three massive galaxies from TNG100 at $z = 0$, which are used to calibrate ϵ_w , in order to produce realistic $z = 0$ (and $z = 2$) galaxy and BH properties (see Appendix B). In the TNG simulations (and in our zoom simulations), haloes and galaxies are, respectively, defined as friend-of-friend groups and gravitationally bound objects, and are determined on-the-fly using the SUBFIND algorithm (Springel et al. 2001). Each galaxy can then be tracked back over time via its merger tree, built in post-processing using the SUBLINK algorithm (Nelson et al. 2015; Rodriguez-Gomez et al. 2015). As summarized in Table 2, at $z = 2$, our 15 target galaxies have stellar masses $M_{\text{stellar}} \geq 6 \times 10^9 M_{\odot}$ and are hosted in haloes with masses in the range $M_{\text{halo}} = 10^{12} - 3 \times 10^{13} M_{\odot}$.⁹ Throughout this section, if not mentioned otherwise, stellar properties are defined from all the gravitationally bound star particles within twice the stellar half-mass radius of the galaxies. We re-simulate this sample of haloes using the zoom-in technique described below, and will refer to these re-simulations as ‘zoom’ simulations even if we keep the same resolution as TNG100.

For each selected halo, we first extract the list of dark matter particles at snapshot 33 from TNG100, corresponding to our target

final redshift $z = 2$. This is with the exception of haloes with group number GrNr 631, 123, and 57, for which ICs are extracted at $z = 0$ (snapshot 99), their final redshift (see Appendix B). The spatial distribution of these particles is then located back in time, down to the initial snapshot zero (which corresponds to $z = 20$), in order to identify the Lagrangian region from which the halo formed. We finally apply to this particle set a perturbation field following the Zel’dovich (1970) approximation, in the same way as done in N-GENIC (Springel 2015), using the same cosmological parameters and power spectrum as used in the TNG100 simulation. The initial conditions generated consist of a list of dark matter particles (from which gas cells are internally generated by AREPO), at TNG100 resolution within the region centred on the target halo, and at a lower resolution outside, such that the computational load is exclusively focused on the region of interest. Within this zoom region, the highest resolution DM particles have a mass of $7.5 \times 10^6 M_{\odot}$, and the DM resolution progressively coarsens away from the region of interest down to $2 \times 10^{12} M_{\odot}$. The central ‘zoom’ region of interest is defined as the sphere containing four times more DM particles than that belonging to the virial radius R_{vir} of the target halo at the final redshift. This extends the resolution around the halo’s immediate surrounding (up to $\simeq 1.6 \times R_{\text{vir}}$) in order to minimize any contamination by low resolution DM particles.

4.2 Impact of AGN-driven winds on $z=2$ massive galaxies

In this section, we employ our suite of cosmological simulations to investigate the effects of AGN winds on massive galaxies. We focus on BH and galaxy properties at $z = 2$, which is the most promising redshift for probing the impact of AGN winds driven by radiatively efficient accretion. Indeed, this redshift, also referred to as ‘cosmic noon’, corresponds to the epoch when both star formation and SMBH growth are at their peaks (Madau & Dickinson 2014). Therefore, AGN feedback is expected to be particularly important, by generating winds that can impact star formation and gas fraction. We compare the two versions of MISTRAL to the TNG THERMAL model, in order to isolate the impact of AGN feedback in the radiatively efficient regime of SMBHs. Unlike the ISOTROPIC THERMAL model used in Section 3, the TNG THERMAL model includes the same fiducial physical models and parameters as in the TNG simulations, and we only disable the switch for the TNG RANDOM KINETIC model at low Eddington ratios. We also provide a comparison of MISTRAL to the FIDUCIAL TNG AGN feedback model where the kinetic winds are enabled at low Eddington ratios. In these two sets of simulations, we use the same radiative and coupling efficiency parameters as in TNG (unlike what was done in Section 3). When using the TNG THERMAL model, the fraction of rest-mass accreted energy that couples to the gas is defined by the product of $\epsilon_r = 0.2$ and $\epsilon_f = 0.1$. When switching to the RANDOM KINETIC model, this fraction is variable but capped at $\epsilon_{\text{kin}} = 0.2$ (see equation 9 from Weinberger et al. 2017). In the simulations with MISTRAL, we adopt $v_w = 10^4 \text{ km s}^{-1}$ and $\epsilon_w = 10^{-3}$, which corresponds to 36 per cent (respectively 64 per cent) of the inflowing gas mass being accreted (kicked as a wind). With both MISTRAL-CONTINUOUS and MISTRAL-STOCHASTIC, this choice of parameters gives reasonable results in terms of stellar and black hole masses compared to $z = 0$ observations (see Appendix B).

We first analyse some of the main galaxy properties compared with observationally derived scaling relations in Section 4.2.1. We then look at the impact of AGN feedback on gas fractions in Sections 4.2.2 and 4.2.3, and on gas flows in Section 4.2.4.

⁹We note that, upon re-simulation with the zoom-in technique, and even if using the same galaxy formation physics and resolution, some properties may vary from the parent galaxy to its zoom counterpart, and BH mass, stellar mass, and SFR will obviously differ when using one AGN feedback model or another, as shown by the different figures from this section.

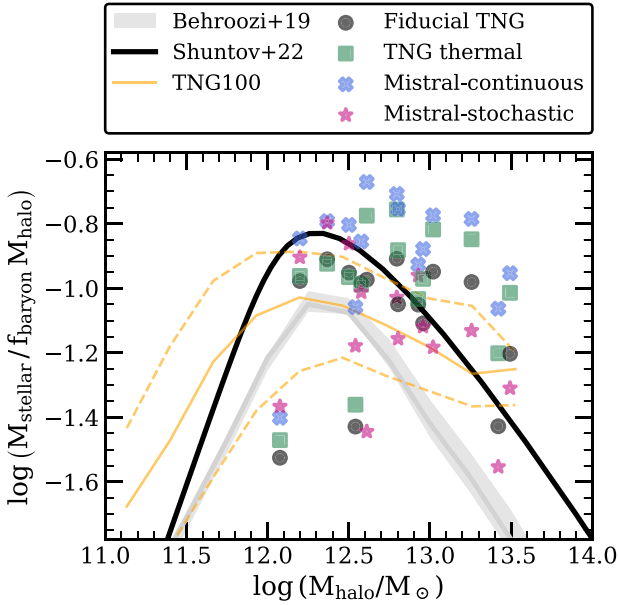


Figure 10. SMHM relation at $z = 2$, for zoom simulations with the FIDUCIAL TNG physics (circles), the TNG THERMAL model (squares), MISTRAL-CONTINUOUS (crosses), and MISTRAL-STOCHASTIC (stars). We also show the stellar to halo mass relation at $z = 2$ from Behroozi et al. (2019) with a grey line and shaded area, and from Shuntov et al. (2022) with a black line. To enhance clarity, the x-axis shows halo masses from zoom simulations with the FIDUCIAL TNG physics, such that, for each galaxy, all four models are displayed at the same x-coordinate. The solid and dashed orange lines correspond to the mean and 1σ scatter from the TNG100 simulation, shown for reference.

4.2.1 Stellar mass, SFR, BH mass, and BH accretion rate at $z=2$

We first focus on how AGN feedback regulates star formation. Fig. 10 shows the stellar mass to halo mass (SMHM) relation, where the y-axis is defined as the ratio between the stellar mass and the theoretical baryonic mass $f_{\text{baryon}} M_{\text{halo}} = 0.17 M_{\text{halo}}$. In addition, we compare our data with observational estimates at $z = 2$ from Behroozi et al. (2019) and Shuntov et al. (2022). As they do, in this section, we use the virial overdensity halo mass definition (Bryan & Norman 1998). With a few exceptions, all galaxies lie around or above the expected relations.¹⁰ The two lowest mass haloes (with GrNr 631 and 511) that are located before the peak of the SMHM relation, with $M_{\text{halo}} < 10^{12.3} M_{\odot}$, have higher stellar masses with MISTRAL than with the TNG AGN feedback models. This trend changes when halo mass increases. MISTRAL-CONTINUOUS (crosses) and the TNG THERMAL model (squares) lead to the highest stellar masses, especially at the most massive end. Enabling the RANDOM KINETIC feedback model in the FIDUCIAL TNG zooms (circles) reduces this tension, and MISTRAL-STOCHASTIC (stars) shows good agreement with the empirical relation from Shuntov et al. (2022), at any halo mass. From this analysis, we conclude that the TNG THERMAL mode alone and MISTRAL-CONTINUOUS are both inefficient at regulating star formation in the most massive galaxies, the regime where AGN feedback is thought to have a dominant role. This motivated the inclusion of a complementary more efficient AGN feedback

¹⁰Simulations with the lowest stellar masses correspond to the galaxies hosted in halo GrNr 190 with MISTRAL-STOCHASTIC and in halo GrNr 123 with the FIDUCIAL TNG and the TNG THERMAL models.

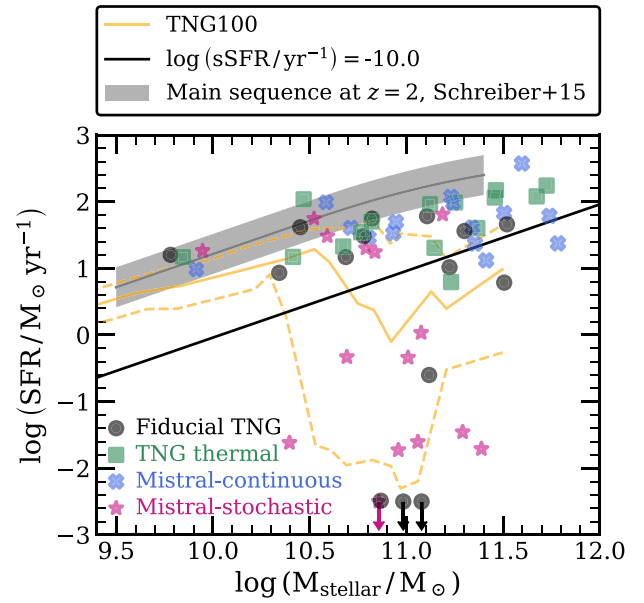


Figure 11. Star formation rate as a function of stellar mass at $z = 2$ for zoom simulations with the FIDUCIAL TNG physics (circles), the TNG THERMAL model (squares), MISTRAL-CONTINUOUS (crosses), and MISTRAL-STOCHASTIC (stars). SFR values that are lower than the limit of the plot (with FIDUCIAL TNG and MISTRAL-STOCHASTIC) are denoted with downwards arrows. The grey line and shaded region show the observationally derived $z = 2$ main sequence and its 1σ scatter, as taken from Schreiber et al. (2015), and the black line corresponds to the $z = 2$ sSFR quiescent threshold defined e.g. in Franx et al. (2008) as the Hubble time divided by 0.3. The solid and dashed orange lines correspond to the mean and 1σ scatter from the TNG100 simulation, shown for reference.

mode in the TNG simulations, which is achieved with the RANDOM KINETIC model when BHs have sufficiently low Eddington ratios given their mass (following equation 6, Weinberger et al. 2017). Here, our MISTRAL-STOCHASTIC model has a similar effect on galaxy growth, and works self-consistently across the whole halo mass range explored without the need to invoke a BH mass-dependent feedback scheme.

We report a similar result in Fig. 11, which shows the SFR as a function of stellar mass. We show the star formation main sequence at $z = 2$, as derived from Schreiber et al. (2015), which is reproduced by the TNG100 simulation. The black line corresponds to the $z = 2$ quiescent threshold of $\text{sSFR} = 10^{-10} \text{ yr}^{-1}$, defined as the Hubble time divided by 0.3, following Franx et al. (2008). At this redshift, almost all galaxies can be considered as star-forming when using the TNG THERMAL model alone, with only one galaxy having an SFR below the quiescence threshold. This number increases to two with MISTRAL-CONTINUOUS, which similarly only suppresses the SFR for some of the most massive galaxies. However, as shown in Fig. 10, this is not sufficient to produce realistic galaxy masses at $z = 2$. Quenching massive galaxies is easier when enabling the RANDOM KINETIC model at low Eddington ratios in the FIDUCIAL TNG runs, with six galaxies out of 15 that drop below the quiescence limit. MISTRAL-STOCHASTIC is even more efficient at suppressing star formation, with 9/15 galaxies quenched at $z = 2$. Notably, MISTRAL-STOCHASTIC efficiently acts over a wider range of galaxy mass compared to the FIDUCIAL TNG model, extending quenching to galaxies with lower stellar masses (down to $M_{\text{star}} = 10^{10.4} M_{\odot}$) than any of the other AGN feedback models. While MISTRAL-STOCHASTIC yields a

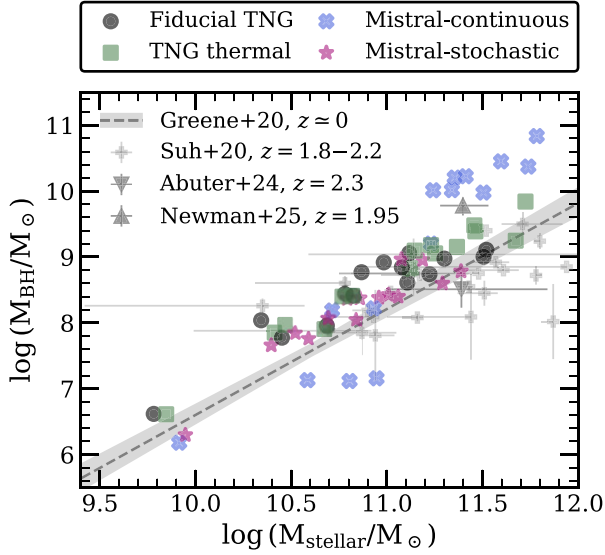


Figure 12. BH mass versus stellar mass at $z = 2$, for zoom simulations with FIDUCIAL TNG (circles), TNG THERMAL (squares), MISTRAL-CONTINUOUS (crosses), and MISTRAL-STOCHASTIC (stars) AGN feedback models. We show the local $z = 0$ black hole–stellar mass relation from Greene, Strader & Ho (2020), and $z \simeq 2$ observational estimates from Suh et al. (2020), Abuter et al. (2024), and Newman et al. (2025).

quenched fraction of 60 per cent at $z = 2$ in our sample, this value lies within the broad observational range of 20–70 per cent reported for similar stellar masses (e.g. Muzzin et al. 2013; Pandya et al. 2017; Weaver et al. 2023). In forthcoming work, we will explore quenching fractions further with a larger simulation sample.

AGN feedback is expected to influence not only the growth of its host galaxy but also the evolution of its SMBH. Fig. 12 shows the BH mass versus stellar mass at $z = 2$, compared to two direct dynamical measurements of BH mass at $z \simeq 2$ from Abuter et al. (2024) and Newman et al. (2025). We also include observational estimates at $z = 1.8$ – 2.2 from Suh et al. (2020), though we caution that these are based on indirect methods (such as virial estimators applied to broad-line AGNs), which are subject to significant systematic uncertainties (e.g. Bertemes et al. 2025), such that their use for comparison should be treated with care. Finally, we include the $z = 0$ BH–stellar mass observational relation from Greene et al. (2020), supported by observational evidence suggesting that this relation has not evolved strongly, if at all, since $z = 2$ (e.g. Suh et al. 2020; Tanaka et al. 2025).

Acknowledging these uncertainties, we find that with MISTRAL-CONTINUOUS, BHs are, respectively, undermassive and overmassive in low- and high-mass galaxies, compared to these observations. However, we note that this result depends on the value of ϵ_w , chosen as the best compromise for producing realistic stellar and black hole masses at $z = 0$ – 2 with both MISTRAL-CONTINUOUS and MISTRAL-STOCHASTIC (Fig. B1 and B2). In Appendix B, we show with three galaxies (GrNr 631, 123, and 57 at $z = 2$) that a higher value of ϵ_w tends to produce too low BH masses with both versions of MISTRAL at $z = 2$, and with MISTRAL-STOCHASTIC at $z = 0$. Conversely, values giving a better match with the observations at $z = 2$ also produce overmassive BHs at $z = 0$. With the value of ϵ_w adopted, MISTRAL-CONTINUOUS therefore prevents BH growth in the lowest mass galaxies from our sample, and fails at regulating BH and galaxy growth in the most massive ones.

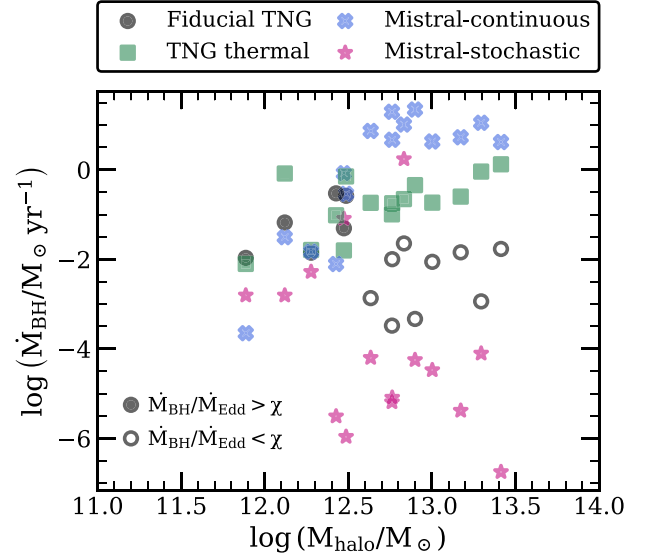


Figure 13. Black hole accretion rate as a function of halo mass at $z = 2$, for zoom simulations with the FIDUCIAL TNG physics (circles), the TNG THERMAL model (squares), MISTRAL-CONTINUOUS (crosses), and MISTRAL-STOCHASTIC (stars). With the former, empty markers indicate that the TNG RANDOM KINETIC model is active, as determined by $\dot{M}_{\text{BH}}/\dot{M}_{\text{Edd}} < \chi$ (see equation 6). As for Fig. 10, the x-axis shows halo masses from zoom simulations with the FIDUCIAL TNG physics, in order to display our four models at the same x-coordinate.

Interestingly, using the TNG THERMAL feedback alone leads to both higher stellar and black hole masses than with the FIDUCIAL TNG physics, in such a way that the BH–galaxy co-evolution remains broadly consistent with the observational estimates. BH masses with MISTRAL-STOCHASTIC are also nicely consistent with these observations, and globally leads to slightly lower BH masses than the FIDUCIAL TNG zoom simulations.

To further evaluate the impact of AGN feedback on BH growth, Fig. 13 shows the BH accretion rate at $z = 2$, as a function of halo mass. For $M_{\text{halo}} < 10^{12.5} M_{\odot}$, the four models lead to similar BH accretion rates, with MISTRAL producing slightly lower values. This indicates that, in this regime, both variants of our model regulate BH growth at $z = 2$ at least as efficiently as, if not more efficiently than, the TNG AGN feedback prescriptions. In more massive haloes, a clear trend emerges: MISTRAL-CONTINUOUS produces the highest BH accretion rates, followed by the TNG THERMAL model, with the FIDUCIAL TNG model leading to even lower values. Notably, at $z = 2$, all SMBHs in haloes with $M_{\text{halo}} > 10^{12.5} M_{\odot}$ have $\dot{M}_{\text{BH}}/\dot{M}_{\text{Edd}} < \chi$ (where χ is defined by equation 6), indicating that AGN feedback operates following the RANDOM KINETIC model. Finally, in these most massive haloes, MISTRAL-STOCHASTIC generally results in the lowest BH accretion rates at this redshift, which is also true for the entire halo mass range explored (with the exception of a few objects).

To summarize, Figs 10, 11, 12, and 13 show that the TNG THERMAL feedback alone and MISTRAL-CONTINUOUS are generally less efficient at regulating black hole and stellar masses in massive galaxies. MISTRAL-STOCHASTIC better matches observational estimates at all halo masses, and is sufficiently strong to quench more than half of the galaxies in our sample by $z = 2$. We note, however, that this study focuses on massive systems, which would be the progenitors of massive quiescent galaxies at $z = 0$. In future work, we will explore how MISTRAL would regulate such galaxies as well as lower

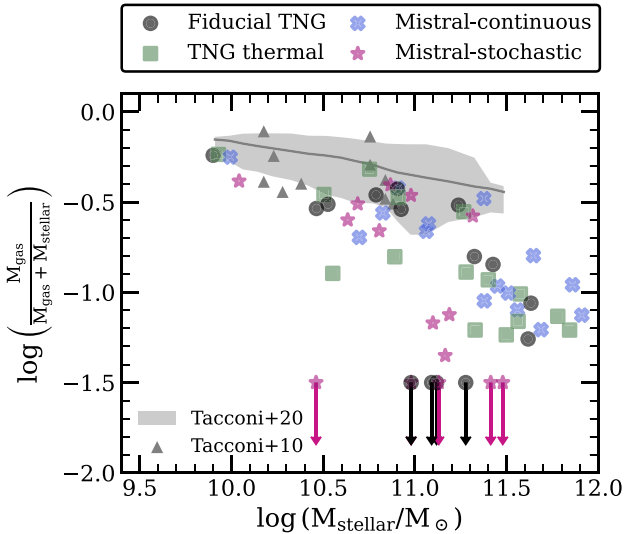


Figure 14. Cold gas fraction versus stellar mass at $z = 2$, for simulations with the FIDUCIAL TNG physics (circles), the TNG THERMAL model (squares), MISTRAL-CONTINUOUS (crosses), and MISTRAL-STOCHASTIC (stars). Both the gas and the stellar masses are measured within a sphere of radius $R_{\text{vir}}/10$ centred on the galaxies. Gas is defined as cold if it has a temperature $T \leq 10^5$ K and a density $\rho \geq 2 \times 10^{-25} \text{ g cm}^{-3}$. The arrows denote gas fractions below the range of the plot. The triangle markers and shaded area show observations of star-forming galaxies at $z = 2$ compiled by Tacconi et al. (2010) and Tacconi, Genzel & Sternberg (2020).

mass objects at lower redshift, especially when the SMBH is in the radiatively inefficient regime.

4.2.2 Cold gas fraction

We now want to study how AGN feedback impacts the gas content in and around galaxies. We start by illustrating how the cold gas fraction (defined as the cold gas to baryonic mass ratio) varies as a function of stellar mass in Fig. 14. We measure the cold gas and stellar masses within the galaxy radius, defined as 10 per cent of the virial radius (e.g. Oser et al. 2010). Gas is defined as cold if it has a temperature $T \leq 10^5$ K and a density $\rho \geq 2 \times 10^{-25} \text{ g cm}^{-3}$ (or equivalently $n_{\text{H}} \geq 0.13 \text{ cm}^{-3}$), which corresponds to the thresholds used to define star-forming gas in the TNG simulations. For reference, we include observations of star-forming galaxies at $z = 2$ from Tacconi et al. (2010) and Tacconi et al. (2020). In observational data, cold gas is usually traced by molecular gas. This is not modelled in our simulations, which rely on a two-phase ISM subgrid prescription instead (Springel & Hernquist 2003). Nevertheless, there is fairly good agreement between observations and gas fractions measured in our simulations, especially at the lower mass end, irrespective of the AGN feedback model used. Larger differences between models are seen for the most massive galaxies, for which the cold gas fraction is expected to decrease as galaxies become increasingly quenched. When $M_{\text{star}} \geq 10^{11} M_{\odot}$, the TNG THERMAL model and MISTRAL-CONTINUOUS globally lead to the highest gas fractions. With the FIDUCIAL TNG physics, the cold gas fraction significantly drops for some of the galaxies, which correspond to the galaxies with the lowest SFR reported in Fig. 11. In these objects, the RANDOM KINETIC model of the low-accretion BH mode is enabled, which efficiently removes copious amounts of gas from the galaxies and prevents gaseous inflows. Finally, the impact of MISTRAL-STOCHASTIC on gas

fraction is even more dramatic, which is also consistent with the larger number of massive quenched galaxies at $z = 2$ produced with this model. The general picture depicted by Fig. 14 can also be related to the phase diagrams shown in Section 3.2.1 for our idealized galaxy simulations (Fig. 5): with the TNG THERMAL model and MISTRAL-CONTINUOUS, galaxies retain more cold gas in their ISM, while a significant fraction of this gas phase is heated, efficiently ejected out of the dense ISM and prevented from being re-accreted with MISTRAL-STOCHASTIC, especially in the highest-mass galaxies.

4.2.3 Hot gas fraction

AGN feedback is thought to play a role in the overall baryon content of galaxy groups and clusters (e.g. Eckert et al. 2021; Oppenheimer et al. 2021), by affecting the amount and thermal state of gas at these large scales, which can also be traced by the resulting thermal X-ray luminosity of the hot gas within the DM halo (McCarthy et al. 2010). Interestingly, it has been shown that the hot gas fraction can vary substantially from one simulation to another depending on the AGN feedback model adopted (e.g. Choi et al. 2015; Weinberger et al. 2017), and despite the fact that these simulations all reproduce galaxy population properties such as the massive end of the local galaxy stellar mass function (Eckert et al. 2021). Therefore, the hot gas fraction is a sensitive probe to test different implementations of AGN feedback.

For this reason, Fig. 15 shows the hot gas fraction at $z = 2$ (top panel) and $z = 0$ (bottom panel), as a function of the total mass M_{500} enclosed within a sphere of radius R_{500} . The latter is defined as the region where the mean density is 500 times the critical density of the Universe, and the hot gas fraction is defined as the ratio of the hot gas mass to the total mass within R_{500} . We define hot gas as gas with $T \geq 10^6$ K and $\rho \leq 2 \times 10^{-25} \text{ g cm}^{-3}$. For reference, we show the fit to the gas fraction–halo mass relation made by Eckert et al. (2021) from a collection of local observations, recently extended by Popesso et al. (2024, see their equation 4) to lower halo masses using eRosita observations at $z < 0.2$.

With a few exceptions, the TNG AGN feedback models and MISTRAL-CONTINUOUS tend to produce higher hot gas fractions compared to MISTRAL-STOCHASTIC. This is not only the case at $z = 2$ but also at $z = 0$, as shown in the bottom panel of Fig. 15, using the three galaxies that we re-simulated down to $z = 0$. With this restricted sample of objects, the FIDUCIAL TNG model, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC all give results reasonably close to the fit to the observations of Popesso et al. (2024) at $z = 0$. Interestingly, the FIDUCIAL TNG model has been shown to tend to overpredict hot gas fractions at $z = 0$ in the TNG simulations (Popesso et al. 2024; Rennehan et al. 2024). Here, MISTRAL-CONTINUOUS and MISTRAL-STOCHASTIC, respectively, overpredict and underpredict the hot gas fraction for one of the three galaxies. To confirm the ability of MISTRAL to reproduce realistic hot gas fractions, we would however need to repeat a similar analysis for a larger sample of massive galaxies, simulated down to $z = 0$. At this redshift, where SMBHs primarily undergo radiatively inefficient accretion, we may need to recalibrate our model, as it was not designed to operate in this regime, which we will explore in future work.

Going back to $z = 2$, simulations with MISTRAL-STOCHASTIC produce the lowest hot gas fractions across different halo mass ranges, compared to the other AGN feedback models tested. This shows that at $z = 2$, AGN winds from radiatively efficient accretion can efficiently expel gas beyond the halo virial scale, and simultaneously regulate galaxy properties such as stellar mass, star formation rate,

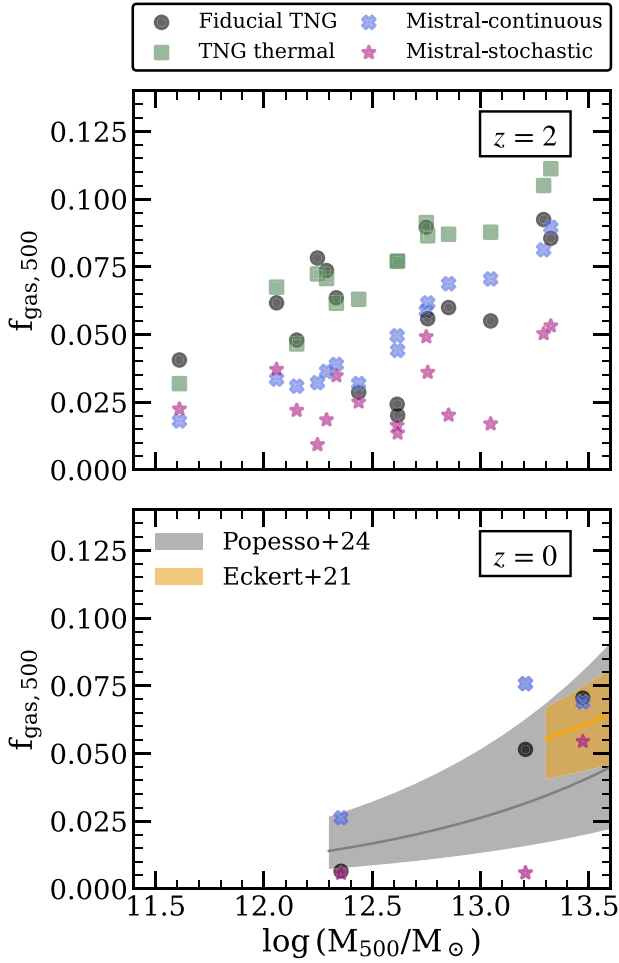


Figure 15. Hot gas fraction in the CGM at $z=2$ (top panel) and $z=0$ (bottom panel), defined as gas with $T \geq 10^6$ K and $\rho \leq 2 \times 10^{-25}$ g cm $^{-3}$, as a function of the total mass M_{500} enclosed within a sphere of radius R_{500} . The x-axis shows M_{500} from zoom simulations with the FIDUCIAL TNG physics, in order to display the four set-ups at the same x-coordinate. The shaded areas show the best fits to the gas fraction–halo mass relation determined from $z=0$ observations by Eckert et al. (2021) and Popesso et al. (2024), respectively. In this work, only three galaxies are simulated down to $z=0$ (see Appendix B).

black hole mass, and cold gas fraction. While we calibrate our model to match BH masses and stellar masses at $z=0$, MISTRAL-STOCHASTIC is not calibrated to reproduce the regulation of star formation and cold and hot gas fractions, which can be challenging for cosmological simulations (Eckert et al. 2021; Popesso et al. 2024; Rennehan et al. 2024). Thus, MISTRAL-STOCHASTIC is a promising tool for reproducing and studying the regulation of galaxy population growth due to AGN feedback.

4.2.4 Inflows and outflows

Finally, we quantify how AGN feedback regulates gas flows at different scales by focusing on mass inflow and outflow rates. To compute these two quantities, we adopt the same prescription as detailed in Section 3.2.3. Mass inflow and outflow rates are defined following equation (26). They are differentiated by the sign of their radial velocity, with positive (negative) values corresponding to outflowing (inflowing) gas. To better account for the different

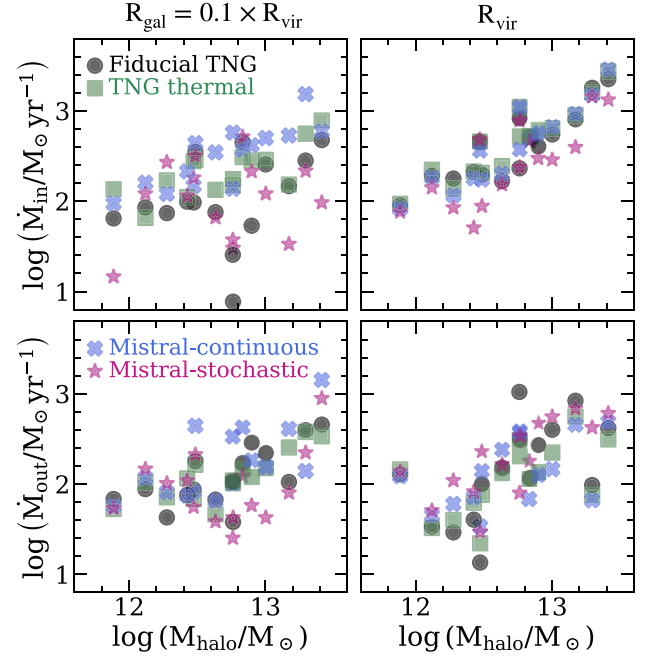


Figure 16. Mass inflow (upper plots) and outflow (bottom plots) rates at $z=2$ as a function of halo mass, measured in spherical shells of $0.01 \times R_{\text{vir}}$ in width, located at the galaxy radius (defined as 10 per cent of the virial radius, leftmost panels) and at the halo virial radius (rightmost panels). The x-axis shows the halo mass from zoom simulations with the FIDUCIAL TNG model such that inflow and outflow rates are displayed at the same x-coordinate.

masses and sizes of our 15 galaxies, we compute mass flow rates at the galaxy and the halo radius, respectively, defined as $R_{\text{gal}} = 0.1 R_{\text{vir}}$ and R_{vir} (with R_{vir} varying from 56 to 210 kpc at $z=2$ from the least to the most massive galaxy from our sample). For the same reason, the width of the shell within which flow rates are calculated is set to $0.01 \times R_{\text{vir}}$.

The resulting mass inflow and outflow rates at $z=2$ are shown as a function of halo mass in Fig. 16. At the galaxy radius, simulations with MISTRAL-CONTINUOUS have the highest mass inflow and outflow rates. This again illustrates the small-scale gaseous fountains driven with this model. With MISTRAL-CONTINUOUS, AGN feedback expels gas that rapidly falls back to the galaxy, which is notably confirmed with mass outflow rates that are lower at R_{vir} than at R_{gal} . Conversely, MISTRAL-STOCHASTIC leads to the opposite behaviour for the most massive galaxies, which, respectively, have among the lowest mass outflow rates at R_{gal} , but the highest mass outflow rates at R_{vir} . This can be explained by the fact that these massive galaxies have very little gas left at $z=2$ (Fig. 14), and experience fewer AGN feedback episodes as a result of the lower BH accretion rates (Fig. 13). When there is enough gas left to feed the SMBH, galaxies with MISTRAL-STOCHASTIC have their outflows enhanced, which is the case for galaxies with $M_{\text{halo}} \leq 10^{12.6} M_{\odot}$. This is also the case at any halo mass at higher redshift. For instance, we checked (but do not show) that mass outflow rates measured at R_{gal} are higher with MISTRAL-STOCHASTIC than with the other models at $z=3$ and $z=4$. In addition, galaxies with MISTRAL-STOCHASTIC have their inflow rates suppressed at both halo and galaxy scales. Compared to MISTRAL-CONTINUOUS and the TNG THERMAL model, the most massive galaxies with MISTRAL-STOCHASTIC have their inflow rates at R_{gal} suppressed by up to 1 dex. AGN feedback with MISTRAL-STOCHASTIC therefore prevents inflowing gas from reaching the

galaxy (by providing pressure support beyond the galaxy scale), and simultaneously efficiently removes material from the halo. MISTRAL-STOCHASTIC acts simultaneously in an ejective and preventive way, just as does the model from Choi et al. (2012), as shown by Brennan et al. (2018), and the TNG model (Zinger et al. 2020; Wright et al. 2024). Our simulations with the FIDUCIAL TNG model similarly have inflows suppressed at the galaxy scale, but both the suppression of inflow rates and the enhancement of mass outflow rates at the virial radius are less pronounced than with MISTRAL-STOCHASTIC.

5 DISCUSSION

In this section, we contextualize our results by comparing them to other mechanical, kinetic AGN feedback models (Section 5.1), before discussing the limitations of our work and potential avenues for improvement (Section 5.2).

5.1 Comparison with AGN feedback models from the literature

MISTRAL builds upon previous efforts to develop more physically motivated models of AGN-driven winds. In this section, we focus on comparing MISTRAL with other AGN feedback prescriptions adopted in cosmological simulations, in order to highlight the unique features of our model.

In this study, we illustrate how the numerical AGN energy injection scheme impacts galaxy properties and evolution. In particular, we show that modelling AGN feedback via continuous thermal energy deposition, as done in the TNG quasar model, is inefficient at driving fast large-scale winds, such as those observed in some AGN (e.g. Bischetti et al. 2019, 2024). This may be the consequence of artificial overcooling, which corresponds to the fact that thermal energy is quickly radiated away, if resolution is not high enough to let the energy propagate away from its injection site (Weinberger et al. 2018). To alleviate this issue and make the quasar feedback more efficient, one possibility is to increase the amount of energy injected, by introducing a duty cycle to store the AGN energy (Booth & Schaye 2009; Henden et al. 2018). This is what is used in the simulations presented by Schaye et al. (2015), who additionally switch from a continuous to a stochastic energy deposition scheme. Another solution, adopted in MISTRAL-CONTINUOUS, is to rely on (continuous) momentum injection. While this improves the situation in generating small-scale galactic winds, the model, as it is, does not reproduce realistic BH and galaxy properties at $z = 2$. By combining a stochastic AGN feedback scheme with momentum injection, MISTRAL-STOCHASTIC produces large-scale winds driven by BH accretion while effectively regulating BH and galaxy evolution. Using idealized simulations of galaxy groups and clusters, Huško et al. (2024) also found that directional kinetic feedback is more efficient at removing cold gas and suppressing star formation than isotropic thermal or kinetic feedback, which is similar to what we find with MISTRAL-STOCHASTIC and the TNG RANDOM KINETIC model, when compared with the TNG ISOTROPIC THERMAL model and MISTRAL-CONTINUOUS.

Although the TNG RANDOM KINETIC model produces similar outflow properties as MISTRAL-STOCHASTIC (Section 3), the two models lead to quantitatively different CGM gas morphology, and MISTRAL-STOCHASTIC needs less cumulative BH energy to produce a noticeable effect on gas fraction and star formation. These differences arise from the timing and location of AGN energy deposition. Because the two models require different conditions for releasing energy, they naturally lead to different AGN feedback duty cycles. In addition, differences in the directionality of the injection affect whether energy

is preferentially coupled to low-density CGM gas or to denser gas in the galaxy disc, which can occasionally occur with the TNG RANDOM KINETIC model. Together, these two aspects impact the efficiency of AGN feedback on large scales. In addition, the TNG RANDOM KINETIC model has only been used in the radiatively inefficient regime in a cosmological context, and it would likely require recalibration in order to operate in both SMBH regimes, as MISTRAL-STOCHASTIC does. We also note that, in this study, we do not intend to quantitatively compare the AGN-driven winds generated by MISTRAL with observations of BAL outflows. We also do not assess how realistically MISTRAL operates in the radiatively inefficient regime, or in regulating galaxies within haloes of $M_{\text{halo}} \leq 7 \times 10^{12} M_{\odot}$, using the same parametrizations for radiatively efficient AGN feedback at different redshifts. These are all aspects that we plan to explore in detail in subsequent works.

MISTRAL-STOCHASTIC is the AREPO analogue of the AGN feedback model developed by Ostriker et al. (2010) and implemented by Choi et al. (2012) in GADGET. First, we should remind the reader that the two astrophysical codes rely on a different computational method for solving the hydrodynamical equations. GADGET (Springel 2005) employs the Lagrangian SPH technique (Monaghan 1992), which considers dark matter, gas, and stars as particles. AREPO (Springel 2010) is a descendant of GADGET, which discretizes the MHD equations on a moving Voronoi mesh, combining the advantages of grid-based and SPH schemes. In both MISTRAL-STOCHASTIC and its GADGET equivalent, gas elements (cells with AREPO and particles with GADGET) are stochastically kicked at a pre-defined wind velocity. In addition, the model from Choi et al. (2012) shares the momentum imparted to the wind particle with up to two neighbouring gas elements. All particles have the same velocity increment, corresponding to a third of the initial wind velocity. Because this correspondingly decreases the total kinetic energy released, the residual energy is then deposited in the thermal form to ensure energy conservation. This feature was primarily developed for its computational advantages: reducing the wind velocity increment limits the reduction of the numerical time-step, and sharing momentum among several particles allows the model to better capture the AGN outflows at low resolution, when the number of gas particles is insufficient to propagate the AGN wind otherwise. Due to its grid-based approach, AREPO treats energy and momentum propagation via fluxes. This allows the AREPO version of the model to capture the AGN wind evolution at low resolution (by enabling cell refinement, if needed), without suffering from the technical limitations of SPH prescriptions. For this reason, we do not utilize momentum sharing in our current implementation of MISTRAL-STOCHASTIC. Unlike Choi et al. (2012), we also do not consider the radiative feedback from the SMBH X-ray radiation, which has been shown to have a negligible effect on the stellar and gas properties of massive galaxies Choi et al. (2017).

Overall, MISTRAL-STOCHASTIC produces similar results as the model from Choi et al. (2012). Using an idealized galaxy disc simulation, Choi et al. (2012) illustrated how their model produces higher wind velocity and energy compared to isotropic thermal feedback models. With $z = 0$ cosmological zoom simulations, Choi et al. (2015, 2017, 2018) showed that their AGN feedback model leads to a realistic regulation of black hole growth, star formation, and galaxy properties such as their sizes and X-ray luminosities. Using the same model, Brennan et al. (2018) analysed mass inflow and outflow rates at galaxy and halo scales, and demonstrated the ability of the Choi et al. (2012) feedback model to launch powerful galactic winds, while simultaneously acting via preventive feedback, similar to what we find and show in Fig. 16 (see their figs 12 and 13). Despite

using different astrophysical codes and galaxy formation physics, our implementation of the model of Choi et al. (2012) therefore produces qualitatively consistent results.

Other kinetic AGN feedback models have been used in cosmological simulations, primarily for the radiatively inefficient regime. In the Horizon-AGN simulation (Dubois et al. 2014), bipolar outflows are launched with a velocity of 10^4 km s^{-1} , following the model implemented in the grid code RAMSES by Dubois et al. (2012). Unlike MISTRAL, this feedback mode only operates when the Eddington ratio is lower than 0.01, and quasar feedback is modelled via thermal energy injection. A bimodal feedback scheme is also adopted in the TNG simulations (Nelson et al. 2019a), in which AGN feedback is modelled with a kinetic prescription only at low Eddington ratios (Weinberger et al. 2017). In addition, they do not impose any bipolarity for the wind: instead, its direction is randomly chosen at each feedback event, leading to collimated outflows that preferentially escape along the path of least resistance (Nelson et al. 2019b). With our model, the bipolarity of the AGN wind is controlled by an angular momentum weight in MISTRAL-CONTINUOUS, and by directly injecting momentum along the direction of the angular momentum of the inner disc with MISTRAL-STOCHASTIC.

MISTRAL-STOCHASTIC is similar in some ways to the AGN feedback model in the Simba simulations (Davé et al. 2019). Simba is the first large-scale cosmological simulation that adopted a kinetic prescription for both the radiatively efficient and inefficient regimes, using the GIZMO mesh-free code (Hopkins 2015). As we do with MISTRAL-STOCHASTIC, gas elements are ejected in a strict bipolar way, following a certain wind probability set by the desired mass loading factor (see equation 9 in Davé et al. 2019). In contrast with our work, the wind velocity varies with BH mass or Eddington ratio, depending on the feedback mode. Lower velocity winds are launched in the high-accretion BH regime (with typical velocities of a few hundreds kilometres per second), and winds as fast as $\sim 10^4 \text{ km s}^{-1}$ (similar to the velocities we use in MISTRAL) can only be launched in the low-accretion rate regime. Also conducted with the GIZMO code, the galaxy zoom simulations from the FIRE project (Hopkins et al. 2014, 2018) include a kinetic model for fast nuclear winds, without distinguishing between BH accretion regimes. Similar to the wind model of Choi et al. (2012) and to MISTRAL, their model, first introduced by Hopkins et al. (2016), assumes that a fraction of the gas accreted on to the BH is blown out as a wind at a certain velocity. However, gas particles are launched in the plane of the BH accretion disc, creating planar outflows instead of the bipolar winds obtained with MISTRAL-STOCHASTIC. More recently, Torrey et al. (2020) updated this model to inject kinetic winds using a particle spawning technique, enabling better resolution of the wind shock. In this approach, each newly spawned wind particle is given a fixed temperature, metallicity, and radially outward velocity. Using this model, Cochrane et al. (2023), Wellons et al. (2023), and Byrne et al. (2024) all show that the resulting AGN outflows regulate galaxy properties, such as their size, stellar mass, and morphology.

Most AGN feedback models used in cosmological simulations to date have been calibrated to regulate massive galaxies, and reproduce a range of $z = 0$ observations. At low redshift, AGN feedback is expected to be dominated by a radio-jet mode, active during episodes of low accretion rates. For this reason, a wealth of theoretical studies try to capture the complex jet–ISM interaction, and its impact on galaxy growth and properties (e.g. Bourne & Sijacki 2017; Talbot et al. 2021; Huško et al. 2022; Borodina et al. 2025). Conversely, less effort has been directed towards the modelling of the radiatively efficient (quasar) mode (but see e.g. Costa et al. 2020; Rennehan et al. 2024), which is much more common in gas rich high-redshift

galaxies. Given the growing number of high-redshift quasar, BAL winds, and galaxy observations that challenge our models of BH and galaxy evolution (e.g. Carnall et al. 2023; Wang et al. 2023; Yang et al. 2023; Bischetti et al. 2024; Matthee et al. 2024), and the difficulty simulations seem to have in producing sufficient numbers of massive quenched galaxies at $z = 3\text{--}7$ (e.g. Valentino et al. 2023; Lagos et al. 2025; Weibel et al. 2025; Weller et al. 2025), improving the physical fidelity of radiatively efficient AGN feedback models is extremely timely.

5.2 Uncertainties and limitations

We now summarize some intrinsic limitations of the MISTRAL model and of our work. To reproduce the outflows thought to be generated by accretion on to SMBH, MISTRAL assumes that a fraction of the inflowing gas mass, corresponding to $M_{\text{BH,wind}}$, is ejected as an AGN wind, while the rest is effectively accreted. As described in Section 2.2.1, this statement is not strictly satisfied with MISTRAL-CONTINUOUS. Instead, the AGN energy is released (in the form of momentum) into a prescribed number of BH neighbouring gas cells, for consistency with what is done (in the form of thermal energy) in the TNG quasar mode. When the total gas mass receiving the BH energy exceeds $M_{\text{BH,wind}}$, this choice overdilutes the SMBH energy, potentially affecting the strength of AGN feedback. This does not apply to MISTRAL-STOCHASTIC: the SMBH energy is distributed over the expected outflowing gas mass, provided that $M_{\text{BH,wind}}$ is higher than the gas mass resolution. Therefore, with MISTRAL-STOCHASTIC, the timing of AGN feedback episodes depends on gas resolution, via the probability of momentum injection (equation 24). In Appendix A, we show that stellar and black hole masses at $z = 3$ are barely impacted by changing the numerical resolution with MISTRAL-STOCHASTIC. However, at higher redshift ($z \gtrsim 7$), the timing of the first AGN feedback event may impact the early stages of BH and galaxy growth.

Capturing the consequences of AGN feedback on galaxy evolution also relates to the details of BH seeding and accretion, which determine the amount of SMBH energy released. In this work, BH accretion is modelled following the Bondi–Hoyle–Littleton formalism. This is a simple model, which predicts growth rates based on the SMBH mass, assuming a spherical symmetry while remaining agnostic about the angular momentum transfer of the gas. Other models, implemented in cosmological simulations, reduce the dependency of the accretion rate on the BH mass. One such example is the torque-limited accretion (Hopkins & Quataert 2011; Anglés-Alcázar, Özel & Davé 2013; Anglés-Alcázar et al. 2015, 2017), which regulates BH growth based on the fact that gas has to lose angular momentum via gravitational instabilities in order to be driven from galaxy scales down to the BH accretion disc. Together with improved BH seeding prescriptions (e.g. Bhowmick et al. 2024), more realistic treatments of BH accretion would improve the physical fidelity of AGN feedback models, and enable a better interpretation of the BH–galaxy co-evolution through cosmic time.

For this study, MISTRAL is implemented and used in the framework of the TNG simulations. Like all cosmological simulations, the TNG simulations (and therefore our work) rely on a number of simplifying assumptions regarding the galaxy formation physics. Ideally, capturing galaxy evolution would involve resolving the complex, multiphase structure of the ISM where star formation and BH accretion take place. These processes happen at sub-parsec scales that are prohibitively expensive to reconcile with cosmological volumes of several tens or hundreds of Mpc in width. To encapsulate the range of physical processes happening at unresolved scales,

the TNG simulations couple star formation and feedback subgrid models with an effective equation of state (Springel & Hernquist 2003; Vogelsberger et al. 2013). Recently, alternatives have been developed, such as SMUGGLE (Marinacci et al. 2019), a model that resolves the multiphase structure of the ISM implemented in AREPO. It additionally includes an improved prescription for stellar feedback by self-consistently treating the outflows generated during supernova explosions (see also Smith, Sijacki & Shen 2018), as opposed to the TNG decoupled stellar wind-model used in our simulations (Pillepich et al. 2018a). Adopting an improved model of the ISM could enhance the robustness of our conclusions, as the structure of the ISM itself can affect the AGN wind propagation, and hence impacts the outflow properties, such as their temperature and velocity (Ward et al. 2024; Sivasankaran et al. 2025).

Another important aspect for accurately simulating the interactions between AGN winds and galaxies is to capture the scales at which SMBH energy couples with the ambient gas. As already mentioned, BH accretion and wind launching happen at sub-parsec scales, a resolution very challenging to reach in combination with the Mpc scales of cosmological simulations. To circumvent this limitation, one promising avenue is to make use of hyper-refinement techniques, which allow one to restrict the computational load of an increased resolution to a specific location and for a limited amount of time. This is precisely the approach adopted by Anglés-Alcázar et al. (2021). Using a galaxy zoom simulation from the FIRE project (Hopkins et al. 2014, 2018), they used a hyper-Lagrangian refinement technique to connect galaxy scales to the sub-parsec scale of the BH accretion disc. However, this is achieved for a short duration of a few Myr, and they do not account for AGN feedback. In the spirit of increasing resolution dynamically around the SMBH, a variety of super-Lagrangian refinement techniques have been implemented and coupled to AGN feedback prescriptions (e.g. Curtis & Sijacki 2016a, b; Bourne & Sijacki 2017; Beckmann, Devriendt & Slyz 2019; Bourne, Sijacki & Puchwein 2019; Koudmani et al. 2019; Costa et al. 2020). However, because of their computational cost, none of them have been applied to large-scale cosmological simulations to date, nor to extended periods of time in cosmological zoom simulations.

6 CONCLUSIONS

Building on the success of the AGN feedback model from Ostriker et al. (2010) and Choi et al. (2012), we introduce and implement the MISTRAL subgrid model of AGN winds from radiatively efficient accretion into the AREPO code. MISTRAL aims to capture AGN feedback from BAL winds, which are driven during BH accretion by physical processes occurring at scales typically unresolved in cosmological simulations. To compare the impact of MISTRAL to previous models that rely on isotropic thermal energy injection, we introduce two versions of MISTRAL. MISTRAL-CONTINUOUS serves as a natural extension to the continuous isotropic thermal energy deposition scheme used at high Eddington ratios in the TNG simulations, by continuously injecting momentum in a similar spherical fashion. In the spirit of the original model developed in GADGET by Choi et al. (2012), MISTRAL-STOCHASTIC imparts momentum stochastically to gas cells, with the direction of the velocity kicks being set to create bipolar winds. We first test both versions of our model in an idealized galaxy simulation, coupling MISTRAL with the TNG galaxy formation physics models for star formation, the ISM, stellar feedback, etc. In this controlled experiment, we compare the effects of MISTRAL to those of the TNG AGN feedback models (ISOTROPIC THERMAL and RANDOM KINETIC models), and report the following results:

(i) MISTRAL generates galaxy-scale winds, whose properties depend on the momentum injection scheme. When momentum is injected continuously and radially with MISTRAL-CONTINUOUS, AGN winds are dense, short-lived, and form a galactic fountain. Alternatively, when momentum is stochastically imparted to gas cells with MISTRAL-STOCHASTIC, large-scale fast bipolar winds form, creating hot and diffuse pathways through the CGM. This result contrasts with traditional AGN feedback models that rely on continuous thermal energy injection: when using the same energy coupling efficiency as with MISTRAL, the TNG ISOTROPIC THERMAL model does not produce a noticeable AGN wind.

(ii) AGN winds from radiatively efficient accretion can significantly impact galaxy evolution. By expelling dense gas that quickly falls back to the ISM of the galaxy, MISTRAL-CONTINUOUS acts as positive feedback, enhancing the SFR through higher gas fractions than the other models. Conversely, the large scale AGN winds generated with MISTRAL-STOCHASTIC quite effectively reduce the cold gas fraction, star formation, and BH accretion.

(iii) Episodic kinetic AGN feedback models are more efficient at generating large-scale outflows than continuous energy or momentum deposition schemes. With MISTRAL-STOCHASTIC and the TNG RANDOM KINETIC model, the episodic AGN feedback events produce persistent winds that propagate to large distances from the galaxy, with mass and energy outflow rates higher than those obtained with the other models. MISTRAL-CONTINUOUS and the TNG ISOTROPIC THERMAL model lead to similar mass and energy outflow rates over time and distance, except within the inner few kpc, where MISTRAL-CONTINUOUS accumulates gas.

In the second part of our work, we examine MISTRAL's ability to reproduce realistic galaxies at $z = 2$. For this purpose, we conduct four sets of cosmological zoom simulations, targeting 15 massive galaxies from the TNG100 simulation. For each galaxy, we run simulations with both versions of MISTRAL, the TNG THERMAL model (i.e. the ISOTROPIC THERMAL and radiative AGN feedback models), and the FIDUCIAL TNG models (which include the RANDOM KINETIC model at low Eddington ratios). Isolating the TNG THERMAL model allows for a more direct comparison of MISTRAL with previous prescriptions that have been used to model AGN feedback in the radiatively efficient SMBH regime. Simulations with the FIDUCIAL TNG physics provide a more balanced comparison between MISTRAL (which covers both regimes of BH accretion) and the well tested AGN feedback model used in TNG. We mainly focus on MISTRAL-STOCHASTIC, as it produces the more realistic and interesting results. Our conclusions are as follows:

(i) MISTRAL-STOCHASTIC efficiently regulates galaxy growth across the entire range of halo masses tested, and can lead to star formation quenching at $z = 2$. Conversely, MISTRAL-CONTINUOUS and the TNG THERMAL model tend to overpredict the stellar mass of the most massive galaxies. With the FIDUCIAL TNG model, enabling the low-accretion BH mode with the RANDOM KINETIC model is essential to suppress star formation and quench galaxies at $z = 2$.

(ii) MISTRAL-STOCHASTIC shows a broadly consistent trend of BH-stellar mass evolution, with simulated BH masses that fall within the range of observational estimates of BH masses at $z = 2$. At $z = 2$, this model produces BH masses that are slightly lower than the FIDUCIAL TNG model, and lie slightly above the $z = 0$ observational SMBH versus stellar mass relation from Greene et al. (2020). All AGN feedback models tested, except MISTRAL-CONTINUOUS, also produce SMBH and stellar masses lying close to this observational relation. This suggests that the efficiency of MISTRAL-CONTINUOUS may need

to vary with time or galaxy mass to regulate BH growth consistently across different galaxy masses.

(iii) MISTRAL-STOCHASTIC realistically modulates gas fractions in massive galaxies. Our model reduces the cold gas content in the ISM of the most massive galaxies and regulates their hot gas content. The hot gas fractions at $z = 0$ are broadly consistent with observational measurements, but this result, based on three simulated haloes, would require a larger statistical sample for a more robust comparison. The regulation of both cold and hot gas fractions with MISTRAL-STOCHASTIC is achieved without requiring an SMBH mass-dependent AGN feedback scheme, unlike the FIDUCIAL TNG model. Both MISTRAL-CONTINUOUS and the TNG THERMAL model fail to suppress cold gas fractions and star formation in massive galaxies, and produce high hot gas fraction at the halo scale.

(iv) MISTRAL-STOCHASTIC acts via both preventive and ejective feedback mechanisms. Indeed, it suppresses inflows at the galaxy and halo scales while driving outflows up to the virial radius of the haloes. MISTRAL-CONTINUOUS primarily enhances mass outflow rates in the most massive galaxies at small scales, but does not provide sufficient pressure support to affect gas inflows. Unlike the TNG THERMAL model alone, the FIDUCIAL TNG model occasionally helps in preventing gas inflows and in generating outflows, although to a lesser extent than MISTRAL-STOCHASTIC does at large scales.

The numerical details of AGN feedback modelling can profoundly impact the properties of simulated galaxies, influencing our understanding of the physics governing the evolution of galaxies. In summary, our MISTRAL-STOCHASTIC model successfully reproduces realistic stellar masses and cold gas fractions at $z = 2$, and simultaneously regulates SMBH masses and hot gas fractions. It enhances outflows, removing gas from the galaxies and the haloes, and suppresses inflows, bringing some galaxies to a quiescent state by $z = 2$. MISTRAL-STOCHASTIC works self-consistently across the entire halo mass range explored, capturing the consequences of AGN feedback in the radiatively efficient regime while avoiding the BH-mass dependent TNG feedback scheme. Due to this success, MISTRAL-STOCHASTIC has been shown to be a promising tool for deciphering the role of AGN winds from radiatively efficient accretion in shaping galaxy properties through cosmic time. Our model paves the way for making robust predictions for the impact of AGN-winds using cosmological simulations. In a following work, we will create synthetic spectra following the approach of Hirschmann et al. (2017, 2019) and Hirschmann, Charlot & Somerville (2023), which will be crucial to interpret the wealth of high-redshift quasar and galaxy observations made with current and upcoming facilities such as *James Webb Space Telescope* (JWST), Euclid, and ELTs. In particular, MISTRAL-STOCHASTIC will be an invaluable tool for exploring the origin of high-redshift massive galaxy quenching, a phenomenon increasingly observed with JWST (Carnall et al. 2023; Valentino et al. 2023; de Graaff et al. 2025), but not yet fully captured by current cosmological simulations (e.g. Weller et al. 2025).

ACKNOWLEDGEMENTS

We thank the anonymous referee for providing helpful comments that improved this paper. This work benefited from the ‘Learning the Universe’ collaboration, supported by the Simons Foundation. The simulations presented in this paper have been partly performed using the facilities of the Scientific IT and Application Support Center (SCITAS) of EPFL, using the cluster Jed. This work was also supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s1231. The analysis of the simulations were

performed using the Yggdrasil machine, from the Baobab HPC service of the University of Geneva. We are grateful to Volker Springel and the TNG collaboration team for providing access to the AREPO code and to the physical modules used in the TNG simulations. We additionally thank Kung-Yi Su for sharing the initial conditions of the idealized galaxy used in this paper, Tiago Costa for his comments on the original draft, and Yohan Dubois and Jenny Greene for instructive discussions. MF and MH acknowledge funding from the Swiss National Science Foundation (SNF) via a PRIMA Grant PR00P2 193577 ‘From cosmic dawn to high noon: the role of black holes for young galaxies’. EC was supported by the National Research Foundation of Korea (NRF-RS-2023-00213322). TN acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2094 - 390783311 from the DFG Cluster of Excellence ‘ORIGINS’. RW acknowledges funding of a Leibniz Junior Research Group (project number J131/2022).

The main roles of the authors were, using the CRediT (Contribution Roles Taxonomy) system:¹¹ Marion Farcy: conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing – original draft, Michaela Hirschmann: conceptualization; funding acquisition; methodology; resources; writing – review and editing, Rachel S. Somerville: conceptualization; methodology; writing – review and editing, Ena Choi: conceptualization; methodology, Sophie Koudmani: conceptualization; methodology; writing – review and editing, Thorsten Naab: conceptualization; methodology; writing – review and editing, Rainer Weinberger: methodology; writing – review and editing, Jake S. Bennett: writing – review and editing, Aklant K. Bhowmick: writing – review and editing, Hyunseop Choi: writing – review and editing, Lars Hernquist: writing – review and editing, Julie Hlavacek-Larrondo: writing – review and editing, Bryan A. Terrazas: resources; writing – review and editing, Francesco Valentino: writing – review and editing.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. The MISTRAL model has primarily been developed by MF, MH, RSS, EC, SK, and TN. We kindly ask that any inquiries regarding the use of MISTRAL or requests for collaboration be directed to these authors. The IllustrisTNG simulations are publicly available and accessible at www.tng-project.org/data (Nelson et al. 2019a).

REFERENCES

- Abuter R. et al., 2024, *Nature*, 627, 281
- Anglés-Alcázar D., Özel F., Davé R., 2013, *ApJ*, 770, 5
- Anglés-Alcázar D., Özel F., Davé R., Katz N., Kollmeier J. A., Oppenheimer B. D., 2015, *ApJ*, 800, 127
- Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017, *MNRAS*, 464, 2840
- Anglés-Alcázar D. et al., 2021, *ApJ*, 917, 53
- Arav N., Liu G., Xu X., Stidham J., Benn C., Chamberlain C., 2018, *ApJ*, 857, 60
- Beckmann R. S., Devriendt J., Slyz A., 2019, *MNRAS*, 483, 3488
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143
- Bertemes C. et al., 2025, *A&A*, 693, A176

¹¹<https://credit.niso.org>

- Bhowmick A. K., Blecha L., Torrey P., Weinberger R., Kelley L. Z., Vogelsberger M., Hernquist L., Somerville R. S., 2024, *MNRAS*, 529, 3768
- Bischetti M., Maiolino R., Carniani S., Fiore F., Piconcelli E., Fluetsch A., 2019, *A&A*, 630, A59
- Bischetti M. et al., 2023, *ApJ*, 952, 44
- Bischetti M. et al., 2024, *ApJ*, 970, 9
- Blandford R. D., Begelman M. C., 1999, *MNRAS*, 303, L1
- Blandford R. D., Znajek R. L., 1977, *MNRAS*, 179, 433
- Bondi H., 1952, *MNRAS*, 112, 195
- Bondi H., Hoyle F., 1944, *MNRAS*, 104, 273
- Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53
- Borodina O. et al., 2025, *ApJ*, 981, 149
- Bourne M. A., Sijacki D., 2017, *MNRAS*, 472, 4707
- Bourne M. A., Sijacki D., Puchwein E., 2019, *MNRAS*, 490, 343
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645
- Brennan R., Choi E., Somerville R. S., Hirschmann M., Naab T., Ostriker J. P., 2018, *ApJ*, 860, 14
- Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
- Byrne L. et al., 2024, *ApJ*, 973, 149
- Carnall A. C. et al., 2023, *Nature*, 619, 716
- Chabrier G., 2003, *PASP*, 115, 763
- Choi E., Ostriker J. P., Naab T., Johansson P. H., 2012, *ApJ*, 754, 125
- Choi E., Naab T., Ostriker J. P., Johansson P. H., Moster B. P., 2014, *MNRAS*, 442, 440
- Choi E., Ostriker J. P., Naab T., Oser L., Moster B. P., 2015, *MNRAS*, 449, 4105
- Choi E., Ostriker J. P., Naab T., Somerville R. S., Hirschmann M., Núñez A., Hu C.-Y., Oser L., 2017, *ApJ*, 844, 31
- Choi E., Somerville R. S., Ostriker J. P., Naab T., Hirschmann M., 2018, *ApJ*, 866, 91
- Choi H., Leighly K. M., Terndrup D. M., Gallagher S. C., Richards G. T., 2020, *ApJ*, 891, 53
- Choi H., Leighly K. M., Terndrup D. M., Dabbieri C., Gallagher S. C., Richards G. T., 2022, *ApJ*, 937, 74
- Cicone C. et al., 2014, *A&A*, 562, A21
- Ciotti L., Ostriker J. P., Proga D., 2010, *ApJ*, 717, 708
- Cochrane R. K. et al., 2023, *MNRAS*, 523, 2409
- Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018, *MNRAS*, 479, 2079
- Costa T., Pakmor R., Springel V., 2020, *MNRAS*, 497, 5229
- Croton D. J. et al., 2006, *MNRAS*, 365, 11
- Curtis M., Sijacki D., 2016a, *MNRAS*, 457, L34
- Curtis M., Sijacki D., 2016b, *MNRAS*, 463, 63
- Dai X., Shankar F., Sivakoff G. R., 2008, *ApJ*, 672, 108
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, 486, 2827
- Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, 433, 604
- Di Matteo T., Khandai N., DeGraf C., Feng Y., Croft R. A. C., Lopez J., Springel V., 2012, *ApJ*, 745, L29
- Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, *MNRAS*, 420, 2662
- Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, *MNRAS*, 433, 3297
- Dubois Y. et al., 2014, *MNRAS*, 444, 1453
- Eckert D., Gaspari M., Gastaldello F., Le Brun A. M. C., O'Sullivan E., 2021, *Universe*, 7, 142
- Fabian A. C., 2012, *ARA&A*, 50, 455
- Fabian A. C. et al., 2011, *MNRAS*, 418, 2154
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Fiore F. et al., 2017, *A&A*, 601, A143
- Förster Schreiber N. M. et al., 2014, *ApJ*, 787, 38
- Franx M., van Dokkum P. G., Förster Schreiber N. M., Wuyts S., Labbé I., Toft S., 2008, *ApJ*, 688, 770
- Gaspari M., Melioli C., Brighenti F., D'Ercole A., 2011, *MNRAS*, 411, 349
- Gaspari M., Ruszkowski M., Sharma P., 2012, *ApJ*, 746, 94
- Genzel R. et al., 2014, *ApJ*, 796, 7
- de Graaff A. et al., 2025, *Nat. Astron.*, 9, 280
- Greene J. E., Strader J., Ho L. C., 2020, *ARA&A*, 58, 257
- Habouzit M. et al., 2021, *MNRAS*, 503, 1940
- Habouzit M. et al., 2022, *MNRAS*, 511, 3751
- Hamann F., Chartas G., Reeves J., Nardini E., 2018, *MNRAS*, 476, 943
- Harrison C. M., 2017, *Nat. Astron.*, 1, 0165
- Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, *MNRAS*, 441, 3306
- Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, *MNRAS*, 479, 5385
- Hernquist L., 1990, *ApJ*, 356, 359
- Herrera-Camus R. et al., 2020, *A&A*, 633, L4
- Hewett P. C., Foltz C. B., 2003, *AJ*, 125, 1784
- Hirschmann M., Khochfar S., Burkert A., Naab T., Genel S., Somerville R. S., 2010, *MNRAS*, 407, 1016
- Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, *MNRAS*, 442, 2304
- Hirschmann M., Charlot S., Feltre A., Naab T., Choi E., Ostriker J. P., Somerville R. S., 2017, *MNRAS*, 472, 2468
- Hirschmann M., Charlot S., Feltre A., Naab T., Somerville R. S., Choi E., 2019, *MNRAS*, 487, 333
- Hirschmann M., Charlot S., Somerville R. S., 2023, *MNRAS*, 526, 3504
- Hlavacek-Larrondo J., Li Y., Churazov E., 2022, in Bambi C., Sanganello A., eds, *Handbook of X-ray and Gamma-ray Astrophysics*. Springer Nature, Singapore, p. 5
- Hopkins P. F., 2015, *MNRAS*, 450, 53
- Hopkins P. F., Quataert E., 2011, *MNRAS*, 415, 1027
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
- Hopkins P. F., Torrey P., Faucher-Giguère C.-A., Quataert E., Murray N., 2016, *MNRAS*, 458, 816
- Hopkins P. F. et al., 2018, *MNRAS*, 480, 800
- Hoyle F., Lyttleton R. A., 1939, *Proc. Cambridge Philos. Soc.*, 35, 405
- Huško F., Lacey C. G., Schaye J., Schaller M., Nobels F. S. J., 2022, *MNRAS*, 516, 3750
- Huško F., Lacey C. G., Schaye J., Nobels F. S. J., Schaller M., 2024, *MNRAS*, 527, 5988
- Huško F., Lacey C. G., Roper W. J., Schaye J., Briggs J. M., Schaller M., 2025, *MNRAS*, 537, 2559
- Jahnke K., Macciò A. V., 2011, *ApJ*, 734, 92
- Kauffmann G., Haehnelt M., 2000, *MNRAS*, 311, 576
- Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, *MNRAS*, 450, 1349
- King A., Pounds K., 2015, *ARA&A*, 53, 115
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Koudmani S., Sijacki D., Bourne M. A., Smith M. C., 2019, *MNRAS*, 484, 2047
- Lagos C. d. P. et al., 2025, *MNRAS*, 536, 2324
- Liu W. et al., 2024, *ApJ*, 976, 33
- Maccarone T. J., 2003, *A&A*, 409, 697
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Magorrian J. et al., 1998, *AJ*, 115, 2285
- Maiolino R. et al., 2024, *A&A*, 691, A145
- Marinacci F., Pakmor R., Springel V., 2014, *MNRAS*, 437, 1750
- Marinacci F. et al., 2018, *MNRAS*, 480, 5113
- Marinacci F., Sales L. V., Vogelsberger M., Torrey P., Springel V., 2019, *MNRAS*, 489, 4233
- Matthee J. et al., 2024, *ApJ*, 963, 129
- Matzeu G. A. et al., 2023, *A&A*, 670, A182
- McCarthy I. G. et al., 2010, *MNRAS*, 406, 822
- McNamara B. R. et al., 2000, *ApJ*, 534, L135
- Merloni A., Heinz S., 2008, *MNRAS*, 388, 1011
- Miller M. J., Bregman J. N., 2013, *ApJ*, 770, 118
- Miller M. J., Bregman J. N., 2015, *ApJ*, 800, 14
- Monaghan J. J., 1992, *ARA&A*, 30, 543
- Muzzin A. et al., 2013, *ApJ*, 777, 18
- Naab T., Ostriker J. P., 2017, *ARA&A*, 55, 59
- Naiman J. P. et al., 2018, *MNRAS*, 477, 1206
- Narayan R., Yi I., 1994, *ApJ*, 428, L13
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Nelson D. et al., 2015, *Astron. Comput.*, 13, 12

Nelson D. et al., 2018, *MNRAS*, 475, 624
 Nelson D. et al., 2019a, *Comput. Astrophys. Cosmol.*, 6, 2
 Nelson D. et al., 2019b, *MNRAS*, 490, 3234
 Newman A. B. et al., 2025, preprint (arXiv:2503.17478)
 Oppenheimer B. D., Babul A., Bahé Y., Butsky I. S., McCarthy I. G., 2021, *Universe*, 7, 209
 Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, *ApJ*, 725, 2312
 Ostriker J. P., Choi E., Ciotti L., Novak G. S., Proga D., 2010, *ApJ*, 722, 642
 Pandya V. et al., 2017, *MNRAS*, 472, 2054
 Pillepich A. et al., 2018a, *MNRAS*, 473, 4077
 Pillepich A. et al., 2018b, *MNRAS*, 475, 648
 Planck Collaboration XIII, 2016, *A&A*, 594, A13
 Popesso P. et al., 2024, preprint (arXiv:2411.16555)
 Reines A. E., Volonteri M., 2015, *ApJ*, 813, 82
 Rennehan D., Babul A., Moa B., Davé R., 2024, *MNRAS*, 532, 4793
 Richards G. T. et al., 2011, *AJ*, 141, 167
 Rodríguez-Gómez V. et al., 2015, *MNRAS*, 449, 49
 Rodríguez Hidalgo P., Khatri A. M., Hall P. B., Haas S., Quintero C., Khatu V., Kowash G., Murray N., 2020, *ApJ*, 896, 151
 Schaye J. et al., 2015, *MNRAS*, 446, 521
 Schaye J. et al., 2023, *MNRAS*, 526, 4978
 Schödel R. et al., 2002, *Nature*, 419, 694
 Schreiber C. et al., 2015, *A&A*, 575, A74
 Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337
 Shuntov M. et al., 2022, *A&A*, 664, A61
 Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, *MNRAS*, 380, 877
 Silk J., Rees M. J., 1998, *A&A*, 331, L1
 Sivasankaran A. et al., 2025, *MNRAS*, 537, 817
 Smith M. C., Sijacki D., Shen S., 2018, *MNRAS*, 478, 302
 Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
 Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, *MNRAS*, 391, 481
 Springel V., 2005, *MNRAS*, 364, 1105
 Springel V., 2010, *MNRAS*, 401, 791
 Springel V., 2015, *Astrophysics Source Code Library*, record ascl:1502.003
 Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
 Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
 Springel V., Di Matteo T., Hernquist L., 2005, *MNRAS*, 361, 776
 Springel V. et al., 2018, *MNRAS*, 475, 676
 Steinborn L. K., Dolag K., Hirschmann M., Prieto M. A., Remus R.-S., 2015, *MNRAS*, 448, 1504
 Su K.-Y. et al., 2019, *MNRAS*, 487, 4393
 Suh H., Civano F., Trakhtenbrot B., Shankar F., Hasinger G., Sanders D. B., Allevato V., 2020, *ApJ*, 889, 32
 Tacconi L. J. et al., 2010, *Nature*, 463, 781
 Tacconi L. J., Genzel R., Sternberg A., 2020, *ARA&A*, 58, 157
 Talbot R. Y., Bourne M. A., Sijacki D., 2021, *MNRAS*, 504, 3619
 Tanaka T. S. et al., 2025, *ApJ*, 979, 215
 Terrazas B. A. et al., 2020, *MNRAS*, 493, 1888
 Tombesi F., Cappi M., Reeves J. N., Palumbo G. G. C., Braitto V., Dadina M., 2011, *ApJ*, 742, 44
 Tombesi F., Cappi M., Reeves J. N., Nemmen R. S., Braitto V., Gaspari M., Reynolds C. S., 2013, *MNRAS*, 430, 1102
 Torrey P. et al., 2020, *MNRAS*, 497, 5292
 Tremaine S. et al., 2002, *ApJ*, 574, 740
 Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, *MNRAS*, 470, 1121
 Valentino F. et al., 2023, *ApJ*, 947, 20
 Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, *A&AR*, 28, 2
 Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031
 Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
 Wang B. et al., 2023, *ApJ*, 957, L34
 Ward S. R., Costa T., Harrison C. M., Mainieri V., 2024, *MNRAS*, 533, 1733
 Weaver J. R. et al., 2023, *A&A*, 677, A184
 Weibel A. et al., 2025, *ApJ*, 983, 11

Weinberger R. et al., 2017, *MNRAS*, 465, 3291
 Weinberger R. et al., 2018, *MNRAS*, 479, 4056
 Weinberger R., Springel V., Pakmor R., 2020, *ApJS*, 248, 32
 Weller E. J., Pacucci F., Ni Y., Hernquist L., Park M., 2025, *ApJ*, 979, 181
 Wellons S. et al., 2023, *MNRAS*, 520, 5394
 Wright R. J., Somerville R. S., Lagos C. d. P., Schaller M., Davé R., Anglés-Alcázar D., Genel S., 2024, *MNRAS*, 532, 3417
 Wylezalek D., Flores A. M., Zakamska N. L., Greene J. E., Riffel R. A., 2020, *MNRAS*, 492, 4680
 Yang J. et al., 2023, *ApJS*, 269, 27
 Zel'dovich Y. B., 1970, *A&A*, 5, 84
 Zinger E. et al., 2020, *MNRAS*, 499, 768

APPENDIX A: TESTING MISTRAL AT HIGHER RESOLUTION

In order to study the resolution convergence of the results discussed in Section 4, we re-simulate three of the 15 galaxies of our sample with an increased resolution. Specifically, we target haloes with group number 61, 30, and 10. We construct their new ICs at two times higher spatial resolution, i.e. we increase the number of particles by a factor of 8, which roughly corresponds to TNG50 resolution (as summarized in Table A1). To limit their computational cost, we run these higher resolution simulations down to $z = 3$. We refer to them as the $\text{res} \times 2$ simulations, as opposed to the $\text{res} \times 1$ simulations studied in Section 4, that adopt the resolution of their parent TNG100 simulation. Apart from resolution, both sets of simulations use the same galaxy formation physics, as described throughout the paper. We only increase the weighted number of gas cells in the BH smoothing volume $n_{\text{BH,ngb}}$ by a factor of 2, as was originally done from TNG100 to TNG50 (and as explained by Weinberger et al. 2017).

Fig. A1 shows the halo, stellar, and black hole masses in the $\text{res} \times 2$ versus the $\text{res} \times 1$ simulations. We compare these three properties at $z = 3$ for zoom simulations using MISTRAL-CONTINUOUS (blue), MISTRAL-STOCHASTIC (purple), as well as the FIDUCIAL TNG physics (black), for reference. Halo and stellar masses are fairly well converged with these three set-ups, as they follow closely the one-to-one relation shown in dashed lines. While MISTRAL-STOCHASTIC leads to similar BH masses regardless of resolution, MISTRAL-CONTINUOUS and the FIDUCIAL TNG model can both lead to more and less massive BHs when changing resolution, and this effect is more pronounced with MISTRAL-CONTINUOUS. Resolution convergence in cosmological simulations is a long-standing issue, due to the complex and non-linear nature of the interactions between feedback processes and galaxy evolution. In order to recover similar evolution

Table A1. Simulation resolution as a function of the zoom factor, with $\text{res} \times 1$ corresponding to TNG100 resolution and $\text{res} \times 2$ to (roughly) TNG50 resolution. From top to bottom: m_{DM} : dark matter mass resolution; m_{baryons} : mean baryon mass resolution; $\epsilon_{\text{DM,stars}}$: Plummer equivalent gravitational softening of the collisionless component; ϵ_{gas} : minimum value of the gravitational softening of the gas; $n_{\text{BH,ngb}}$: weighted number of gas cells in the BH smoothing volume; z_{end} : redshift reached at the end of the run, for the last snapshot.

Resolution	TNG100 ($\text{res} \times 1$)	\sim TNG50 ($\text{res} \times 2$)
$m_{\text{DM}} [\text{M}_{\odot}]$	7.5×10^6	9.4×10^5
$m_{\text{baryons}} [\text{M}_{\odot}]$	1.4×10^6	1.7×10^5
$\epsilon_{\text{DM,stars}} [\text{cpc}/h]$	1000	500
$\epsilon_{\text{gas}} [\text{cpc}/h]$	125	62.5
$n_{\text{BH,ngb}}$	256	512
z_{end}	2	3

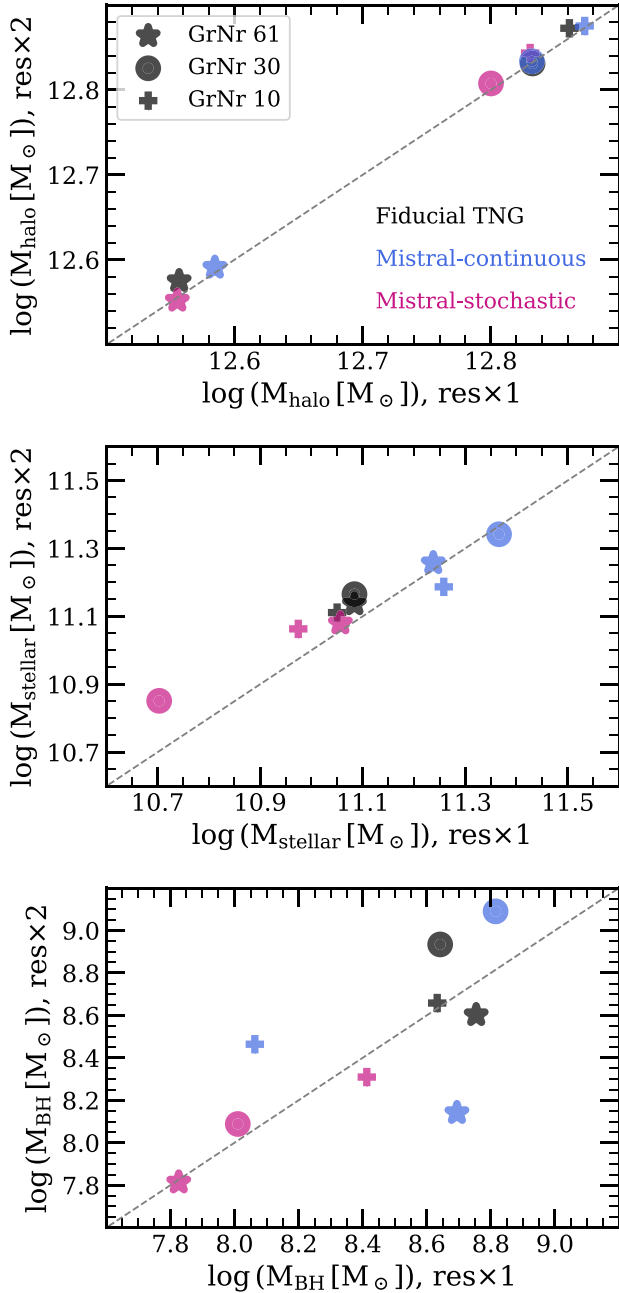


Figure A1. Halo mass, stellar mass, and black hole mass at $z=3$ for three galaxy simulations at TNG100 (res $\times 1$) and TNG50 (res $\times 2$) resolution. Zoom simulations with the FIDUCIAL TNG physics, MISTRAL-CONTINUOUS, and MISTRAL-STOCHASTIC are shown in black, blue, and purple, respectively. In each panel, the dashed line shows the 1:1 relation.

of stellar and black hole masses with time, it may be necessary to recalibrate the efficiency of AGN feedback against $z=0$ observations when changing resolution, which we will address in a future work.

APPENDIX B: CALIBRATION OF THE WIND EFFICIENCY PARAMETER FOR MISTRAL

Our MISTRAL model contains two free parameters: the average wind velocity v_w , and the AGN feedback efficiency ϵ_w . Throughout this paper, we keep $v_w = 10^4 \text{ km s}^{-1}$, as done by Choi et al. (2012, 2014, 2015, 2017, 2018). In order to determine which ϵ_w should be used in our cosmological zoom simulations, we explore the impact of varying this parameter on three of the 15 galaxies from our sample. More specifically, we focus on galaxies hosted in the haloes with group number 631, 123, and 57 at $z=2$ (snapshot 33 from TNG100), which correspond to GrNr = 692, 101, and 64 at $z=0$ (snapshot 99 from TNG100, from which their ICs are extracted). We simulate these three haloes with MISTRAL-CONTINUOUS and MISTRAL-STOCHASTIC, down to $z=0$, and with $\epsilon_w = 5 \times 10^{-4}$, 10^{-3} , and 5×10^{-3} .

In Fig. B1 and B2, we show the results when increasing the value of ϵ_w from light to darker shades of blue and purple, for MISTRAL-CONTINUOUS and MISTRAL-STOCHASTIC, respectively. For reference, we also rerun these three objects using the zoom-in technique with the FIDUCIAL TNG physics, which is displayed with black symbols. From left to right, the two figures show SFR and BH mass versus stellar mass, and the SMHM relation at $z=0$ (Fig. B1) and $z=2$ (Fig. B2). Overall, and with both versions of MISTRAL, increasing ϵ_w leads to smaller black hole masses and to slightly higher stellar masses. This is because the fraction of gas mass accreted by the BH decreases when ϵ_w increases (while the fraction of gas mass receiving BH energy increases, see also Fig. 1), which reduces the total amount of BH energy released and, as a result, reduces the ability of AGN feedback to impact star formation. For this reason, galaxies tend to have lower SFRs and to be more easily quenched at $z=0$ in the runs with $\epsilon_w = 5 \times 10^{-4}$. In this paper, we adopt $\epsilon_w = 10^{-3}$ as the best compromise to produce realistic black hole and stellar masses at $z=0$ and $z=2$, simultaneously suppressing star formation at $z=0$ for most simulations. A higher (respectively, lower) value of ϵ_w could have been used with MISTRAL-STOCHASTIC (MISTRAL-CONTINUOUS), but we prefer to use the same value in both models. In particular, this choice is motivated by the fact that $\epsilon_w = 5 \times 10^{-3}$ with MISTRAL produces low BH masses at $z=2$, which would make it less relevant for studying the impact of AGN-driven winds at cosmic noon.

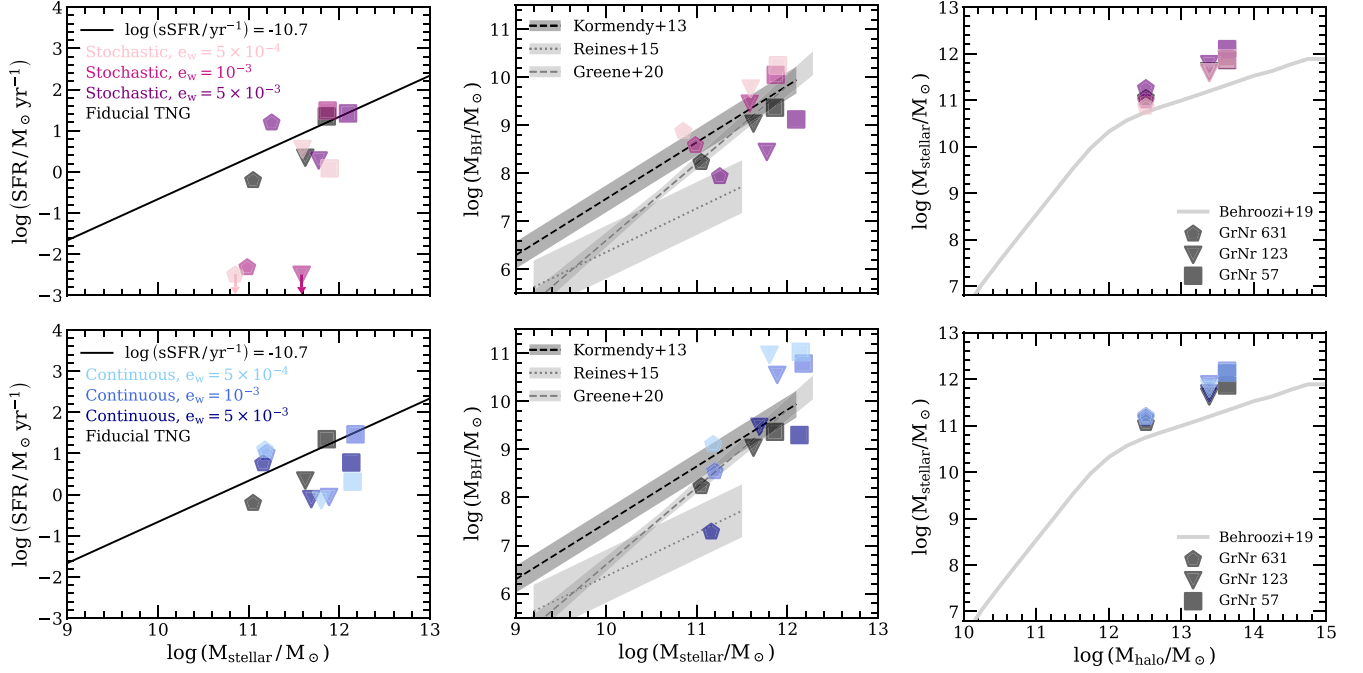


Figure B1. Scaling relations at $z = 0$ for zoom simulations with MISTRAL-STOCHASTIC in the upper row (respectively, MISTRAL-CONTINUOUS in the bottom row), with $\epsilon_w = 5 \times 10^{-4}$, 10^{-3} , 5×10^{-3} in pink (light blue), purple (blue), and dark purple (dark blue). For reference, we show the counterpart zoom simulations with the FIDUCIAL TNG physics in black. From left to right: star formation rate and BH mass versus stellar mass (with observational relations from Kormendy & Ho 2013; Reines & Volonteri 2015; Greene et al. 2020), and SMHM relation (displayed at the same x -coordinate, compared with the empirical relation from Behroozi et al. 2019). The arrows in the upper left plot denote an SFR lower than the limit of the plot. The legend with the markers indicates the halo group number at $z = 2$, for consistency with the rest of the paper.

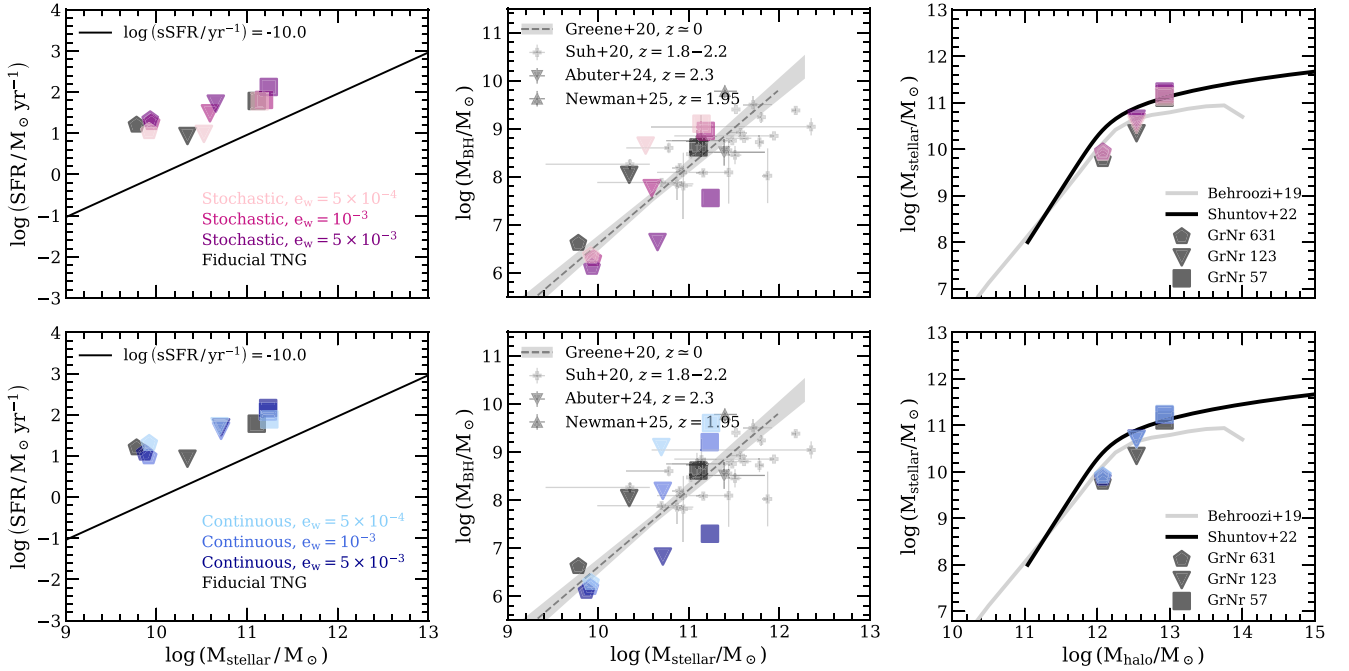


Figure B2. Same as Fig. B1 but at $z = 2$, with observational estimates from Suh et al. (2020), Abuter et al. (2024), and Newman et al. (2025) in the middle panels and from Shuntov et al. (2022) in the right ones.

This paper has been typeset from a \LaTeX file prepared by the author.