






Modeling gravitational wave sources in the MillenniumTNG simulations

Federico Marinacci^{1,2,*}, Marco Baldi^{1,2,3}, Giuliano Iorio⁴, M. Celeste Artale⁵, Michela Mapelli^{6,7},
Volker Springel⁸, Sownak Bose⁹, and Lars Hernquist¹⁰

¹ Dipartimento di Fisica e Astronomia “Augusto Righi”, Università di Bologna, Via Piero Gobetti 93/2, I-40129 Bologna, Italy

² INAF, Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, I-40129 Bologna, Italy

³ INFN, Sezione di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy

⁴ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), c. Martí i Franquès 1, 08028 Barcelona, Spain

⁵ Universidad Andres Bello, Facultad de Ciencias Exactas, Departamento de Física y Astronomía, Instituto de Astrofísica, Fernandez Concha 700, Las Condes, Santiago RM, Chile

⁶ Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert Ueberle Str. 2, D-69120 Heidelberg, Germany

⁷ Universität Heidelberg, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, D-69120 Heidelberg, Germany

⁸ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, D-85748 Garching, Germany

⁹ Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH13LE, UK

¹⁰ Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA

Received 6 October 2025 / Accepted 9 January 2026

ABSTRACT

We introduce a flexible framework for building gravitational wave (GW) event catalogs in hydrodynamic simulations of galaxy formation. Our framework couples the state-of-the-art binary population synthesis code SEVN with AREPO-GW – a module fully integrated into the moving-mesh code AREPO – to assign merger events of binary compact objects to stellar particles in simulations by stochastically sampling merger tables generated with SEVN. AREPO-GW supports both on-the-fly operation, producing event catalogs during simulations, and post-processing, using snapshots from existing runs. The algorithm is fully parallel and can be readily adapted to outputs from simulation codes beyond AREPO. To demonstrate the capabilities of our new framework, we applied AREPO-GW in post-processing to simulations from the MillenniumTNG suite, including its flagship box – one of the largest full-physics cosmological simulations to date. We investigate key properties of the resulting GW event catalog, built on SEVN predictions, focusing on comoving merger rates, formation efficiencies, delay-time distributions, and progenitor mass and metallicity distributions. We also examine how these properties vary with simulated volume. We find that GW progenitor rates closely track simulated star formation histories and are generally consistent with current observational constraints at low redshift, aside from an excess – by a factor of ~ 4.5 – in binary black hole mergers, in line with trends reported in the literature. Moreover, our binary black hole merger rates decline more slowly with redshift than current observational estimates for $z \lesssim 1$. Finally, the analysis of progenitor mass functions across different formation channels reveals only mild redshift evolution, in agreement with earlier studies, while the binary black hole mass function displays features compatible with current observational determinations. These findings highlight the potential of our novel framework to enable detailed predictions for upcoming GW surveys within a full cosmological context.

Key words. gravitational waves – methods: numerical – binaries: close – stars: black holes – stars: neutron – cosmology: theory

1. Introduction

The first direct detection of gravitational waves (GWs) by LIGO in 2015 ushered in the era of GW astronomy. Since then, a growing number of detections has enabled detailed studies of binary compact objects, such as binary black holes (BBHs), binary neutron stars (BNSs), and black hole–neutron star (BHNS) systems. The fourth observing run (O4) of the LIGO-Virgo-KAGRA (LVK) collaboration has further increased the number of observed events recently collected in the Gravitational-Wave Transient Catalog 4 (GWTC-4; [The LIGO Scientific Collaboration, the Virgo Collaboration, & the KAGRA Collaboration 2025](#), hereafter [LVK25](#)). These new results significantly improve our ability to probe the population demographics of merging compact objects and their astrophysical environments (see also [Abbott et al. 2019, 2021, 2023a](#), for previous

GWTC releases). Next-generation GW detectors will expand the observable volume dramatically, allowing exploration of BBH mergers out to very high redshifts, potentially reaching $z \lesssim 100$ (see, e.g., [Evans et al. 2021](#); [Branchesi et al. 2023](#); [Abac et al. 2025](#)).

To interpret the population of GW events, it is essential to establish a physically motivated framework that connects compact binary formation to the cosmic history of star formation and the chemical enrichment of the Universe. Several methods have been proposed to address this challenge. A possible solution is to use semi-analytic techniques or observationally calibrated star formation histories, convolved with parameterized binary formation prescriptions and merger time distributions, to estimate merger rate densities across cosmic time ([Adhikari et al. 2020](#); [Rauf et al. 2023](#); [Vijaykumar et al. 2024](#); [Dehghani et al. 2025](#)). Although computationally inexpensive, these approaches often cannot track mergers within the complex dynamical environments of evolving galaxies and halos.

* Corresponding author: federico.marinacci2@unibo.it

A more recent class of models has attempted to overcome this issue by combining merger trees – derived from N -body or hydrodynamic simulations – or empirical galaxy scaling relations with (simplified) binary evolution models (Dominik et al. 2012, 2013; Lamberts et al. 2016; Boco et al. 2019; Neijssel et al. 2019; Tang et al. 2020; Boco et al. 2021; Santoliquido et al. 2022; van Son et al. 2023; Boesky et al. 2024; Chruślińska 2024; Sgalletta et al. 2025). This hybrid approach captures certain aspects of hierarchical galaxy growth and metallicity enrichment but still lacks a mapping of the spatial distribution of events and relies on assumptions underlying the empirical scaling relations and binary evolution models. Similar techniques have also been applied to estimate the GW background produced by mergers of supermassive black holes at the centers of galaxies (see, e.g., Kelley et al. 2017a,b; Siwek et al. 2020).

Finally, a more comprehensive (and computationally expensive) method involves the post-processing of cosmological hydrodynamic simulations with stellar evolution and binary population synthesis models. In these studies, stellar particles are assigned binary merger probabilities based on metallicity, age, and star formation history, enabling predictions for merger rates and host galaxy properties (Mapelli et al. 2017; Schneider et al. 2017; Mapelli et al. 2018; Artale et al. 2019, 2020a,b; Rose et al. 2021; Perna et al. 2022; Sgalletta et al. 2023).

Despite their success, all these strategies face several limitations. First, simulations with small cosmological volumes ($\lesssim 100^3 \text{ Mpc}^3$) often fail to sample rare, high-density environments such as galaxy clusters, which may host elevated merger rates due to their dense star-forming regions or dynamical assembly histories. Also, semi-analytical and empirical schemes are unable to capture feedback-regulated star formation and enrichment in a self-consistent manner, which may be essential for robust predictions of merger rates, delay times, and host galaxy properties.

To overcome these challenges, we advocate for an approach that couples binary population synthesis models directly into large-volume cosmological hydrodynamic simulations, enabling compact binary formation and evolution to be computed self-consistently from the simulation properties, either on-the-fly during the simulation runtime or in post-processing starting from the simulation outputs. This method ensures that the local physical conditions such as gas density, metallicity, stellar age, and dynamical environment, are accurately reflected in the properties and evolution of the resulting compact binary populations. In this framework, the cosmic evolution of GW sources is naturally coupled with galaxy formation physics, feedback, and structure growth, also providing a way to directly probe their spatial distribution, clustering, and bias with respect to the stellar and matter density fields.

In this work, we integrated results obtained from the binary population synthesis code SEVN (Stellar EVolution for N -body Spera et al. 2015, 2019; Mapelli et al. 2020; Iorio et al. 2023) directly into the large-volume hydrodynamic simulations performed with the moving-mesh code AREPO (Springel 2010; Pakmor et al. 2016) through a fully integrated module referred to as AREPO-GW. As a first application, we employed this framework on the cosmological simulations of the MillenniumTNG (hereafter MTNG) project (Pakmor et al. 2023). These simulations span a comoving volume of $\approx 740^3 \text{ Mpc}^3$ with full hydrodynamics, metal-dependent cooling, stellar and active galactic nucleus feedback, and a calibrated galaxy formation model consistent with key observables. By combining SEVN with the MTNG simulation suite, we track the formation, evolution, and mergers of compact binaries in self-consistent connection with

the local star formation rate, gas-phase metallicity, and galactic environment of each stellar particle.

This work presents the methodology of – and initial results obtained with – our new framework. Specifically, we focus on predictions for the redshift evolution of GW merger rates, the metallicity distribution of progenitors, and the dependence of merger features on simulated volume and the global properties of the MTNG flagship simulation. Our goal is to provide a physically motivated model to interpret current and future GW observations, while also enabling cross-correlation with electromagnetic surveys and cosmological tracers. Ultimately, our approach opens new avenues for using GWs not only as probes of stellar evolution but also as tools to investigate the structure and evolution of the Universe.

The paper is organized as follows. In Section 2, we provide a brief overview of the MTNG project, including a description of the key physical processes implemented in its galaxy formation model and the properties of the simulation boxes analyzed in this work. Section 3 outlines the methodology developed to generate GW event catalogs by associating stellar binary mergers with star particles in hydrodynamic galaxy formation simulations. In Section 4, we present the main results of our analysis, focusing on the key properties of GW progenitor events in the flagship simulation box of the MTNG project. Finally, in Section 5, we summarize our findings and conclude. Appendix A presents a quantitative comparison between the two operational modes of the AREPO-GW framework, while Appendix B extends the analysis of GW progenitor properties to all the other simulation boxes from the MTNG suite included in this study.

2. The MillenniumTNG project

The MTNG project is a suite of cosmological simulations that merge the large simulated volume of the Millennium simulation (Springel et al. 2005) with a state-of-the-art galaxy formation physics model. Therefore, MTNG represents a unique combination between the need for extremely large simulation volumes for cosmological applications and recent advancements in hydrodynamic simulations that enable the self-consistent generation of a realistic galaxy population.

The galaxy formation physics model of MTNG is based on that of the highly successful IllustrisTNG simulations and includes as core processes: (i) an explicit treatment of the interstellar medium based on a physically motivated effective equation of state (Springel & Hernquist 2003), (ii) a stochastic prescription of star formation (Springel & Hernquist 2003), (iii) primordial and metal line cooling (Vogelsberger et al. 2013; Pillepich et al. 2018a), (iv) mass return from stars and metal enrichment of the interstellar gas by core-collapse supernovae, type Ia supernovae, and asymptotic giant branch stars (Vogelsberger et al. 2013), (v) an effective model for scaling the properties (speed, mass, and metal loading) of galactic winds (Pillepich et al. 2018a), and (vi) a model for the seeding, growth and feedback from supermassive black holes (Weinberger et al. 2017). We refer the interested reader to Weinberger et al. (2017) and Pillepich et al. (2018a) for further details about the physical processes implemented in IllustrisTNG.

Here, we analyze the hydrodynamic full physics boxes of the MTNG project, which are evolved with the moving-mesh code AREPO (Springel 2010; Pakmor et al. 2016). We focus in particular on the flagship run MTNG740, which has been presented in detail in Pakmor et al. (2023). This realization of the MTNG simulations features a cubic periodic box with side length $500 h^{-1} \text{ Mpc}$ ($\approx 740 \text{ Mpc}$). The initial conditions were

generated at $z = 63$ and were sampled with 4320^3 dark matter particles of mass $1.7 \times 10^8 M_\odot$ and 4320^3 gas cells with an initial mass of $3.1 \times 10^7 M_\odot$, which is the target baryonic mass resolution¹. The gravitational softening length of collisionless particles (dark matter and stars) was set to $\epsilon_{\text{DM},*} = 3.7$ kpc, whereas the minimum (comoving) gravitational softening length for gas cells was set to $\epsilon_{\text{gas},\text{min}} = 370$ pc. The gravitational softening length of gas cells is adaptive, and it was scaled according to their size.

We also examine three additional full physics boxes to assess the influence of the simulated volume on our results. We indicate these additional runs with MTNG46, MTNG92, and MTNG185, respectively. They differ from the flagship simulation in terms of the simulated volume only, which is a periodic cubic box with a side length of $31.25 h^{-1}$ Mpc (≈ 46 Mpc) for the MTNG46 simulation, $62.5 h^{-1}$ Mpc (≈ 92 Mpc) for the MTNG92 realization, and $125 h^{-1}$ Mpc (≈ 185 Mpc) for the MTNG185 run. To maintain the same mass resolution as the flagship simulation, the number of particles was adjusted accordingly. Specifically, MTNG46 uses 2×270^3 dark matter particles and gas cells, MTNG92 employs 2×540^3 total particles, while MTNG185 contains 2×1080^3 resolution elements overall. This ensures that each dark matter particle and gas cell retains the same initial mass as in the flagship simulation. All other parameters were left unchanged.

Finally, we note that MTNG employs the cosmological parameters from [Planck Collaboration XIII \(2016\)](#). Specifically, the adopted values were $\sigma_8 = 0.8159$, $n = 0.9667$, $h = 0.6774$, $\Omega_{\text{m}} = 0.3089$, $\Omega_{\text{b}} = 0.0486$, and $\Omega_{\Lambda} = 0.6911$. This choice was made to facilitate the comparison of the simulation outcomes with those obtained in the IllustrisTNG suite, which is based on the same cosmology. Moreover, since our approach to modeling the GW signal in MTNG can be seamlessly applied to the original IllustrisTNG simulations (see Sect. 3.2), this also allows us to further assess the robustness of our results while factoring out the influence of cosmological parameter choice.

3. Methods

3.1. SEVN: Performing state-of-the-art binary population synthesis simulations

Modeling the GW signal in MTNG requires a catalog of compact binary mergers originating from stellar populations. To build this catalog, we used the archival dataset *SEVNbenchmark2401*² that was produced with the state-of-the-art binary population synthesis code SEVN³ ([Spera et al. 2015, 2019](#); [Mapelli et al. 2020](#); [Iorio et al. 2023](#)). The dataset contains the evolution of 2×10^7 stellar binary systems for each of 15 representative metallicities in the range $0.0001 \leq Z \leq 0.03$ ⁴, for a total of 3×10^8 binaries.

¹ AREPO implements a refinement scheme that keeps the mass of a gas cell within a factor of 2 from this reference value (m_{target}). Cells with mass $m > 2 \times m_{\text{target}}$ are split into two separate cells, whereas cells with mass $m < 0.5 \times m_{\text{target}}$ are dissolved into their neighbors. As star particles are generated from gas cells, their initial mass distribution are also in this range.

² The dataset is publicly available on Zenodo at the following link: <https://doi.org/10.5281/zenodo.16587145>.

³ Our analysis employed SEVN version 2.7.5 (commit [7bb74a89](#)). More information on SEVN can be found at <https://sevn.codes.gitlab.io/sevn/>.

⁴ Z denotes the mass fraction of elements heavier than helium. The specific values considered are $Z = 0.0001, 0.0002, 0.0004, 0.0006, 0.0008, 0.001, 0.002, 0.004, 0.006, 0.008, 0.01, 0.014, 0.017, 0.02, \text{ and } 0.03$.

SEVN models the stellar evolution by interpolating (in mass and metallicity) sets of precomputed stellar tracks for H-rich stars and pure-He stars (i.e., stars stripped by their H-rich envelope). Further information on the stellar models and interpolation schemes implemented in SEVN is provided in [Iorio et al. \(2023\)](#) and [Korb et al. \(2025\)](#). To produce the *SEVNbenchmark2401* dataset, the updated PARSEC stellar evolution models ([Bressan et al. 2012](#); [Chen et al. 2015](#); [Costa et al. 2019](#); [Nguyen et al. 2022](#)) as presented in [Costa et al. \(2025\)](#), were adopted among the available stellar tracks.

SEVN assumes that all stars with a final carbon–oxygen core mass exceeding $1.38 M_\odot$ undergo core collapse, forming a compact remnant, either a black hole or a neutron star. To model the core-collapse phase, the *DelayedGauNS* prescription (see Appendix A2 in [Iorio et al. 2023](#)) was used. This prescription is based on the delayed supernova model by [Fryer et al. \(2012\)](#) and produces no mass gap between neutron stars and black holes, in agreement with recent observational evidence (e.g., [Abbott et al. 2020](#); [Wyrzykowski & Mandel 2020](#); [Ray et al. 2025](#)). Compact remnants with masses above $3 M_\odot$ were classified as black holes, while those below this threshold were considered neutron stars. Black hole masses were further reduced to account for neutrino-driven mass loss, applying the correction proposed by [Zevin et al. \(2020, Equation \(13\) in Iorio et al. 2023\)](#). The [Fryer et al. \(2012\)](#) models are known to produce neutron star masses that do not reproduce the observed distribution in the Milky Way (see, e.g., [Giacobbo & Mapelli 2018](#); [Kruckow et al. 2018](#); [Vigna-Gómez et al. 2018](#)). Therefore, to generate the dataset, neutron star masses were resampled from a Gaussian distribution with a mean of $1.33 M_\odot$, a standard deviation of $0.09 M_\odot$, and lower and upper limits of $1.1 M_\odot$ and $3 M_\odot$, respectively, so that their mass distribution is consistent with Galactic observations ([Özel & Freire 2016](#)).

Toward the end of their evolution, massive stars can experience hydrodynamic instabilities triggered by the efficient production of electron-positron pairs. This process can lead either to pulsational pair-instability, characterized by repeated mass-loss episodes ([Yoshida et al. 2016](#); [Woosley 2017](#)), or to pair-instability supernovae, in which a single, extremely energetic pulse completely disrupts the star, leaving no remnant ([Heger et al. 2003](#); [Costa et al. 2021](#)). In SEVN these effects are accounted for by the fiducial *M20* model ([Mapelli et al. 2020](#); [Iorio et al. 2023](#)), which is based on hydrodynamic simulations of [Woosley \(2017\)](#). In this model, the fate of a massive star is determined by the mass of its helium core (m_{He}) at the onset of carbon burning. Stars with $32 M_\odot < m_{\text{He}} < 64 M_\odot$ undergo pulsational pair-instability, while those with $64 M_\odot < m_{\text{He}} < 135 M_\odot$ explode as pair-instability supernovae, leaving no remnant. Stars with even more massive helium cores collapse directly into black holes.

At the formation of a black hole or neutron star, the remnant receives a natal kick. These kicks were modeled following [Giacobbo & Mapelli \(2020\)](#). Specifically, the kick magnitude was drawn from a Maxwellian distribution with a root-mean-square velocity of 265 km s^{-1} ([Hobbs et al. 2005](#)), and then rescaled by a factor proportional to $m_{\text{ej}}/m_{\text{rem}}$, with m_{ej} indicating the mass of the ejecta and m_{rem} the remnant mass. The kick direction was randomly sampled assuming spherical symmetry. In the case of direct collapse ($m_{\text{ej}} = 0$) the resulting kick is zero. Within this framework, massive black holes and remnants formed from stripped-envelope stars typically receive smaller natal kicks. As a consequence of the kick, SEVN updates the orbital properties of the binary, including the possibility that the system becomes unbound. If the binary is disrupted,

SEVN continues to evolve the two stars as effectively single objects.

SEVN includes all the relevant binary processes (mass transfer, common envelope, stellar tides, GW decay) by implementing analytic and semi-analytic formalism to model the change on stellar (mass, radius, rotation) and binary (separation, eccentricity) properties. Mass transfer processes in SEVN include both accretion from stellar winds and Roche-lobe (RL) overflow. Wind mass transfer is always active, while RL overflow mass transfer is triggered when one of the two stars (the donor) expands beyond its RL (Eggleton 1983), transferring mass to its companion (the accretor). The adopted SEVN setup assumes that, during RL overflow mass transfer, 50% of the mass lost by the donor is accreted by the accretor. If the accretor is a compact object (i.e., a black hole or neutron star), accretion, whether from stellar winds or RL overflow, is limited by the Eddington accretion rate (see Eq. (17) in Iorio et al. 2023). Any mass lost by the donor that is not accreted leaves the system as a fast, isotropic wind from the donor (in case of the winds) or from the accretor (in case of RL overflow), carrying away angular momentum from the binary.

Certain binary configurations (short separations, high mass-ratios, donor with deep convective envelopes) can lead to mass transfer that is unstable on dynamical timescales. In SEVN, the stability of RL overflow mass transfer is assessed by comparing the donor-to-accretor mass ratio to a critical value, q_{crit} (Webbink 1985). Stars with deep convective envelopes, such as evolved giants, are more prone to unstable mass transfer and thus have lower q_{crit} values (typically $\lesssim 1$). In contrast, stars with radiative envelopes typically allow stable mass transfer, corresponding to higher q_{crit} values (up to $\gtrsim 3$).

The value of q_{crit} was estimated via the SEVN QCRS default prescription (see Iorio et al. 2023). This is equivalent to the q_{crit} implementation in Hurley et al. (2002), with the modification that main sequence and Hertzsprung gap donors are always assumed to undergo stable mass transfer. If the mass transfer is unstable, the binary enters a common envelope (CE) phase (Ivanova et al. 2013). The CE phase is treated using the energy formalism (Webbink 1984), in which the outcome, envelope ejection and orbital shrinking or merger, is determined by comparing the orbital energy released during the inspiral to the binding energy of the donor’s envelope. A CE efficiency parameter of $\alpha = 1$ was employed, corresponding to the assumption that 100% of the available orbital energy is used to unbind the envelope. The envelope binding energy was computed using the prescription in Claeys et al. (2014).

SEVN models stellar tides using the weak friction approximation developed by Hut (1981), as implemented in Hurley et al. (2002). Tidal effects were active throughout all phases of the simulation. If tides were not strong enough to circularize the orbit by the onset of RL overflow, circularization at periastron was enforced by setting the new orbital semimajor axis equal to the periastron distance.

To produce the *SEVNbenchmark2401* dataset, the initial properties of the binary population were selected in the same way as in Korb et al. (2025, see their Appendix A.2). Primary masses were drawn from a Kroupa (2001) initial mass function ($\propto m^{-2.35}$) in the range $5 < m_1/M_{\odot} < 150$. Secondary masses were assigned using a mass ratio $q = m_2/m_1$ drawn from a power-law distribution ($\propto q^{-0.1}$), with a minimum secondary mass of $2.2 M_{\odot}$. Orbital periods were sampled from a distribution $\propto (\log P)^{-0.55}$ over the range 1.4 days to 866 years, and eccentricities from a distribution $\propto e^{-0.42}$ in the interval [0,

0.9], with a correction for the eccentricity–period relation as described in Moe & Di Stefano (2017 their Eq. (3)).

This setup efficiently sampled the rare population of compact-object binaries hosting black holes and neutron stars (see, e.g., Iorio et al. 2023; Sgalletta et al. 2023, 2025; Korb et al. 2025). All binary systems were evolved until both stars reached the end of their evolution, either forming compact remnants (white dwarfs, neutron stars, or black holes), or leaving no compact remnants due to stellar (e.g., pair-instability supernovae) or binary (e.g., mergers) processes. We refer the reader to Iorio et al. (2023) for a comprehensive description of SEVN, including alternative prescriptions and their impact on the population of binary compact objects.

From the *SEVNbenchmark2401* dataset, we extracted all binary compact objects that merge within 14 Gyr, and we stored the relevant properties required to model the GW events in MTNG (see Sect. 3.2). These include: a unique identifier for each merger event, the time of merger, measured from the birth of the binary system, the individual masses of the merging objects, and the mass of the merger remnant. Given that the MTNG simulations adopt a Chabrier (2003) initial mass function with a maximum stellar mass of $M_{\text{max}} = 100 M_{\odot}$, we excluded from the selection all systems whose progenitors include a primary star exceeding this mass limit. For each metallicity, we produced three separate tables, corresponding to the three GW progenitor channels considered in our analysis: BBH, BHNS, and BNS mergers. The events in each table were sorted in ascending order of merger time, to facilitate the incorporation of compact binary mergers into the MTNG cosmological simulations, as described below.

3.2. AREPO-GW: Building GW catalogs in galaxy formation simulations

To predict the distribution of GW events from the evolution of stellar populations modeled within full-physics galaxy formation simulations, we developed AREPO-GW, a flexible and parallel module that is fully integrated into the AREPO code. Our module generates the distribution of GW events by stochastically sampling mergers of compact binary systems – that can be thought of as progenitors of GW emission – occurring within a simple stellar population (SSP). This sampling relies on the merger tables produced by the SEVN code (described in Sect. 3.1) and follows an approach similar to that adopted in previous studies (see, e.g., Mapelli et al. 2018; Artale et al. 2019).

Specifically, AREPO-GW builds upon the stellar evolution infrastructure already implemented in the galaxy formation physics model of the MTNG suite. In this framework, each star particle in a simulation is treated as a SSP following a Chabrier (2003) initial mass function and is characterized by its age and metallicity. Based on the metallicity value of a given star particle, we selected the table generated from SEVN simulations that most closely approximates its metal content. Once a table was selected, we evaluated – for each of the three GW progenitor channels – the probability p^* that a merger event occurs as

$$p^* = \frac{M_* f_{\text{bin}} f_{\text{IMF}}}{M_{\text{sim}}}, \quad (1)$$

where $M_{\text{sim}} \approx 4.1 \times 10^8 M_{\odot}$ is the total mass of simulated binary systems at the selected metallicity and for the considered GW progenitor channel (excluding systems with stars more massive than $100 M_{\odot}$, see Sect. 3.1); M_* is the birth mass of the star particle; $f_{\text{bin}} = 0.628$ is the expected binary fraction

(Moe & Di Stefano 2017)⁵; and $f_{\text{IMF}} = 0.243$ is a correction factor accounting for the incomplete sampling of the initial mass function due to the lower mass limits applied to primary and secondary stars in the sampled binaries ($5 M_{\odot}$ and $2.2 M_{\odot}$, respectively)⁶. The first recorded event was then considered. If the current age of the stellar particle was greater than the recorded merger time – or in other words if enough time had passed since the birth of that stellar population to have a potential binary merger – we extracted a uniformly distributed random number $p \in [0, 1)$. If $p < p^*$, we stored key binary merger event properties, namely its position within the simulation domain, time at which the event occurred, unique identifier (so that other properties of the merger could be retrieved from the SEVN tables for further analysis), merger type (BBH, BHNS, BNS), and the unique identifier of the star particle in the simulations to which the merger is associated. Otherwise, we moved on to the next record in the table. We stopped processing the star particle when its age was less than the merger time of the record being examined. This procedure was repeated for each star particle in the simulation and iterated over all time steps. For each star particle, merger events that had already been examined in previous time steps were excluded from the list of potential events.

We implement two versions of AREPO-GW. The first, which has been described above, works on-the-fly while a simulation is progressing. The other version can be used to post-process a snapshot of an existing large-scale simulation and it becomes particularly useful when dealing with computationally expensive simulations producing a large data volume that cannot be easily rerun to include the GW signal. This is exactly the case that we are exploring in this work. In the post-processing mode, AREPO-GW evolves each star particle of the simulation on a timestep of a predetermined length (that can be selected by the user) from its birth to the time at which the snapshot is produced. Therefore, to have a complete catalog of GW progenitors event, it is in principle sufficient to analyze the snapshot produced by the simulation at the lowest available redshift.

To allow the analysis of very large snapshots on relatively small machines, we designed the post-processing mode of AREPO-GW such that a snapshot could be processed one or several chunks at a time. In this way, we were able to produce the GW catalog of the flagship MTNG740 simulation analyzed in this work starting from its redshift-zero snapshot (≈ 7 TB of data volume) on our local HPC machine. Moreover, for such large-scale applications we adjusted the output strategy to minimize the total volume of saved data. First, we processed star particles in the same order as they appeared in the snapshot, allowing the storage of their unique identifiers to be optionally disabled. For each individual star particle, we then saved the associated merger events contiguously and record their total number. In this way we were able to uniquely associate each binary merger event with its parent star particle in the simulation. Second, we did not store the binary merger positions as they are identical for all the merger events generated by the same star particle. However, the position of the parent star particle can still be retrieved at the time (or redshift) of a specific merger event via snapshot interpolation (if enough snapshots exist) or if light-cone-like outputs are available. Despite these optimizations, which lead to significant

storage saving, the total volume of the GW catalog produced for the MTNG740 simulation still amounts to ≈ 30 TB.

Both operational modes of AREPO-GW introduce only minimal computational overhead because the module is fully embedded within the stellar evolution routines already implemented in AREPO (see Vogelsberger et al. 2013, for details). The generation of GW events requires only a table lookup and a single random draw per event, along with a lightweight output buffer. As a result, any cosmological simulation that can be run with AREPO can also be run with the on-the-fly mode of AREPO-GW. The primary cost difference between the two modes – in addition to the time requested to carry out a specific simulation when the algorithm is used on-the-fly – is the storage requested per event: the on-the-fly mode uses 30 bytes to store a single event compared to 8 bytes needed in post-processing. Since the number of GW progenitor events scales directly with the total stellar mass formed in the simulation (see Eq. (1)), it is largely the simulation volume (and to a lesser extent its resolution) that determines the size of the resulting GW catalog. However, the larger storage footprint of the on-the-fly mode rarely poses a practical limitation to the applicability of the method. For example, using MTNG-like physics and resolution would produce an event catalog of an estimated size of ≈ 30 TB for a simulation with a volume of $(480 \text{ Mpc})^3$ run with AREPO-GW on-the-fly. Simulations of this size are state-of-the-art and comparable to the MTNG740 flagship box analyzed in this work. The on-the-fly mode also has the advantage that precise spatial positions of GW-progenitor events are recorded at formation, eliminating the need for snapshot or light-cone interpolation. This makes the on-the-fly approach particularly well-suited for studies investigating the impact of different galaxy formation physics models, where rerunning the underlying hydrodynamic simulations is required in any case.

To first order, the only difference between the post-processing and on-the-fly modes of AREPO-GW is that, as previously discussed, in the former it is not possible to recover the exact positions of the merger events associated with the same star particle. However, upon closer inspection, one cannot expect an identical sequence of merger events for a given star particle in the two modes, even when analyzing the same simulation. This discrepancy is caused by two main effects. First, the time step over which a star particle evolves is fixed in post-processing but adaptive in the on-the-fly mode. Although both approaches rely on the same power-of-two time-step hierarchy (see Springel 2010, for more details), this can still introduce small variations in the recorded merger times. Second, our modeling is inherently stochastic: SEVN tables were sampled with a probability p^* , defined in Eq. (1), which was then compared to a uniformly distributed random number. Both operational modes of AREPO-GW parallelize execution through the Message Passing Interface (MPI) library and random numbers were generated independently by each MPI task. Consequently, the precise sequence of merger events also depends on the MPI task processing the parent star particle. In the on-the-fly mode, this picture becomes more complex because (star) particles can be exchanged among MPI tasks during runtime to maintain optimal workload balance in AREPO (Springel 2010). In practice, however, provided that merger events are sufficiently well sampled, stochasticity should only introduce negligible differences between the two modes. We present a more quantitative assessment of how the two operational modes of AREPO-GW affect the properties of GW progenitor events within the same simulation in Appendix A.

We would like to stress that the approach to generating GW catalogs described in this section has a broader scope of

⁵ The binary fraction in Moe & Di Stefano (2017) depends on the mass of the primary. The reported value was obtained by considering the whole initial mass function.

⁶ To estimate the correction factor, we assumed a Chabrier (2003) initial mass function ranging from 0.01 to $100 M_{\odot}$, and a mass ratio distribution with q between 0.1 and 1 .

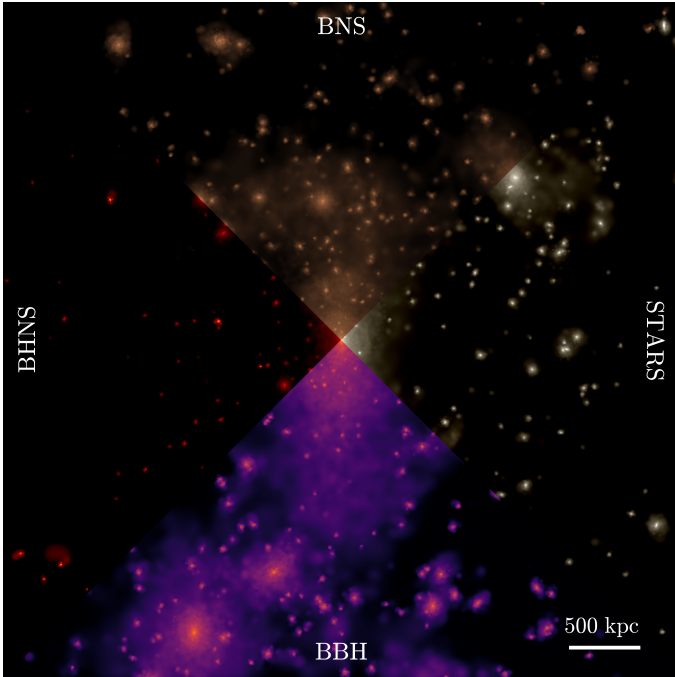


Fig. 1. Maps of projected stellar surface brightness and GW progenitor merger surface density events for the BBH, BHNS, and BNS channels (clockwise from the right, as indicated in the figure). The maps are centered on the most massive FoF halo of the MTNG740 run and have a side length of 5 Mpc. The stellar map was obtained by mapping the g , r , and z photometric bands to the blue, green, and red channels of the figure, respectively, with intensities linearly scaled over the interval $[20, 30]$ mag arcsec $^{-2}$. For the GW progenitor merger rate events, only compact binary mergers occurring at $z < 0.1$ are considered. The values are logarithmically scaled over the interval $[10^{-8}, 10^{-2}]$ pc $^{-2}$.

applicability than the MTNG simulation suite. Indeed, although the algorithm we implemented depends on the stellar evolution routines adopted in the MTNG simulations, such routines are at the foundations of several galaxy formation physics models implemented in AREPO that are targeted to study different aspects of structure formation. These include, among others, the Illustris and IllustrisTNG models (Vogelsberger et al. 2014; Pillepich et al. 2018a), on which the MTNG simulations are based and that focus on formation and evolution of a complete galaxy population, the Auriga project (Grand et al. 2017), which is centered on the formation of galaxies such as the Milky Way, and the SMUGGLE model that includes more realistic prescriptions for ISM structure and stellar feedback (Marinacci et al. 2019). Moreover, our approach can be easily ported to other hydrodynamic simulation codes or used to post-process calculations not performed with AREPO given the limited amount of information on star particles (e.g., initial mass, birth time, metallicity, and assumed IMF) necessary as an input.

4. Results

4.1. The spatial distribution of GW progenitors events

To visually demonstrate the capabilities of the AREPO-GW post-processing module, Fig. 1 presents maps of projected stellar surface brightness and the surface density of GW progenitor mergers for the three different channels – BBH, BHNS, and BNS – shown clockwise from the right. The maps are centered on the most massive friends-of-friends (FoF) structure in the

MTNG740 volume and span a side length of 5 Mpc. The stellar map is a grz composite image, with photometry derived from the ages and metallicities of star particles using the spectral evolution model of Bruzual & Charlot (2003). The maps of GW progenitor merger rates include only compact binary mergers occurring at redshifts $z < 0.1$ and are scaled to the same range of surface density values to highlight relative differences between the channels.

It is apparent that the rates of GW progenitors are closely linked to the stellar distribution of the forming structures in the simulation. Specifically, the merger events tend to cluster around regions of high stellar surface brightness, which correspond to substructures – namely, individual galaxies. We note that the structure shown in Fig. 1 is representative of a galaxy cluster with a virial mass⁷ of $M_{200,c} \approx 8.5 \times 10^{14} M_{\odot}$. However, a noticeable difference emerges in the spatial distribution of the GW progenitor surface densities across the different channels. BBH and BNS events feature more diffuse distributions, while BHNS mergers are more concentrated near peaks in stellar surface brightness. This behavior reflects the lower rates of BHNS mergers at low redshift in our model, which are a factor of ~ 2 – 4 lower than the corresponding BBH and BNS rates (see Fig. 3). Although the $z < 0.1$ merger rates of BBH and BNS are within a factor of ≈ 2 , the BBH merger surface density appears more spatially extended than that of BNS, despite the higher overall occurrence of BNS events. This suggests a possible difference in the delay time distributions of their progenitors, which may influence the spatial distribution of merger sites. Further details on distribution of delay times of the different GW progenitors are provided in Fig. 5.

4.2. Evolution of GW event rates

An important quantity that can be derived from the modeling of GW sources in cosmological simulations is the time evolution of the comoving merger rate of compact binary systems across different channels. Given the approach adopted here, we expect the evolution of the comoving merger rate to be closely tied to the evolution of the cosmic star formation rate density (SFRD). Specifically, the comoving merger rate for a given channel can be expressed as

$$R_{\text{ch}}(t) = \int_0^t \text{SFRD}(t') \Theta_{\text{ch}}(t - t') dt', \quad (2)$$

where $R_{\text{ch}}(t)$ is the comoving merger rate at cosmic time t for a specific progenitor channel (BBH, BHNS, BNS), $\text{SFRD}(t')$ is the comoving SFRD at t' , and $\Theta_{\text{ch}}(t - t')$ is related to both the merger efficiency of the progenitors per unit of stellar mass formed (Fig. 4) and their delay time distribution⁸ (Fig. 5). In our modeling, the SFRD is directly extracted from the hydrodynamic simulation, while Θ_{ch} is encoded in the compact binary merger tables described in Sect. 3.1.

Figure 2 shows the SFRD obtained from the MTNG suite (solid colored lines), compared with two widely used determinations: Madau & Dickinson (2014, dashed line) and Driver et al. (2018, points with error bars). Both SFRD determinations have been rescaled to the value of the Hubble parameter adopted in the

⁷ We define the virial mass of a halo as the total mass enclosed within a radius where the mean density is 200 times the critical density of the Universe.

⁸ In Eq. (2) the delay time is given by $t - t'$, where t indicates the time at which a compact binary merger occurs and t' is the formation time of the associated stellar population.

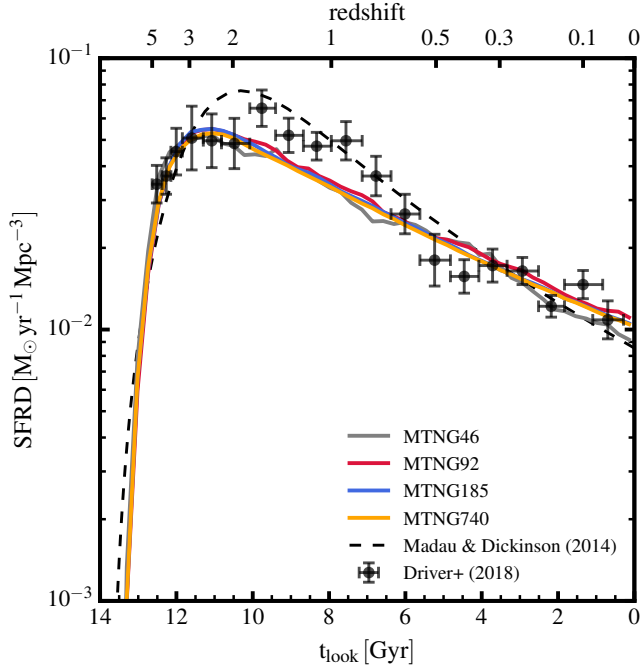


Fig. 2. Comoving SFRD as a function of time and redshift in four different boxes of the MTNG simulations, as indicated in the legend. The determinations of the cosmic SFRD as reported in [Madau & Dickinson \(2014\)](#) and [Driver et al. \(2018\)](#) are also shown for comparison. Owing to the constant mass resolution of the analyzed MTNG boxes, the SFRD is well-converged, with negligible differences among the runs. The MTNG SFRD peaks at earlier redshifts ($z \sim 3$, rather than $z \approx 2$) and shows a less steep decline toward lower redshift compared to the results presented in [Madau & Dickinson \(2014\)](#). Overall, the SFRD values in MTNG are broadly consistent with the determinations of [Madau & Dickinson \(2014\)](#) and [Driver et al. \(2018\)](#), although they tend to be lower by up to a factor of $\lesssim 2$ in the redshift range $z \sim 1-2$.

MTNG cosmology, by multiplying them by the factors $(h/0.7)^3$ and $(h/0.7)$, respectively. The [Madau & Dickinson \(2014\)](#) measurements are further corrected for the different choice of the IMF ([Salpeter 1955](#) vs. [Chabrier 2003](#)), applying a conversion factor of 0.63 (see Table 1 in [Driver et al. 2013](#)). The SFRD of the MTNG740 run was estimated by binning stellar particles according to their ages into 50 bins in the range 0–14 Gyr and summing for each bin their initial mass⁹. The total stellar mass of each bin was then divided by the bin width (0.28 Gyr) and the comoving volume of the box to get the SFRD value.

The most notable feature of the SFRD derived from different MTNG realizations is its independence from the simulated volume. In other words, the SFRD is well-converged, with negligible differences among all runs. This is due to the fact that the mass resolution was kept constant across the different simulation boxes. This convergence suggests that even the smallest simulated volume in the MTNG suite provides a reasonably representative view of the cosmic star formation history. The overall shape of the SFRD is similar in both the MTNG runs and the observational determinations, featuring a peak in the redshift range $z \approx 2-3$ and a general decline toward lower redshift. The SFRD values in MTNG are also broadly consistent with [Madau & Dickinson \(2014\)](#) and [Driver et al. \(2018\)](#) estimates, though differences by up to a factor of $\lesssim 2$ appear in the redshift

⁹ In the MTNG galaxy formation framework, star particles lose mass as a consequence of stellar evolution processes.

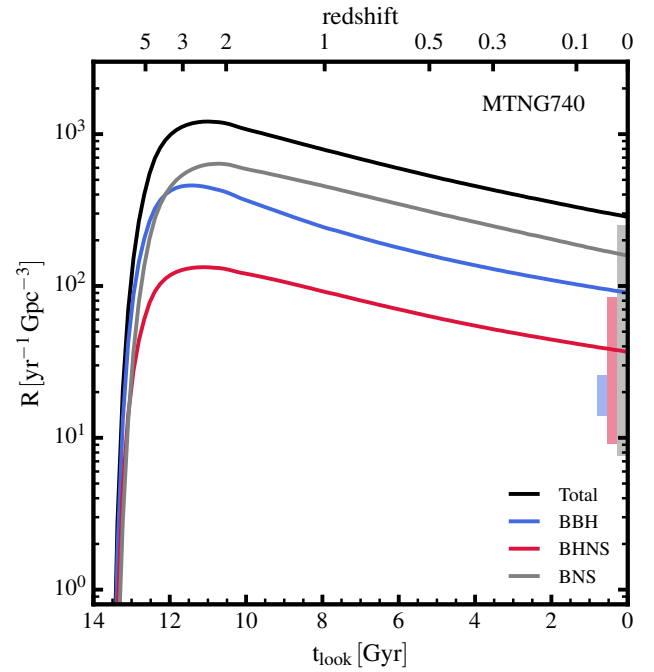


Fig. 3. Rate of GW progenitor events per comoving volume as a function of lookback time and redshift in the simulation MTNG740. The total progenitor merger rate and the contribution of different compact binary merger channels are displayed as indicated in the legend. Colored bands represent constraints on the low-redshift ($z \approx 0$) merger rate of the different channels, as inferred from GWTC-4 high-confidence events ([LVK25](#)).

range $z \sim 1-2$. However, there are important differences worth noting. First, the position of the SFRD peak differs: it occurs at lower redshift ($z \approx 2$) in the [Madau & Dickinson \(2014\)](#) determination, whereas the MTNG suite favors a slightly higher redshift value – the position of the peak is not well-defined in the [Driver et al. \(2018\)](#) measurements. In addition, the cosmic SFRD in [Madau & Dickinson \(2014\)](#) exhibits a steeper decline toward low redshift relative to both MTNG and [Driver et al. \(2018\)](#). A least-squares fit of the form $\text{SFRD} \propto (1+z)^{\kappa_{\text{SFRD}}}$ for $z < 1$ yields $\kappa_{\text{SFRD}} = 1.74 \pm 0.02$ for MTNG, $\kappa_{\text{SFRD}} = 2.26 \pm 0.56$ for [Driver et al. \(2018\)](#), and $\kappa_{\text{SFRD}} = 2.58 \pm 0.02$ for [Madau & Dickinson \(2014\)](#), with errors reported as 90% confidence intervals on the slope. This quantitative comparison highlights a tension in the low-redshift SFRD evolution between MTNG and the [Madau & Dickinson \(2014\)](#) results, while showing better agreement with the values from [Driver et al. \(2018\)](#). These differences in redshift evolution may significantly affect the modeling of GW event rates in simulations and impact the interpretation of their evolution based on observational constraints.

Figure 3 shows the rate of GW progenitor events per comoving volume as a function of time for the different compact binary merger channels (solid colored lines) and the total rate (solid black line) in the flagship MTNG740 simulation. These quantities were obtained with the same binning procedure used to compute the SFRD presented in Fig. 2. Also shown are local ($z \approx 0$) rate estimates from the high-confidence events in the GWTC-4 catalog of the LVK collaboration ([LVK25](#), colored boxes). Thanks to the consistent mass resolution across simulation boxes, the results from other MTNG realizations are essentially unchanged and are presented in Fig. B.1 in Appendix B.

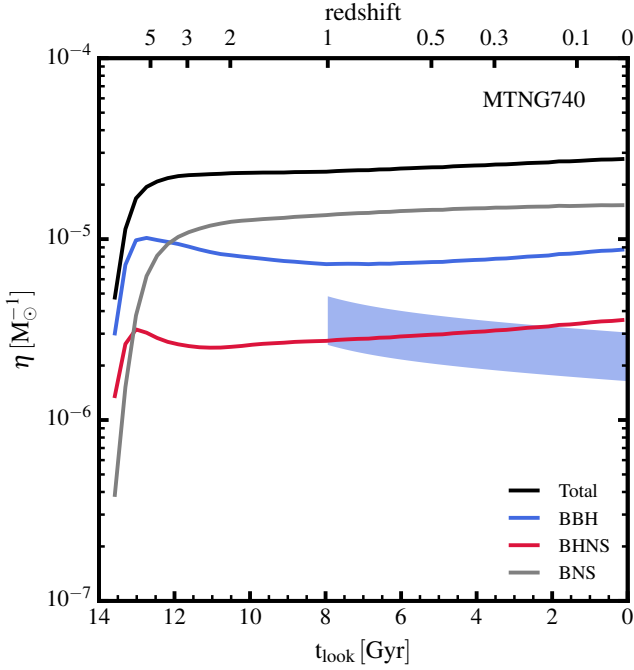


Fig. 4. Merger efficiency of compact binary objects per unit of stellar mass formed as a function of time and redshift in the simulation MTNG740. The total merger efficiency, along with the contributions from different merger channels, is displayed as indicated in the legend. The shaded region represents constraints on the evolution of BBH merger efficiency at low redshifts ($z \lesssim 1$) derived by extrapolating the LVK collaboration limits shown in Fig. 3. This extrapolation assumes a redshift evolution of the form $(1+z)^{3.2}$ (LVK25) and adopts the SFRD evolution given by Madau & Dickinson (2014), as reported in Fig. 2.

We first note that the overall shape of the comoving GW progenitor rates across all three channels is very similar to that of the SFRD. The main difference is in the conversion efficiency of the SFRD into event rates, which is higher for the BNS channel and lower for the BHNS channel. This consistency in shape is expected given the close relationship between the SFRD and the comoving event rate described in Eq. (2). However, upon closer inspection, some notable differences emerge. Beyond the variation in the efficiency, it is evident that the peak of the BBH and BHNS rates occurs at earlier redshifts ($z \approx 3$). In contrast, the BNS rate peaks at a slightly lower redshift, closer to $z \approx 2$. Both values broadly align with the peak of the SFRD, although the latter lies approximately in the middle of this redshift interval. We speculate that this shift is related to the delay time distributions of the different formation channels. Indeed, the vast majority of progenitor stars in BNS systems typically have very short delay times ($t_{\text{delay}} \lesssim 0.1$ Gyr), whereas BBH and BHNS show a nonnegligible tail extending toward longer delay times (see Fig. 5 and the related discussion below). Another important difference lies in the predicted event rates at low redshift. In particular, the local BHNS and BNS rates from MTNG740 are consistent with current LVK observational constraints – though these constraints remain relatively weak. In contrast, the BBH channel shows a significant discrepancy: our SEVN-based modeling overestimates the local BBH rate by a factor of ~ 4.5 . This discrepancy is somewhat reduced – although still present with the excess reduced to a factor of ~ 2.4 – when the constraints from the previous version of the catalog (GWTC-3; Abbott et al. 2023b) are adopted. The overprediction of the local rate of BBH

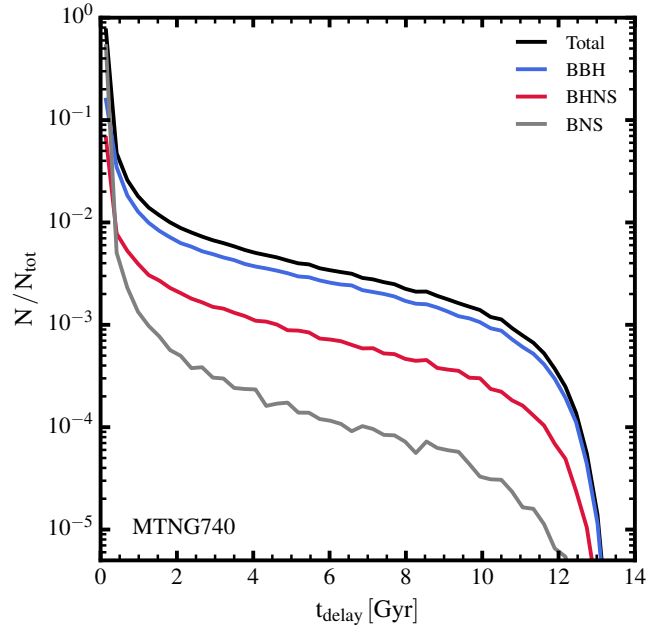


Fig. 5. Delay time distribution of GW progenitor events in the simulation MTNG740. The delay time distribution of the total merger events, along with the contributions from different compact binary merger channels, is displayed as indicated in the legend. All distributions are normalized to the total number of merger events (i.e., $N_{\text{tot}} = \text{BBH} + \text{BHNS} + \text{BNS}$), so that each delay time bin represents the fractional contribution of that specific delay time to the overall number of events.

mergers is a known issue with current state-of-the-art binary population synthesis models, particularly when used in combination with more accurate modeling of the metal-dependent cosmic SFRD, and arises from uncertainties in massive binary star evolution (Boesky et al. 2024; Sgalletta et al. 2025). In addition, part of the excess in BBH mergers predicted by such models may result from an overestimation of the formation rate of merging BBHs through CE evolution (see, e.g., Marchant et al. 2021).

Finally, it is worth pointing out that, as with the SFRD, one can characterize the decline toward low redshift of the comoving GW event rate for each channel by measuring its slope as a function of redshift for $z < 1$. Performing a least-squares fit in each case yields $\kappa_{\text{BBH}} = 1.44 \pm 0.01$, $\kappa_{\text{BHNS}} = 1.34 \pm 0.01$, and $\kappa_{\text{BNS}} = 1.55 \pm 0.02$, with errors reported as 90% confidence intervals on the slope. In all cases, this decline is shallower than that of the SFRD. This indicates that the merger efficiency, defined as the number of merger events per unit stellar mass formed, tends to increase slightly with time (see also Fig. 4). In the case of BBH, we can compare this redshift evolution to observational estimates from LVK25, who report a slope of $\kappa_{\text{BBH}} = 3.2^{+0.94}_{-1.00}$ for $z \lesssim 1$. Although the uncertainty on the slope is appreciable, it is evident that the observational data appear to favor a more rapidly declining trend with redshift than the one predicted by our model.

Figure 4 presents the binary merger efficiency per unit stellar mass formed as a function of time for the total GW progenitor events (solid black line) along with individual contributions of the different compact binary merger channels in the MTNG740 box (solid colored lines). The merger efficiency is computed as the ratio between the comoving rate of GW progenitor events (Fig. 3) and the comoving SFRD (Fig. 2), thus representing the conversion factor at any given cosmic time t between these two

quantities. We also show, as a shaded region, the observational constraints on the BBH merger efficiency at $z \lesssim 1$ obtained by extrapolating the limits displayed in Fig. 3 to higher redshift. This procedure assumes a redshift evolution of the BBH rates $\propto (1+z)^{3.2}$ LVK25 and adopts the Madau & Dickinson (2014) SFRD reported in Fig. 2. Results for all the MTNG boxes analyzed in this work are presented in Fig. B.2 in Appendix B. We anticipate that also in this case, owing to the consistent mass resolution across boxes, the results from other MTNG realizations are essentially identical.

All progenitor channels feature a merger efficiency evolution characterized by an initial phase of rapid increase, followed by an approximately constant behavior for $z \lesssim 5$. Below this redshift, variations in efficiency remain within a factor of 2 for all channels over roughly a Hubble time. This stability further highlights the close connection – expressed in Eq. (2) – between star formation and the production of binary mergers responsible for GW signal in the Universe. Upon closer inspection, some differences in the evolution of merger efficiency among the various channels become apparent. First, the efficiency values differ, reflecting the relative rates of GW progenitor events across channels: BNS mergers exhibit the highest overall rate, followed by BBH and BHNS mergers. A second notable difference lies in the shape of the efficiency evolution. While the BNS efficiency increases monotonically – albeit slowly – over time, the BBH and BHNS efficiencies display non-monotonic trend. Specifically, BBH efficiency declines from a local peak at $z \approx 5$ to a minimum at $z \approx 1$, while BHNS efficiency reaches a minimum at $z \approx 3$ before both resume a slow increase toward $z = 0$. These distinct trends are due to a combination of the delay time distribution of merger events (Fig. 5) and the dependence of event rates on the metallicity of the underlying stellar populations (see Fig. 6 for more details).

Owing to the current lack of robust observational constraints on the redshift evolution of BHNS and BNS merger rates, it is only possible to compare the efficiency evolution of BBH mergers to observational estimates. The most recent LVK observations (LVK25), when combined with the SFRD from Madau & Dickinson (2014), appear to favor a slowly decreasing efficiency over time for this merger channel. However, it is important to note the considerable uncertainty in the slope of the redshift evolution of the BBH merger rate, $\kappa_{\text{BBH}} = 3.2^{+0.94}_{-1.00}$, which implies that the trend predicted by the MTNG results remains somewhat consistent with current observational constraints. Nevertheless, the efficiency values derived from the observations are systematically lower, by a factor of approximately 2.9–5.3 at $z = 0$, compared to our results. This tension is slightly enhanced with respect to earlier LVK constraints obtained from the GWTC-3 catalog (Abbott et al. 2023b) at $z = 0$, where efficiency values were smaller by factors ranging from 1.2 to 4.7 relative to MTNG predictions.

Figure 5 shows the delay time distribution of GW progenitor events in the MTNG740 simulation, both for the total number of events (solid black line) and for the individual compact binary merger channels (solid colored lines). The distributions were computed using the same binning procedure as in Figs. 2 and 3, and were normalized to the total number of events. As a result, each bin represents the fractional contribution of a specific delay time and merger channel to the overall event population. The shape of the delay time distribution is broadly similar across all progenitor channels. Most events are concentrated at short delay times ($t_{\text{delay}} \lesssim 0.5$ Gyr), with a tail extending

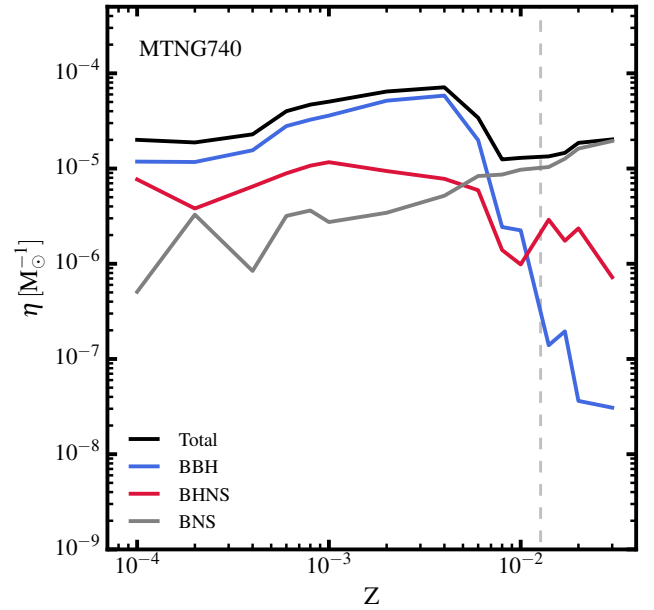


Fig. 6. Merger efficiency of compact binary objects per unit of stellar mass formed as a function of the tabulated metallicity of binary systems simulated with the SEVN code in the MTNG740 simulation. The total merger efficiency, along with the contributions from different merger channels, is displayed as indicated in the legend. The vertical dashed line indicates the solar metallicity value adopted in MTNG, $Z_{\odot} = 0.0127$.

to $t_{\text{delay}} \sim 12$ Gyr before sharply cutting off. Among the three channels, BBH mergers exhibit the most prominent long-delay tail, followed by BHNS and BNS mergers. This extended tail accounts for only a negligible fraction of BNS events beyond $t_{\text{delay}} > 0.5$ Gyr, whereas for BBH and BHNS events, the corresponding fraction rises to $\gtrsim 40\%$. Moreover, the median delay time for BBH mergers is approximately 0.284 Gyr. In comparison, slightly more than 50% of BHNS events occur within 0.14 Gyr, while for BNS mergers this fraction increases to nearly 96%. Similar results are obtained for all the MTNG boxes analyzed in this work and are reported in Fig. B.3 in Appendix B.

These differences in the delay time distribution may help explain the broader spatial extent of BBH mergers relative to BNS events, as shown in Fig. 1. The vast majority of BNS mergers occur within a very short time span ($\lesssim 0.1$ Gyr), making them more tightly linked to recent star formation. Since star formation is typically concentrated in high-density regions near halo centers, BNS mergers are correspondingly more localized. On the other hand, the longer delay times of BBH mergers allow their progenitors to move farther away from their birth sites, resulting in a more extended spatial distribution. The shorter delay times for BNS systems arise from the requirement that they undergo at least one CE phase, which efficiently reduces their orbital separation and leads to faster mergers. In contrast, a fraction of BHNS and BBH systems merge via stable mass transfer on longer timescales (see Fig. 14 in Iorio et al. 2023).

The presence of a more extended tail in the delay time distribution for the BBH and BHNS channels, combined with the declining number of merger events per unit stellar mass formed at metallicities $Z \gtrsim 4 \times 10^{-3}$ (see Fig. 6), may shed light on the non-monotonic evolution of the binary merger efficiency presented in Fig. 4. Specifically, the decline in binary merger efficiency for these channels is partially offset by the long delay time

events, which contribute to a recovery after the minima observed at $z \approx 3$ (for BHNS) and $z \approx 1$ (for BBH). In contrast, the monotonic increase in the efficiency of the BNS channel can be fully accounted for by the rising number of BNS merger events with increasing metallicity, coupled with the fact that almost all BNS mergers occur within ~ 0.1 Gyr after the formation of their parent stellar population.

4.3. Metallicity dependence of GW progenitor events

As formalized in Eq. (2), the rates of GW progenitor events can be expressed as a convolution between the SFRD and a channel-dependent factor that encapsulates both the evolving binary merger efficiency per unit stellar mass formed and the delay time distribution of the events. However, Eq. (2) does not explicitly show the dependence of these quantities on metallicity. In fact, it represents a metallicity-integrated version of the underlying components. The aim of this section is to analyze this metallicity dependence and examine how it affects the contributions of different stellar generations to the GW event rates across the BBH, BHNS, and BNS progenitor channels.

In Fig. 6, we present the merger efficiency of GW progenitor events per unit of stellar mass formed in the MTNG740 simulation as a function metallicity. The metallicity is binned such that the GW events are assigned to a grid of the binary population synthesis models computed with SEVN, selecting the table that most closely matches the metallicity of their parent star particle. This is the same table used to stochastically sample the GW event rate for that particle (see Sect. 3.2). The figure displays both the total efficiency (solid black line) together with contributions from individual compact binary merger channels (solid colored lines). In all cases, the efficiency is defined as the ratio of the number of GW progenitor events at a given metallicity to the total stellar mass formed at that metallicity. The vertical dashed line is located at $Z = Z_{\odot}$, with $Z_{\odot} = 0.0127$ in MTNG (Asplund et al. 2009). Results from the other MTNG simulations analyzed in this work are essentially unchanged with respect the flagship box and are reported in Appendix B (Fig. B.4).

The overall trend of the merger efficiency as a function of metallicity begins at $\eta \sim 2 \times 10^{-5} M_{\odot}^{-1}$ at the lowest tabulated metallicity and steadily increases to a peak of $\eta \sim 6 \times 10^{-5} M_{\odot}^{-1}$ at $Z \approx 4 \times 10^{-3}$. Beyond this point, the efficiency declines, reaching a minimum of $\eta \approx 10^{-5} M_{\odot}^{-1}$ at $Z = 10^{-2}$, followed by a slight rise – by approximately a factor of 2 – at the highest metallicity values considered. This non-monotonic behavior is driven by the varying contributions of different merger channels. At low metallicities ($Z < 4 \times 10^{-3}$), the BBH channel dominates, contributing roughly an order of magnitude more than the BNS channel and a factor of 3–5 more than the BHNS channel. The decline in the total efficiency at $Z > 4 \times 10^{-3}$ is primarily due to a sharp drop in BBH merger efficiency, from $\eta \sim 4 \times 10^{-5} M_{\odot}^{-1}$ to $\eta \sim 2 \times 10^{-8} M_{\odot}^{-1}$ at $Z = 0.03$. This decrease is partially offset at high metallicities by a steady rise in BNS merger efficiency, which increases by nearly a factor of 30 across the full metallicity range. The BHNS channel exhibits a qualitatively similar trend to BBH, although with generally lower efficiency and a less pronounced decline at high metallicities (approximately a factor of 6 across the range $7 \times 10^{-3} < Z < 10^{-2}$). Since younger stellar populations tend to be more metal-rich – because they form from gas that has been enriched by previous generations of stars – we expect the BBH and BHNS channels, and particularly the former, to exhibit a pronounced decline in their comoving merger rates at late times. In contrast, due to the increasing

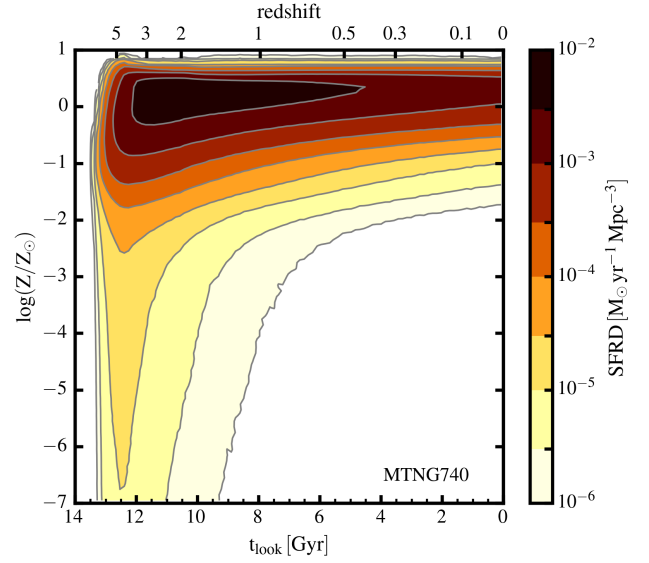


Fig. 7. 2D distribution of metallicity and (lookback) birth times of stellar particles in the MTNG740 simulation. The color scale indicates the corresponding (comoving) SFRD. Contours are placed at $[1, 3, 10, 30, 100, 300, 1000, 3000, 10^4] \times 10^{-6} M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$ to help visualize the contributions of metallicity and stellar ages to the overall SFRD of the simulation.

merger efficiency of BNS systems with metallicity, the decline in their comoving merger rate at late times is primarily driven by the decreasing SFRD beyond $z \sim 2$.

However, this simple picture is made more complex by two key points: (i) as discussed in relation to Fig. 5, GW progenitor events do not occur instantaneously after the birth of their parent stellar population, but follow a distribution in delay times that can be rather extended – for instance, in MTNG about 40% of BBH and BHNS mergers occur for $t_{\text{delay}} \gtrsim 0.5$ Gyr; (ii) at fixed age, stellar populations can exhibit a relatively broad spread in metallicity, due to local variations in chemical enrichment histories (Gallazzi et al. 2005; Haywood et al. 2013; Vincenzo & Kobayashi 2020; Agertz et al. 2021; Vincenzo et al. 2021; Buck et al. 2023; Carrillo et al. 2023; Graf et al. 2025). These two effects can compensate for the sharp decrease in efficiency at high ($Z > 4 \times 10^{-3}$) metallicity and soften the declining trend of the comoving merger rates. In the case of BBH mergers, Fig. 3 hints at a flattening of the profile for $z \lesssim 1$, which could potentially result in an overproduction of BBH merger events relative to current LVK constraints. While point (i) arises from the specific modeling choices made in the binary population synthesis simulations used to generate the GW progenitor tables (see Sec. 3), point (ii) reflects the chemical enrichment history of stellar populations in the MTNG simulation suite. Chemical enrichment depends on both the star formation history and the adopted stellar yields, which are themselves subject to significant uncertainties. Indeed, discrepancies as large as a factor of 2 or more are not uncommon among different yield models reported in the literature (see, e.g., Pillepich et al. 2018b; Hopkins et al. 2023).

To better evaluate the degree of chemical enrichment of the stellar component in the flagship MTNG740 simulation, we show in Fig. 7 the 2D distribution of stellar metallicity versus lookback birth time (i.e., stellar age) for all star particles. This 2D histogram is weighted by the comoving SFRD in each bin, computed as the total initial stellar mass formed within the bin, divided by the time bin width and the volume of the

simulated box. Results from the other MTNG simulations analyzed in this work are qualitatively similar and are presented in Fig. B.5 in Appendix B. The main difference across the boxes is an improved sampling of the low-metallicity, low-age region of the diagram in simulations with larger volumes, due to their higher star particle count.

From Fig. 7, it is evident that a correlation exists between stellar age and metallicity: in general, younger stars tend to exhibit higher metallicities compared to older ones. In particular, the figure shows that the bulk of star formation in the MTNG740 box occurs at metallicity around $Z \sim Z_{\odot}$ already by redshift $z \sim 3$, where the cosmic SFRD peaks (see Fig. 2). Star formation at these metallicities then continues toward $z = 0$, albeit at lower rates. Lower metallicity environments contribute progressively less to the total SFRD over time, with very low metallicities ($Z < 10^{-3} Z_{\odot}$) becoming relevant only at lookback times greater than ~ 8 –9 Gyr. Nevertheless, star formation at relatively low metallicity ($Z \sim 10^{-2} Z_{\odot}$) persists down to the present epoch.

The distribution of the SFRD as a function of metallicity and lookback time helps clarify the evolution of the GW progenitor comoving merger rates shown in Fig. 3. As illustrated in Fig. 7, the median metallicity at which most star formation occurs increases rapidly at early times, rising from $\sim 0.1 Z_{\odot}$ at $t_{\text{look}} \sim 13$ Gyr to approximately solar metallicity at $t_{\text{look}} \sim 11$ Gyr. Beyond that point, the median metallicity continues to grow steadily, reaching $Z \sim 1.5 Z_{\odot}$ by $z = 0$. When combined with the metallicity dependence of the binary merger efficiency for each channel (Fig. 6), this steady rise in the typical metallicity of stellar populations explains the observed timing of the peaks in comoving merger rates. Specifically, the BBH and BHNS rates peak slightly earlier than the SFRD itself (at $z \approx 3$), because their efficiencies decline sharply for $Z \gtrsim 4 \times 10^{-3}$. In contrast, BNS mergers peak later (at $z \approx 2$), because more time is needed for the formation of metal-rich stellar populations, which are more efficient at producing this type of events.

It is also crucial to quantify the spread in metallicity at fixed stellar age of the SFRD, especially for the BBH and BHNS channels, which are characterized by broader delay time distributions (Fig. 5). This contrasts with BNS systems, whose mergers typically occur with shorter delays, making their rates more directly coupled with the median metallicity of newly formed stars at a given lookback time. The metallicity spread – computed as half the width of the 16th–84th percentile range at each lookback time – is approximately 0.27–0.35 dex in the range $t_{\text{look}} \sim 0$ –10 Gyr, increasing up to ~ 0.9 dex beyond $z = 5$. These values indicate that stellar populations with metallicities a factor of ~ 2 below the median are common at essentially all cosmic times. When considering the strong dependence of BBH (and, to a lesser extent, BHNS) binary merger efficiency on metallicity, along with their relatively broad delay time distributions, this spread implies that a nonnegligible fraction of low-metallicity progenitors can still contribute to the merger rate even at late times. As previously noted, this delayed component of merger events can reduce the steepness of the decline of BBH merger rates at late times, potentially leading to an overproduction of events relative to current observational constraints.

To conclude our discussion of the metallicity dependence of GW progenitor events, Fig. 8 shows the distribution of the number of progenitors as a function of the metallicity of their parent stellar particle for the MTNG740 simulation. We present both the total distribution of merger events (solid black line) and the contributions from individual compact binary merger channels (solid colored lines). All distributions are normalized to the total

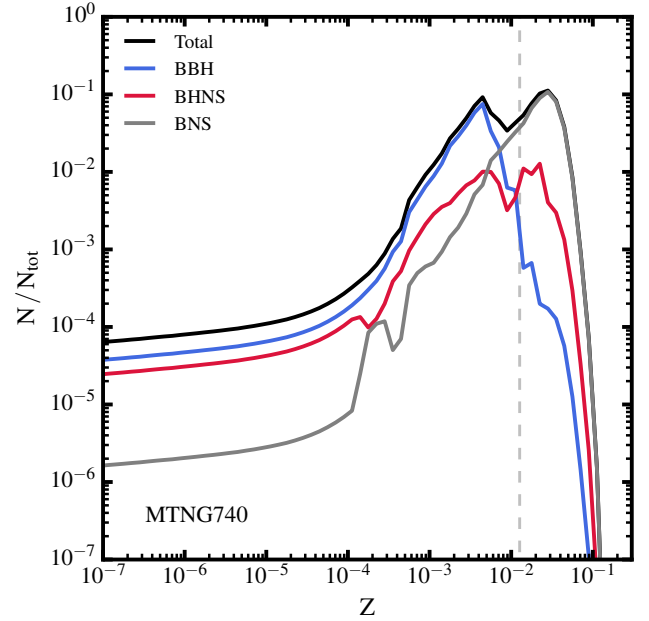


Fig. 8. Distribution of GW progenitor events as a function of the metallicity of the parent stellar particle in the MTNG740 simulation. The metallicity distribution of the total merger events, along with the contributions from different compact binary merger channels, is displayed as indicated in the legend. All distributions are normalized to the total number of merger events (i.e., $N_{\text{tot}} = \text{BBH} + \text{BHNS} + \text{BNS}$), so that each metallicity bin represents the fractional contribution of that specific metallicity to the overall number of events. As in Fig. 6, the vertical dashed line marks the value of Z_{\odot} adopted in MTNG.

number of events in the simulated box and the vertical dashed line marks the location of Z_{\odot} . Results for the other MTNG boxes analyzed in this work are shown in Fig. B.6 in Appendix B. The trends are consistent across all simulated volumes, with the larger volumes providing an improved sampling of the low-metallicity regime due to their higher star particle count.

To some extent, Fig. 8 can be viewed as the “integrated” counterpart – in terms of formed stellar mass at a given metallicity – of the binary merger efficiency curves presented in Fig. 6. Consequently, for all channels, we expect a relatively small fraction of merger events at very low metallicities ($Z \lesssim 10^{-4}$), due to the limited star formation occurring in that regime. The varying contributions of different channels at these low metallicities directly reflect the differences in binary merger efficiency (see Fig. 6). As metallicity increases above $Z = 10^{-4}$, the total fraction of merger events grows accordingly, following the rise in star formation activity discussed in Fig. 7. For BBH mergers, the fraction of events peaks at $Z \approx 4 \times 10^{-3}$, followed by a sharp decline driven by the steep drop in merger efficiency at higher metallicities. In contrast, BNS mergers peak at supersolar metallicities ($Z \sim 1.5 Z_{\odot}$), which roughly corresponds to the median metallicity where most star formation occurs in the simulation for the last ~ 10 Gyr. This is consistent with the BNS merger efficiency, which increases monotonically with metallicity. The subsequent decline in the BNS distribution is primarily due to the small number of stellar particles forming at $Z \gtrsim 1.5 Z_{\odot}$. BHNS mergers are subdominant compared to BBH at $Z \lesssim 4 \times 10^{-3}$ and to BNS at higher metallicities. Their distribution shows a general behavior similar to the other channels, with a rise in event fraction up to a maximum between $Z \sim 0.004$ and $Z \sim 0.02$, followed by a rapid decline. Interestingly, the BHNS distribution

exhibits two distinct peaks – at $Z \approx 2 \times 10^{-2}$ and $Z \approx 4 \times 10^{-3}$ – rather than a single maximum. This pattern is consistent with the shape of the merger efficiency for this channel, which shows a more gradual decline at high metallicities compared to BBH, and a mild increase (by a factor of ~ 2) in the range $Z = 0.01$ – 0.02 .

4.4. Mass functions of progenitor events and their evolution

Another key property that can be explored using the binary merger catalogs derived from the MTNG simulations is the mass function of progenitor events and its evolution with redshift. We present these mass functions in the MTNG740 box for BBH progenitors in Fig. 9 and for the BHNS progenitors in Fig. 10. Each figure contains two panels: the top panel shows the mass distribution of the primary object in the compact binary merger (m_1), while the bottom panel displays the distribution of the mass ratio ($q \equiv m_2/m_1$) of the binary system. Mass functions are shown for the entire redshift range covered by the simulation (black lines), as well as for three distinct redshift intervals: $z < 1$ (blue), $1 < z < 3$ (red), and $z > 3$ (gray). All distributions are normalized to the comoving volume of the simulation box and to the duration (in gigayears) of the respective redshift interval. Since the mass of BNS systems is resampled from a Gaussian distribution to reproduce the Galactic BNS population (see Sect. 3.1), we do not discuss its properties here. Results from the other MTNG simulation boxes are consistent with those of the flagship MTNG740 run and are shown in Appendix B in Fig. B.7 for BBH progenitors and Fig. B.8 for BHNS progenitors. For completeness, Appendix B also presents the mass functions of the BNS merger channel in Fig. B.9.

For BBH mergers (Fig. 9), it is evident that the mass distribution – both in terms of the primary mass and the mass ratio – evolves only mildly with redshift. The main variation lies in the normalization of the distributions, while their overall shape remains largely unchanged. These differences in normalization reflect the evolution of the average comoving BBH merger rates across cosmic time (see Fig. 3). Specifically, the lowest normalization is observed at $z < 1$, consistent with the monotonic decline in the BBH comoving merger rate at late times. In contrast, higher redshifts, particularly the $1 < z < 3$ interval where the merger rate peaks (around $z \approx 3$), show the largest values of the mass distribution. The distribution computed over the entire simulation time span falls between these two extremes, as expected. The mild evolution of the BBH mass function with redshift is consistent with the findings of Mapelli et al. (2019), who performed a similar analysis – albeit with more numerous and less wide redshift bins – using the Illustris-1 simulation. Notably, their study employed a comparable methodology to generate GW progenitor catalogs, though based on an earlier version of the binary population synthesis model (MOBSE) and applied to a cosmological simulation with different galaxy formation physics prescriptions.

The BBH mass distribution exhibits a steadily decreasing trend with the mass of the primary black hole, spanning from $m_1 \approx 4 M_\odot$ up to $m_1 \approx 70 M_\odot$. A particularly notable feature is a sharp drop at $m_1 \approx 44 M_\odot$, where the mass functions decreases by about two orders of magnitude. This is a consequence of the complex interplay between pair-instability supernovae and envelope stripping in binary star evolution, which produce a plateau at around 30 – $40 M_\odot$, and binary evolution processes – such as the CE phase and mass transfer – that remove mass from the system, hindering the formation of BHs with masses $\gtrsim 50 M_\odot$

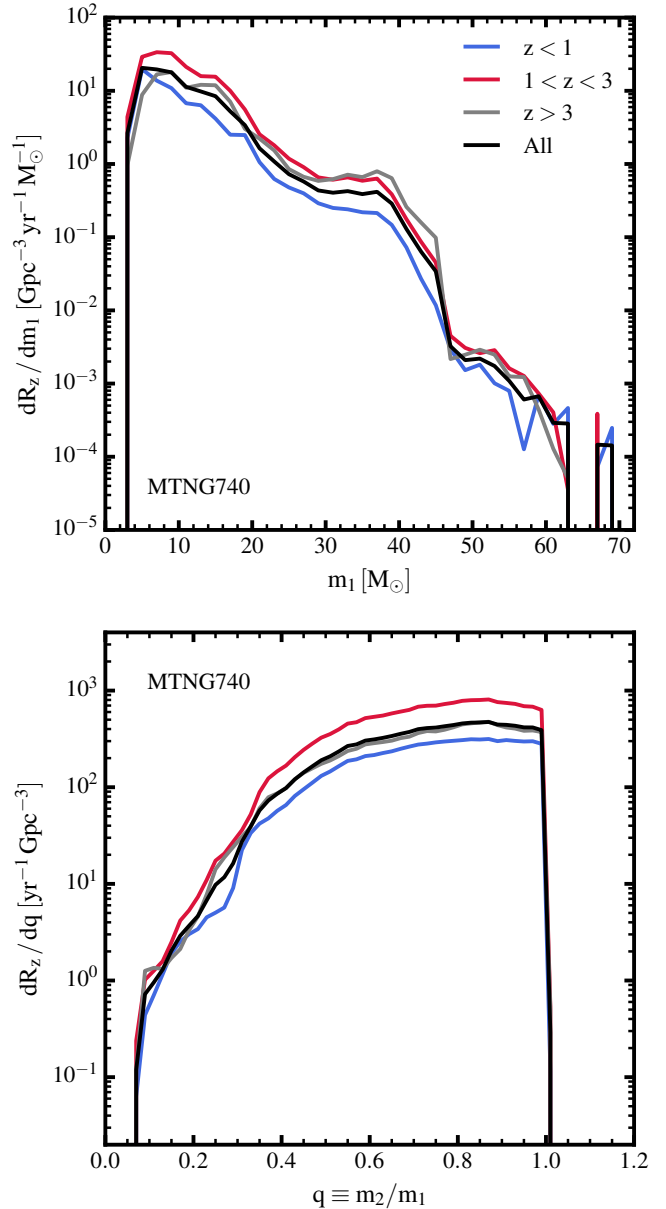


Fig. 9. Redshift evolution of differential merger rates of BBH merger GW progenitor events in the MTNG740 box. Top: distributions displayed as a function of mass of the primary black hole (m_1) of the merging compact binary systems. Bottom: Distributions in terms of mass ratio $q \equiv m_2/m_1$. Merger rates are normalized relative to the comoving volume of the simulation and to the time span (in gigayears) covered by the redshift intervals reported in the legend.

(Bouffanais et al. 2021; Iorio et al. 2023). In agreement with the latest observational constraints from LVK (LVK25, their Fig. 3; see also Fig. 11 of Abbott et al. 2023b for results obtained with the previous version of the GWTC catalog), the distribution shows local maxima around $m_1 \approx 10 M_\odot$ and $m_1 \approx 36 M_\odot$, although the prominence of the latter varies across redshift bins. Our results also indicate that the distribution extends to lower primary masses compared to LVK estimates. The absence of BBH merger progenitors with $m_1 \gtrsim 70 M_\odot$ is likely a consequence of not including the dynamical merger channel, which can produce systems with such high primary masses (e.g., Mapelli 2016; Rodriguez et al. 2016; Di Carlo et al. 2020), in our models. Regarding the mass ratio distribution, we observe

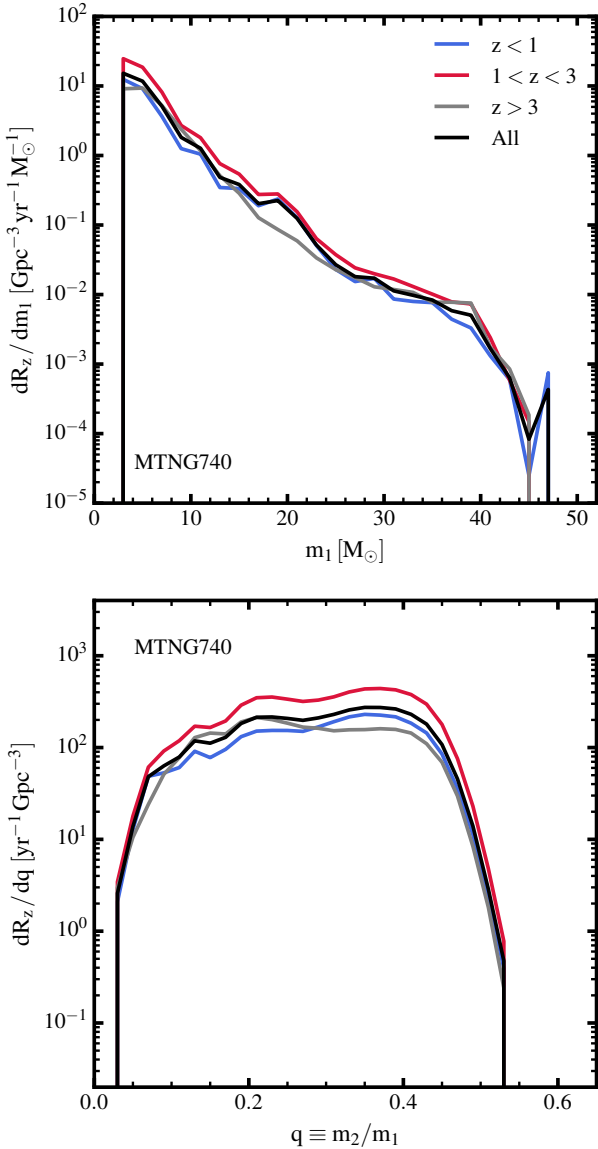


Fig. 10. Same as Fig. 9, but for BHNS merger GW progenitor events.

an opposite trend: the differential merger rate increases with q , rising by approximately two orders of magnitude in the range $0.2 < q < 0.4$ and then continuing to grow more gradually up to $q \sim 1$. This indicates a preference for BBH mergers involving near-equal mass components and may be a signature of mergers predominantly driven by a CE envelope evolutionary phase. However, near-equal mass BBH mergers in SEVN typically arise from a combination of binary evolution processes and supernova models (see [Iorio et al. 2023](#), their Fig. 14).

A mild evolution of the mass distribution with redshift is also found for the BHNS channel (Fig. 10). In this case, the mass distribution of the primary component – i.e., the black hole – exhibits a shape broadly similar to that of BBH mergers. The distribution decreases monotonically with increasing m_1 , spanning a narrower mass range from $m_1 \approx 4 M_\odot$ to $m_1 \approx 48 M_\odot$. A noticeable feature is a bump around $m_1 \approx 20 M_\odot$ for mergers occurring at redshifts $z < 3$. The distribution of BHNS mergers as a function of the mass ratio q is approximately flat for $q \gtrsim 0.2$ up to $q \approx 0.45$, beyond which it drops sharply. This decline at high q values is expected, given that the neutron star is signifi-

cantly less massive than its black hole companion. An exception to this trend is observed in the $z > 3$ redshift interval, where the flat portion of the distribution begins at $q \gtrsim 0.1$. The normalization trends across different redshift ranges are consistent with those found for the BBH channel. The relative flatness of the q distribution suggests that BHNS mergers occur over a wide range of mass ratios, without a strong preference for any particular value of q . This reflects the broader diversity in mass configurations characteristic of these mixed compact object binaries.

5. Summary and conclusions

In this paper, we introduced a flexible and fully parallel framework to build GW event catalogs in full-physics simulations of galaxy formation. Our framework works by associating binary merger events across three different channels (BBH, BHNS, and BNS) with stellar particles representing stellar populations in cosmological simulations. This association was performed by AREPO-GW – a module integrated in the moving-mesh code AREPO – by stochastically sampling (based on the mass, age and metallicity of the underlying stellar population) binary mergers tables obtained with the state-of-the-art population-synthesis code SEVN. AREPO-GW can operate in two main modes: (i) on-the-fly, performing the association of merger events during the course of a (hydrodynamic) galaxy formation simulation, and (ii) in post-processing, using as input a snapshot from an existing run. Our framework can be easily adapted to work with simulation outputs written in formats other than the standard AREPO format, as it requires only minimal input – namely, the initial mass, metallicity, and age of the star particles. Moreover, it also offers broad applicability, as it can be used with different sets of binary merger tables produced by SEVN or other stellar population synthesis codes.

To demonstrate the capabilities of our new framework, we applied AREPO-GW in its post-processing mode to the MTNG simulation suite, one of the largest and most advanced cosmological simulation projects to date. We built GW progenitor events catalogs for all four simulation boxes in the MTNG suite, including the flagship MTNG740 run, yielding a total data volume of ≈ 30 TB. In this work, we focused on the key global properties of the GW events (for instance, their comoving rates, formation efficiencies, delay time distributions, and the metallicity and mass distributions of their progenitors) and examined how these properties vary as a function of the simulated volume. Our main results can be summarized as follows:

- (i) The comoving rate of GW progenitor events follows the evolution of the cosmic SFRD, presenting a peak at early cosmic times ($z \approx 3$ for BBH and BHNS; $z \approx 2$ for BNS) followed by a decline toward $z = 0$. The BBH and BHNS events tend to peak slightly earlier than the SFRD, whereas BNS events more closely track it. The decline in the comoving rate for all GW progenitor channels is somewhat shallower than that of the SFRD, suggesting increasing merger efficiency per unit stellar mass formed at lower redshifts.
- (ii) The rates of BHNS and BNS at $z \sim 0$ are consistent with recent LVK observational determinations. In contrast, BBH rates are overproduced by a factor of ~ 4.5 , a discrepancy also seen in other predictions based on different stellar population synthesis models. In particular, LVK observations seem to favor a steeper slope in the redshift evolution of the BBH merger rate for $z \lesssim 1$ compared to that found in MTNG, although considerable uncertainty remains in the slope derived from observations.
- (iii) The GW progenitor event rates are shaped by two key factors: (i) their delay time distribution, which strongly depends on the assumptions made in stellar population synthesis modeling and (ii) the metallicity dependence

- of merger efficiency, which varies across channels and reflects the chemical enrichment history of the simulation. In MTNG, the quite rapid metal enrichment of stellar populations, together with a declining merger efficiency at higher metallicity for BBH and BHNS systems, explains their earlier peak relative to the SFRD. The narrow delay time distribution of BNS events ($t_{\text{delay}} \lesssim 0.1$ Gyr) results in a redshift evolution closely linked to recent star formation. Conversely, the long-delay tail of BBH events may contribute to their excess at late times relative to current constraints from the LVK collaboration.
- (iv) We also examined the mass functions of the different GW progenitor channels, both in terms of the distribution of primary binary masses (m_1) and merger mass ratios (q), as well as their evolution with redshift. In the three redshift bins analyzed, we found only a mild evolution in the mass functions, in agreement with previous studies adopting different binary population synthesis models and galaxy formation prescriptions. In particular, the shape of the mass functions remains largely unchanged with redshift, with only modest differences in normalization, consistent with the expected redshift evolution of the merger rates. This mild evolution is driven by the rapid metal enrichment of stellar populations in MTNG, which reach metallicities of $Z \sim Z_{\odot}$ by $z \sim 3$ and then stay roughly at the same levels down to $z = 0$. This in turn implies that the mass distributions of binary mergers remain nearly unchanged over time.
 - (v) The results discussed above remain consistent across all four MTNG simulation boxes analyzed in this work. This is not surprising, as the simulations cover progressively larger volumes of the Universe while maintaining the same mass resolution. Consequently, converged results for the SFRD are obtained across different boxes, leading to similar levels of metal enrichment and, therefore, comparable statistics for GW progenitor properties. It is worth noting that the MTNG model is based on the IllustrisTNG galaxy formation framework, which has demonstrated good convergence properties in general. However, drifts of up to $\sim 30 - 40\%$ in stellar mass with resolution have been observed, particularly for low-mass galaxies (Pillepich et al. 2018a). Therefore, star formation rates for low-mass galaxies are numerically biased low in MTNG, so some impact on the amplitude of the event rates can be expected at higher resolution.

The analysis presented in this paper represents a first step in exploiting the GW event catalogs produced for the MTNG simulations with AREPO-GW in conjunction with SEVN predictions. In particular, full light-cone outputs are available for MTNG, which can be combined with the GW event catalogs to generate light-cone outputs for GW events. This type of output will enable studies of their spatial distribution, clustering properties, correlations, and bias relative to the underlying stellar and matter density fields across multiple angular scales.

Another promising avenue made possible by the combination of the MTNG cosmological simulations and our GW event catalog is to associate the properties of the GW signals with those of the host galaxies of the events progenitors. This would allow investigations into how GW signals are influenced by local galaxy properties and the environments in which these objects reside. Such insights could also facilitate the implementation of a method similar to that presented in this work within semi-analytical models of galaxy formation (e.g., Somerville et al. 2015; Hirschmann et al. 2016; Lacey et al. 2016; Henriques et al. 2020). Semi-analytical models offer a computationally cheaper approach than full-physics

simulations for exploring how the initial properties of simulated binary systems and the modeling of key physical processes driving binary evolution – such as those adopted in SEVN or other stellar population synthesis codes – affect compact binary merger rates. The accuracy of these underlying assumptions represents a significant source of uncertainty in the results presented here.

Indeed, a current limitation of AREPO-GW is that only one binary population synthesis model can be used at a time. However, this does not diminish the value of applying the framework to large-scale cosmological simulations. Variations in galaxy formation physics can lead to substantial differences in stellar masses, star formation histories, and metallicity distributions, all of which directly affect the predicted population of GW progenitor events even when the same binary evolution model is used. Because different galaxy formation models must be run as separate simulations in any case, the on-the-fly mode offers an effective means to capture these differences self-consistently. By contrast, exploring the parameter space of binary evolution models remains computationally more efficient in post-processing, since a single hydrodynamic simulation can be reused. Future versions of AREPO-GW may support multiple population synthesis models or multiple independent samplings of the same model. Such extensions are expected to require only a moderate additional memory cost and would enable more comprehensive uncertainty quantification while keeping the framework computationally feasible.

Finally, starting from our GW event catalog, it would be extremely informative to produce mock observations of GW signals that fully incorporate observational uncertainties and selection effects to enable more robust comparisons and help with interpreting actual GW data. We will explore these research directions in future work, thereby advancing our understanding of GW signals and their connection to galaxy formation.

Acknowledgements. We thank the MTNG collaboration for kindly providing advance access to the simulation data. FM and MB gratefully acknowledge funding from the European Union – NextGenerationEU, in the framework of the HPC project – “National Centre for HPC, Big Data and Quantum Computing” (PNRR – M4C2 – 11.4 – CN00000013 — CUP J33C22001170001). GI was supported by a fellowship grant from the la Caixa Foundation (ID 100010434) under grant agreement LCF/BQ/PI24/12040020. MCA acknowledges financial support from FONDECYT Iniciación 11240540 and ANID BASAL project FB210003. MM acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2181/1 – 390900948 (the Heidelberg STRUCTURES Excellence Cluster) and from the European Research Council for the ERC Consolidator grant DEMOBLACK, under contract no. 770017. SB is supported by the UK Research and Innovation (UKRI) Future Leaders Fellowship [grant number MR/V023381/1]. VS and LH acknowledge support from the Simons Foundation through the “Learning the Universe” initiative. The initial conditions for the binary population synthesis simulations were generated with the IC4POPSYN code (<https://github.com/GiacobboNicola/IC4popsyn/tree/main>). SEVN simulations used in this work were made possible through a CINECA-INFN (TEONGRAV) agreement, providing access to resources on Leonardo at CINECA. For the analysis and visualization of simulation data, we used the Python packages NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Matplotlib (Hunter 2007), and mpi4py (Dalcin et al. 2005). The analysis presented in this paper was performed on the HPC computing cluster of the Open Physics Hub project (<https://site.unibo.it/openphysicshub/en>) hosted by the Department of Physics and Astronomy at the University of Bologna.

References

- Abac, A., Abramo, R., Albanesi, S., et al. 2025, ArXiv e-prints [arXiv:2503.12263]
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, *Phys. Rev. X*, 9, 031040
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2020, *ApJ*, 896, L44

- Abbott, R., Abbott, T. D., Abraham, S., et al. 2021, *Phys. Rev. X*, **11**, 021053
- Abbott, R., Abbott, T. D., Acernese, F., et al. 2023a, *Phys. Rev. X*, **13**, 041039
- Abbott, R., Abbott, T. D., Acernese, F., et al. 2023b, *Phys. Rev. X*, **13**, 011048
- Adhikari, S., Fishbach, M., Holz, D. E., Wechsler, R. H., & Fang, Z. 2020, *ApJ*, **905**, 21
- Agertz, O., Renaud, F., Feltzing, S., et al. 2021, *MNRAS*, **503**, 5826
- Artale, M. C., Bouffanais, Y., Mapelli, M., et al. 2020a, *MNRAS*, **495**, 1841
- Artale, M. C., Mapelli, M., Bouffanais, Y., et al. 2020b, *MNRAS*, **491**, 3419
- Artale, M. C., Mapelli, M., Giacobbo, N., et al. 2019, *MNRAS*, **487**, 1675
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Boco, L., Lapi, A., Goswami, S., et al. 2019, *ApJ*, **881**, 157
- Boco, L., Lapi, A., Chruslinska, M., et al. 2021, *ApJ*, **907**, 110
- Boesky, A. P., Broekgaarden, F. S., & Berger, E. 2024, *ApJ*, **976**, 23
- Bouffanais, Y., Mapelli, M., Santoliquido, F., et al. 2021, *MNRAS*, **505**, 3873
- Branchesi, M., Maggiore, M., Alonso, D., et al. 2023, *JCAP*, **2023**, 068
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, **427**, 127
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Buck, T., Obreja, A., Ratcliffe, B., et al. 2023, *MNRAS*, **523**, 1565
- Carrillo, A., Ness, M. K., Hawkins, K., et al. 2023, *ApJ*, **942**, 35
- Chabrier, G. 2003, *PASP*, **115**, 763
- Chen, Y., Bressan, A., Girardi, L., et al. 2015, *MNRAS*, **452**, 1068
- Chruslińska, M. 2024, *Ann. Phys.*, **536**, 2200170
- Claeys, J. S. W., Pols, O. R., Izzard, R. G., Vink, J., & Verbunt, F. W. M. 2014, *A&A*, **563**, A83
- Costa, G., Girardi, L., Bressan, A., et al. 2019, *MNRAS*, **485**, 4641
- Costa, G., Bressan, A., Mapelli, M., et al. 2021, *MNRAS*, **501**, 4514
- Costa, G., Shepherd, K. G., Bressan, A., et al. 2025, *A&A*, **694**, A193
- Dalcin, L., Paz, R., & Storti, M. 2005, *J. Paralle. Distrib. Comput.*, **65**, 1108
- Dehghani, A., Kim, J. L., Hosseini, D. S., et al. 2025, *JCAP*, **2025**, 056
- Di Carlo, U. N., Mapelli, M., Bouffanais, Y., et al. 2020, *MNRAS*, **497**, 1043
- Dominik, M., Belczynski, K., Fryer, C., et al. 2012, *ApJ*, **759**, 52
- Dominik, M., Belczynski, K., Fryer, C., et al. 2013, *ApJ*, **779**, 72
- Driver, S. P., Robotham, A. S. G., Bland-Hawthorn, J., et al. 2013, *MNRAS*, **430**, 2622
- Driver, S. P., Andrews, S. K., da Cunha, E., et al. 2018, *MNRAS*, **475**, 2891
- Eggleton, P. P. 1983, *ApJ*, **268**, 368
- Evans, M., Adhikari, R. X., Afle, C., et al. 2021, ArXiv e-prints [arXiv:2109.09882]
- Fryer, C. L., Belczynski, K., Wiktorowicz, G., et al. 2012, *ApJ*, **749**, 91
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, **362**, 41
- Giacobbo, N., & Mapelli, M. 2018, *MNRAS*, **480**, 2011
- Giacobbo, N., & Mapelli, M. 2020, *ApJ*, **891**, 141
- Graf, R. L., Wetzel, A., Bellardini, M. A., & Bailin, J. 2025, *ApJ*, **981**, 47
- Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, *MNRAS*, **467**, 179
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, **585**, 357
- Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. 2013, *A&A*, **560**, A109
- Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, *ApJ*, **591**, 288
- Henriques, B. M. B., Yates, R. M., Fu, J., et al. 2020, *MNRAS*, **491**, 5795
- Hirschmann, M., De Lucia, G., & Fontanot, F. 2016, *MNRAS*, **461**, 1760
- Hobbs, G., Lorimer, D. R., Lyne, A. G., & Kramer, M. 2005, *MNRAS*, **360**, 974
- Hopkins, P. F., Wetzel, A., Wheeler, C., et al. 2023, *MNRAS*, **519**, 3154
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, *MNRAS*, **329**, 897
- Hut, P. 1981, *A&A*, **99**, 126
- Iorio, G., Mapelli, M., Costa, G., et al. 2023, *MNRAS*, **524**, 426
- Ivanova, N., Justham, S., Chen, X., et al. 2013, *A&ARv*, **21**, 59
- Kelley, L. Z., Blecha, L., & Hernquist, L. 2017a, *MNRAS*, **464**, 3131
- Kelley, L. Z., Blecha, L., Hernquist, L., Sesana, A., & Taylor, S. R. 2017b, *MNRAS*, **471**, 4508
- Korb, E., Mapelli, M., Iorio, G., Costa, G., & Dall'Amico, M. 2025, *A&A*, **695**, A199
- Kroupa, P. 2001, *ASP Conf. Ser.*, **228**, 187
- Kruckow, M. U., Tauris, T. M., Langer, N., Kramer, M., & Izzard, R. G. 2018, *MNRAS*, **481**, 1908
- Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, *MNRAS*, **462**, 3854
- Lamberts, A., Garrison-Kimmel, S., Clausen, D. R., & Hopkins, P. F. 2016, *MNRAS*, **463**, L31
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Mapelli, M. 2016, *MNRAS*, **459**, 3432
- Mapelli, M., Giacobbo, N., Ripamonti, E., & Spera, M. 2017, *MNRAS*, **472**, 2422
- Mapelli, M., Giacobbo, N., Toffano, M., et al. 2018, *MNRAS*, **481**, 5324
- Mapelli, M., Giacobbo, N., Santoliquido, F., & Artale, M. C. 2019, *MNRAS*, **487**, 2
- Mapelli, M., Spera, M., Montanari, E., et al. 2020, *ApJ*, **888**, 76
- Marchant, P., Pappas, K. M. W., Gallegos-Garcia, M., et al. 2021, *A&A*, **650**, A107
- Marinacci, F., Sales, L. V., Vogelsberger, M., Torrey, P., & Springel, V. 2019, *MNRAS*, **489**, 4233
- Moe, M., & Di Stefano, R. 2017, *ApJS*, **230**, 15
- Neijssel, C. J., Vigna-Gómez, A., Stevenson, S., et al. 2019, *MNRAS*, **490**, 3740
- Nguyen, C. T., Costa, G., Girardi, L., et al. 2022, *A&A*, **665**, A126
- Özel, F., & Freire, P. 2016, *ARA&A*, **54**, 401
- Pakmor, R., Springel, V., Bauer, A., et al. 2016, *MNRAS*, **455**, 1134
- Pakmor, R., Springel, V., Coles, J. P., et al. 2023, *MNRAS*, **524**, 2539
- Perna, R., Artale, M. C., Wang, Y.-H., et al. 2022, *MNRAS*, **512**, 2654
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, *MNRAS*, **475**, 648
- Pillepich, A., Springel, V., Nelson, D., et al. 2018b, *MNRAS*, **473**, 4077
- Planck Collaboration XIII. 2016, *A&A*, **594**, A13
- Rauf, L., Howlett, C., Davis, T. M., & Lagos, C. D. P. 2023, *MNRAS*, **523**, 5719
- Ray, A., Farr, W., & Kalogera, V. 2025, ArXiv e-prints [arXiv:2507.09099]
- Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016, *Phys. Rev. D*, **93**, 084029
- Rose, J. C., Torrey, P., Lee, K. H., & Bartos, I. 2021, *ApJ*, **909**, 207
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Santoliquido, F., Mapelli, M., Artale, M. C., & Boco, L. 2022, *MNRAS*, **516**, 3297
- Schneider, R., Graziani, L., Marassi, S., et al. 2017, *MNRAS*, **471**, L105
- Sgalletta, C., Iorio, G., Mapelli, M., et al. 2023, *MNRAS*, **526**, 2210
- Sgalletta, C., Mapelli, M., Boco, L., et al. 2025, *A&A*, **698**, A144
- Siwiek, M. S., Kelley, L. Z., & Hernquist, L. 2020, *MNRAS*, **498**, 537
- Somerville, R. S., Popping, G., & Trager, S. C. 2015, *MNRAS*, **453**, 4337
- Spera, M., Mapelli, M., & Bressan, A. 2015, *MNRAS*, **451**, 4086
- Spera, M., Mapelli, M., Giacobbo, N., et al. 2019, *MNRAS*, **485**, 889
- Springel, V. 2010, *MNRAS*, **401**, 791
- Springel, V., & Hernquist, L. 2003, *MNRAS*, **339**, 289
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, **435**, 629
- Tang, P. N., Eldridge, J. J., Stanway, E. R., & Bray, J. C. 2020, *MNRAS*, **493**, L6
- The LIGO Scientific Collaboration, the Virgo Collaboration, & the KAGRA Collaboration 2025, ArXiv e-prints [arXiv:2508.18083] (LVK25)
- van Son, L. A. C., de Mink, S. E., Chruslińska, M., et al. 2023, *ApJ*, **948**, 105
- Vigna-Gómez, A., Neijssel, C. J., Stevenson, S., et al. 2018, *MNRAS*, **481**, 4009
- Vijaykumar, A., Fishbach, M., Adhikari, S., & Holz, D. E. 2024, *ApJ*, **972**, 157
- Vincenzo, F., & Kobayashi, C. 2020, *MNRAS*, **496**, 80
- Vincenzo, F., Weinberg, D. H., Miglio, A., Lane, R. R., & Roman-Lopes, A. 2021, *MNRAS*, **508**, 5903
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, **17**, 261
- Vogelsberger, M., Genel, S., Sijacki, D., et al. 2013, *MNRAS*, **436**, 3031
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, **444**, 1518
- Webbink, R. F. 1984, *ApJ*, **277**, 355
- Webbink, R. F. 1985, in *Interacting Binary Stars*, eds. J. E. Pringle, & R. A. Wade, 39
- Weinberger, R., Springel, V., Hernquist, L., et al. 2017, *MNRAS*, **465**, 3291
- Woosley, S. E. 2017, *ApJ*, **836**, 244
- Wyrzykowski, Ł., & Mandel, I. 2020, *A&A*, **636**, A20
- Yoshida, T., Umeda, H., Maeda, K., & Ishii, T. 2016, *MNRAS*, **457**, 351
- Zevin, M., Spera, M., Berry, C. P. L., & Kalogera, V. 2020, *ApJ*, **899**, L1

Appendix A: Quantitative comparison between the on-the-fly and post-processing modes of Arepo-GW

As discussed in Sec. 3.2, the two operational modes of AREPO-GW process the stellar particles of a specific simulation using a slightly different approach. Given that the association between star particles and GW progenitor events is performed stochastically in both modes, it is not possible to obtain the exact same sequence of binary merger events even when analyzing the same simulation. The goal of this Appendix is to compare the on-the-fly and post-processing modes and quantify the extent of the resulting discrepancies.

For this comparison, we decided to rerun the MTNG46 box with the on-the-fly mode of AREPO-GW enabled and post-process the redshift zero snapshot of this new simulation for consistency. This setup enables a direct, controlled comparison between the two GW catalogs produced with the two approaches, ensuring that the same set of stellar particles (with identical formation masses, formation times, and metallicities) is used in both cases. Therefore, any differences in the outcome can be attributed solely to the operational mode of AREPO-GW.

Below, we report the results of this analysis. We present a comparison of the comoving rate of GW progenitor events as a function of lookback time (Fig. A.1, top left panel), the delay time distribution of GW progenitor events (Fig. A.1, top middle panel), the merger efficiency of compact binary objects as a function of the tabulated metallicity of binary systems (Fig. A.1, top right panel), the distribution of GW progenitor events as a function of the metallicity of the parent stellar particle (Fig. A.1, bottom left panel), and the BBH and BHNS mass functions (Fig. A.1, bottom middle and bottom right panel, respectively). For each quantity, the panels show the relative difference between the two approaches, adopting the post-processing mode as the reference case. These results clearly demonstrate that the impact of the specific operational mode of AREPO-GW on the statistical properties of the population of the GW progenitor events is minimal. The observed differences are at most on the order of a few percent and can be attributed to the stochastic nature of the association procedure between simulated star particles and compact binary mergers implemented in AREPO-GW. Occasionally, larger relative differences – up to $\sim 10 - 20\%$ – can be present, but only for rare events, such as systems with long delay times (≥ 12 Gyr) or extremely metal-rich ($Z \geq 0.1$) or high-mass ($m_1 \geq 55 M_\odot$) progenitors. Larger simulation boxes are expected to yield comparable or more consistent results between the two modes, due to the increased number (and therefore better sampling) of GW progenitor events.

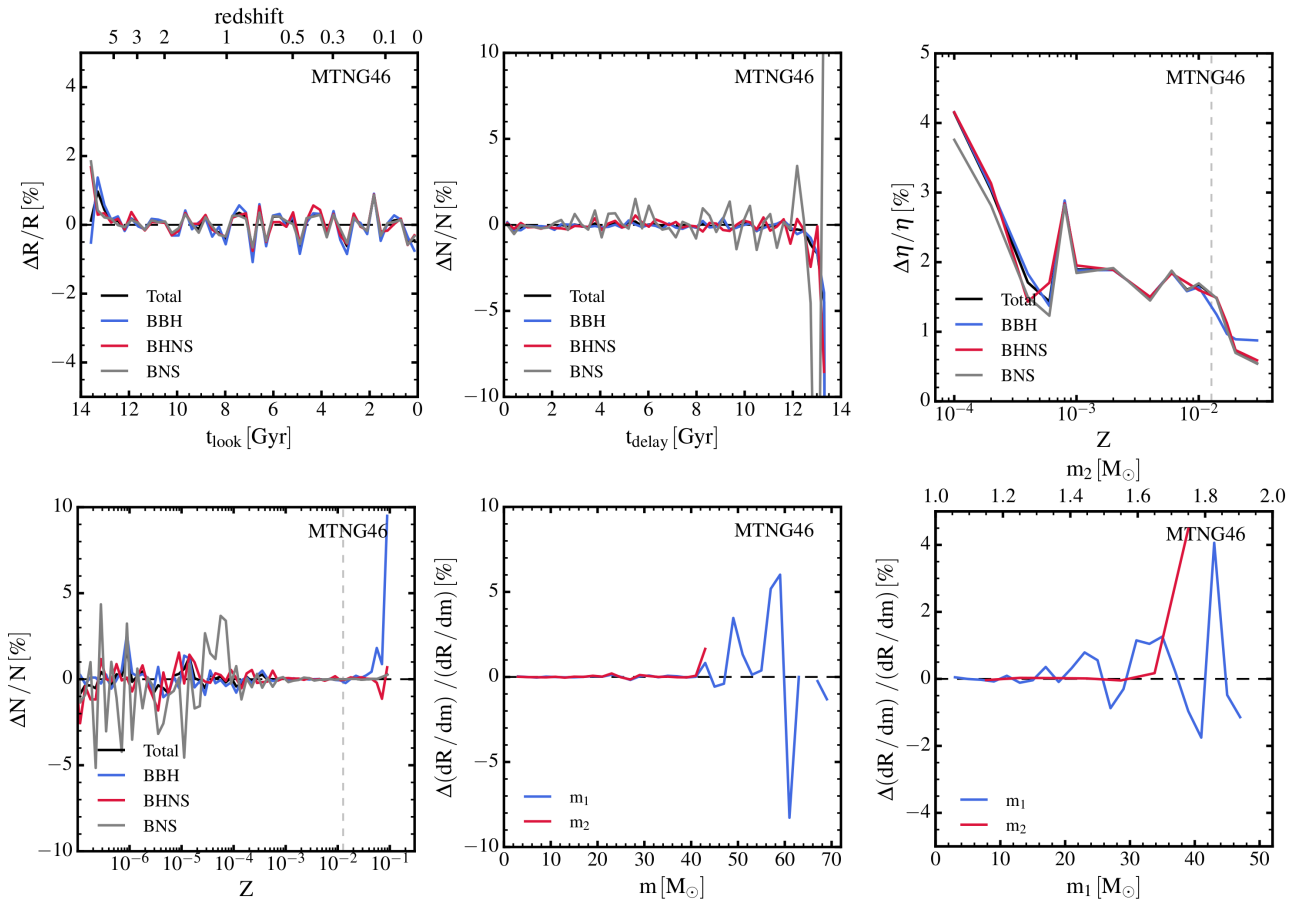


Fig. A.1. Relative differences between properties of GW progenitor events obtained with the on-the-fly and post-processing modes of AREPO-GW for a rerun of the MTNG46 simulation. Differences are expressed as percentages, with the post-processing mode taken as the reference case. Top row: Comoving rates as a function of time (left panel), delay time distribution (middle panel), and merger efficiency as a function of tabulated SEVN metallicity (right panel). Bottom row: Distribution of GW progenitor events as a function of the metallicity of the parent star particle (left panel), BBH (middle panel), and BHNS (right panel) mass functions.

Appendix B: GW progenitor properties across different MTNG boxes

In this Appendix, we present the same set of results shown in the main body of the paper for the flagship MTNG740 simulation, extended to all the boxes of the MTNG suite for which we produced a GW progenitor catalog. As mentioned above, the four analyzed boxes differ only in terms of the simulated volume, but adopt the same prescriptions and parameterization for galaxy formation physics, as well as the same mass resolution.

Each figure below is organized consistently. For each row, panels from left to the right display results for simulations with progressively increasing volume. We decided to also include MTNG740 results (already presented in the main body of the paper) as reference to facilitate comparison. The figures are displayed in the same order as discussed in the main text: comoving merger rates (Fig. B.1), evolution of the binary merger efficiency per unit stellar mass (Fig. B.2), delay time distribution (Fig. B.3), binary merger efficiency as a function of metallicity (Fig. B.4), 2D histogram of SFRD as a function of lookback time and metallicity (Fig. B.5), distribution of merger events as a function of the metallicity of the parent stellar particle (Fig. B.6), and mass functions for BBH (Fig. B.7), BHNS (Fig. B.8), and BNS (Fig. B.9, included here for completeness) mergers.

Owing to the constant mass resolution across all simulations, the results are largely consistent among the different volumes. The only noticeable difference arises in the smaller simulation volumes, where sampling noise becomes more evident because fewer GW progenitor events are generated. This effect is most pronounced in the MTNG46 box, which represents a relatively small cosmological volume, but it becomes progressively less significant in larger boxes. Therefore, relatively modest simulated volumes – which are less computationally expensive to run, post-process and analyze – can be used to calibrate AREPO-GW parameters and test alternative version of merger tables produced by binary population synthesis simulations. This approach is extremely helpful for assessing the performance of different GW progenitor models, which can be employed, for instance, to address the tension with LVK results regarding the expected rate of BBH mergers.

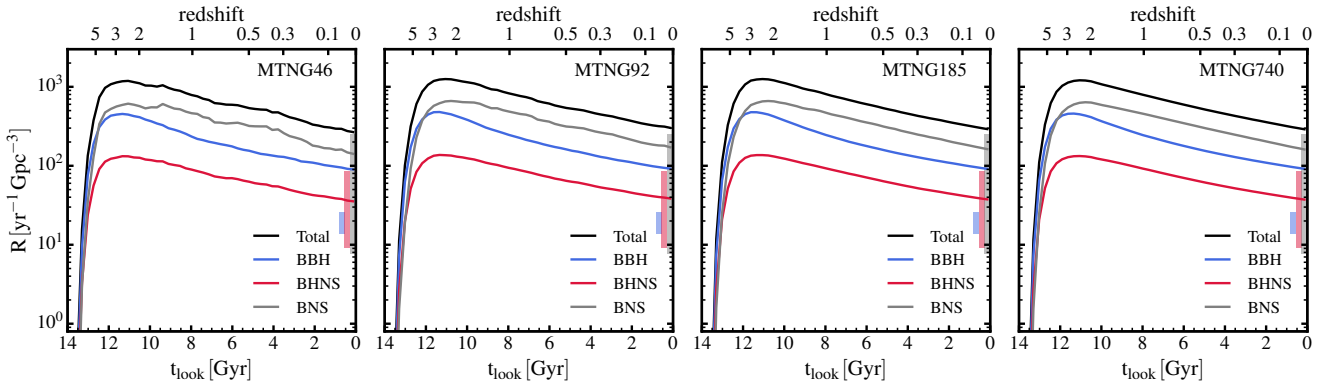


Fig. B.1. Rate of GW progenitor events per comoving volume as a function of lookback time and redshift in four different boxes of the MTNG simulation suite: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). The total progenitor merger rate and the contribution of different compact binary merger channels are displayed as indicated in the legend. Colored bands represent constraints on the low-redshift ($z \approx 0$) merger rate of the different channels, as inferred from GWTC-4 high-confidence events (LVK25).

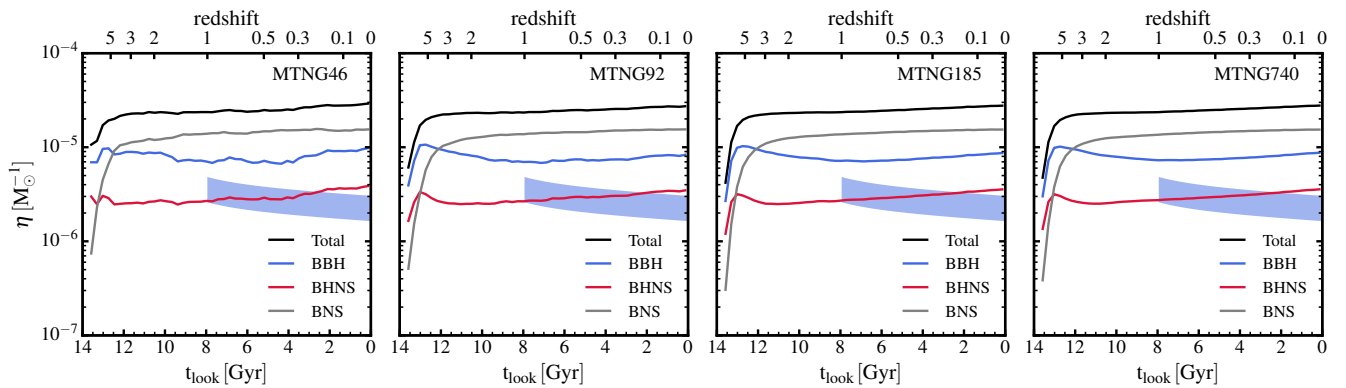


Fig. B.2. Merger efficiency of compact binary objects per unit of stellar mass formed as a function of time and redshift in four different boxes of the MTNG simulations: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). The total merger efficiency, along with the contributions from different merger channels, is displayed as indicated in the legend. Shaded regions represent constraints on the evolution of BBH merger efficiency at low redshifts ($z \leq 1$), derived by extrapolating the (LVK25) limits as in Fig. 4.

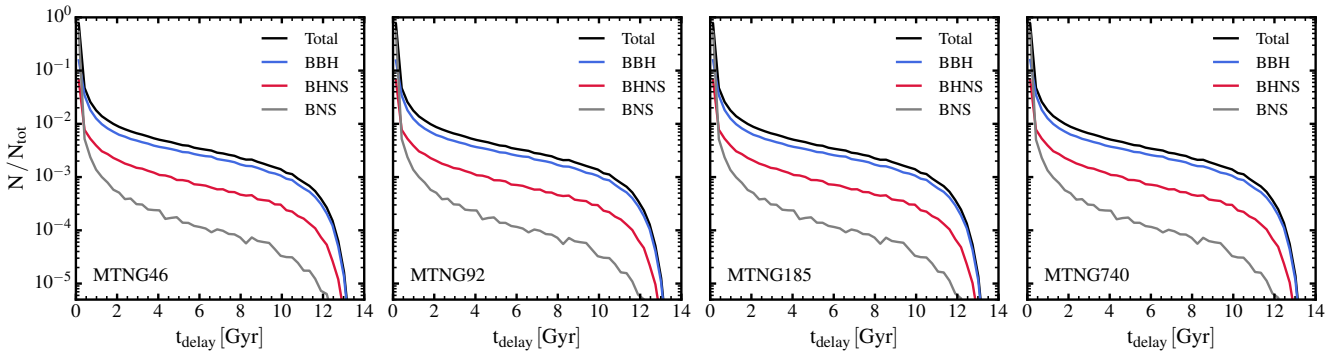


Fig. B.3. Delay time distribution of GW progenitor events in four different boxes of the MTNG simulations: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). The delay time distribution of the total merger events, along with the contributions from different compact binary merger channels, is displayed as indicated in the legend. All distributions are normalized to the total number of merger events.

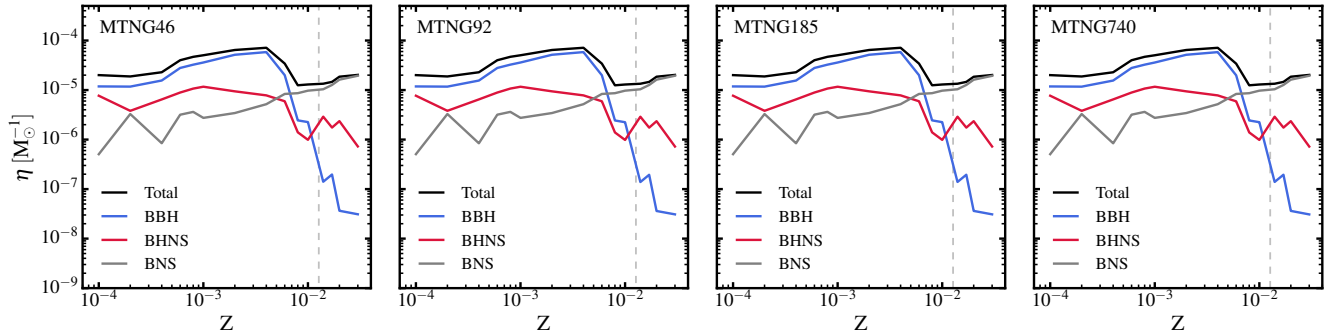


Fig. B.4. Merger efficiency of compact binary objects per unit of stellar mass formed as a function of tabulated metallicity of binary systems simulated with the SEVN code in four different boxes of the MTNG simulations: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). The total merger efficiency, along with the contributions from different merger channels, is displayed as indicated in the legend. The vertical dashed lines indicate the solar metallicity value, $Z_{\odot} = 0.0127$, adopted in MTNG.

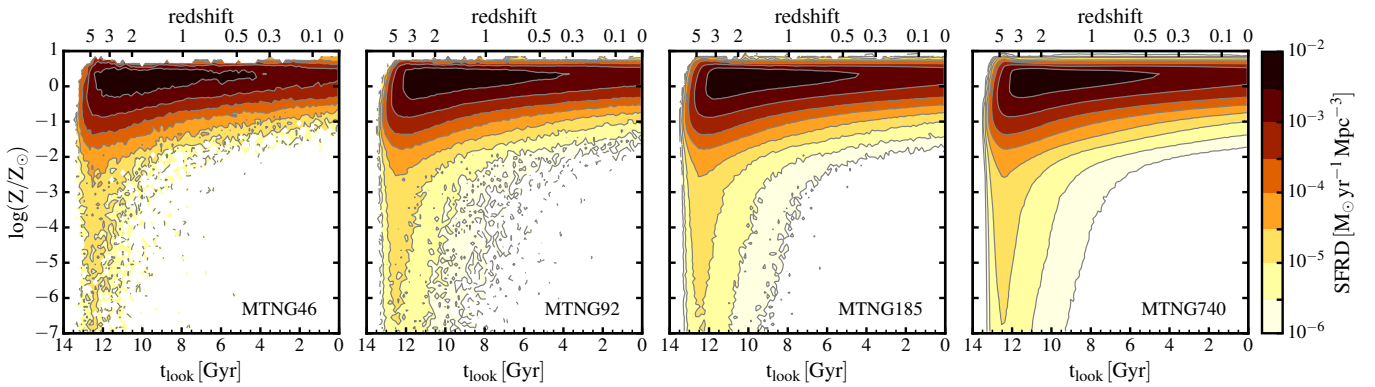


Fig. B.5. 2D distribution of metallicity and (lookback) birth times of stellar particles in four different boxes of the MTNG simulations: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). The color scale indicates the corresponding (comoving) SFRD. Contours are placed at $[1, 3, 10, 30, 100, 300, 1000, 3000, 10^4] \times 10^{-6} M_{\odot} \text{yr}^{-1} \text{Mpc}^{-3}$.

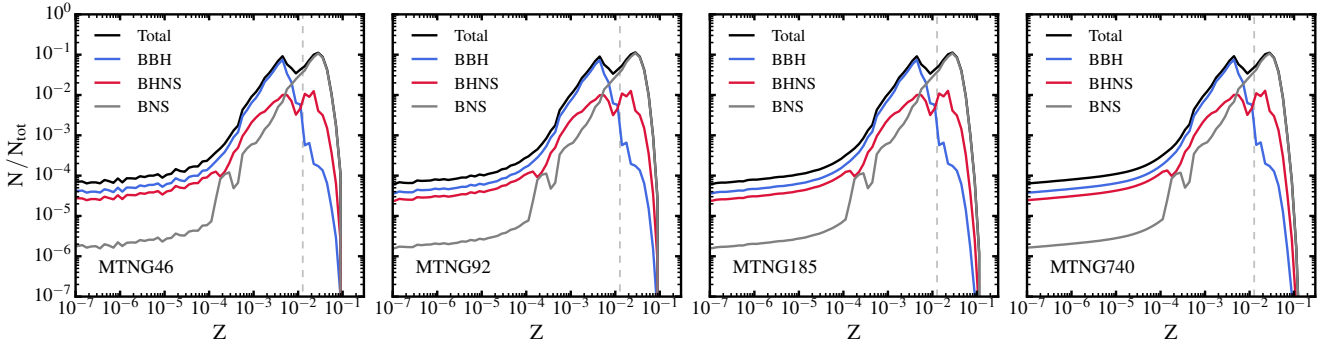


Fig. B.6. Distribution of GW progenitor events as a function of the metallicity of the parent stellar particle in four different boxes of the MTNG simulations: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). The metallicity distribution of the total merger events, along with the contributions from different compact binary merger channels, is displayed as indicated in the legend. All distributions are normalized to the total number of merger events. As in Fig. B.4, the vertical dashed lines mark the value of Z_{\odot} adopted in the simulations.

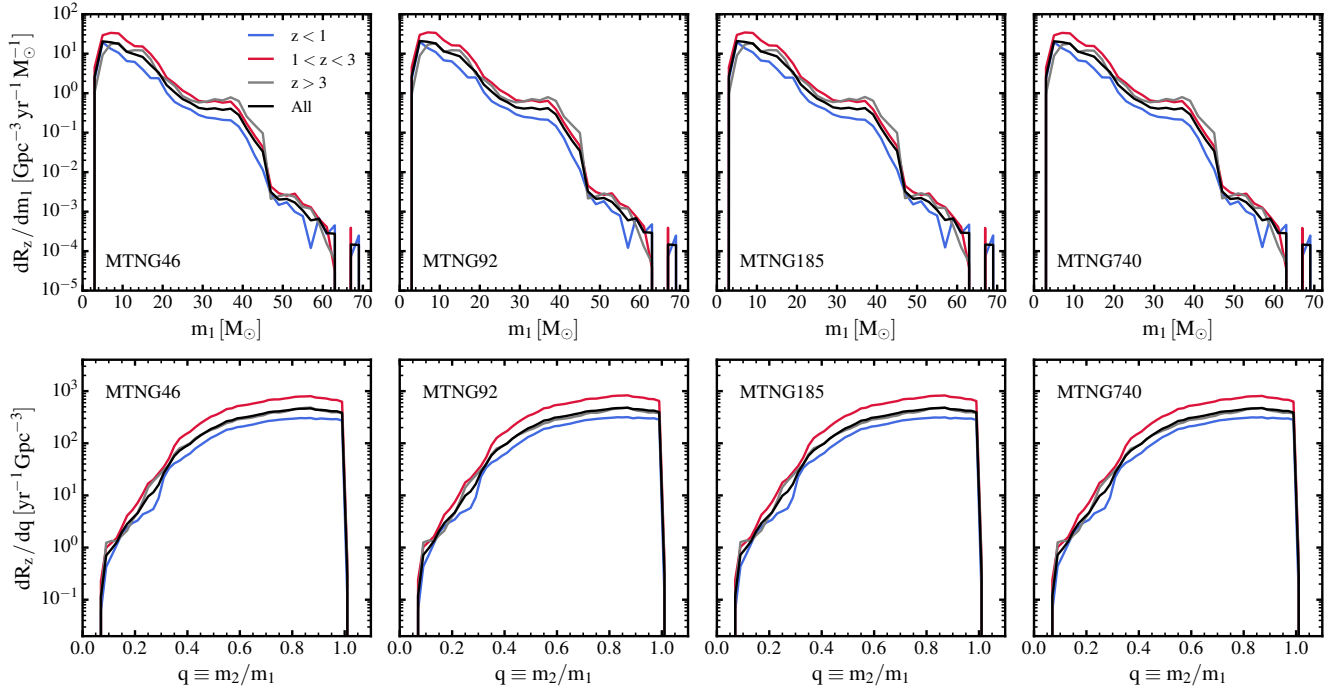


Fig. B.7. Redshift evolution of differential merger rates of BBH merger GW progenitor events in four different boxes of the MTNG simulations: MTNG46, MTNG92, MTNG185, and MTNG740 (from left to right). Top: Distributions displayed as a function of the mass of the primary black hole (m_1) of the merging compact binary systems. Bottom: Distributions in terms of mass ratio $q \equiv m_2/m_1$. Merger rates are normalized relative to the comoving volume of each simulation and to the time span (in gigayears) covered by the redshift intervals reported in the legend.

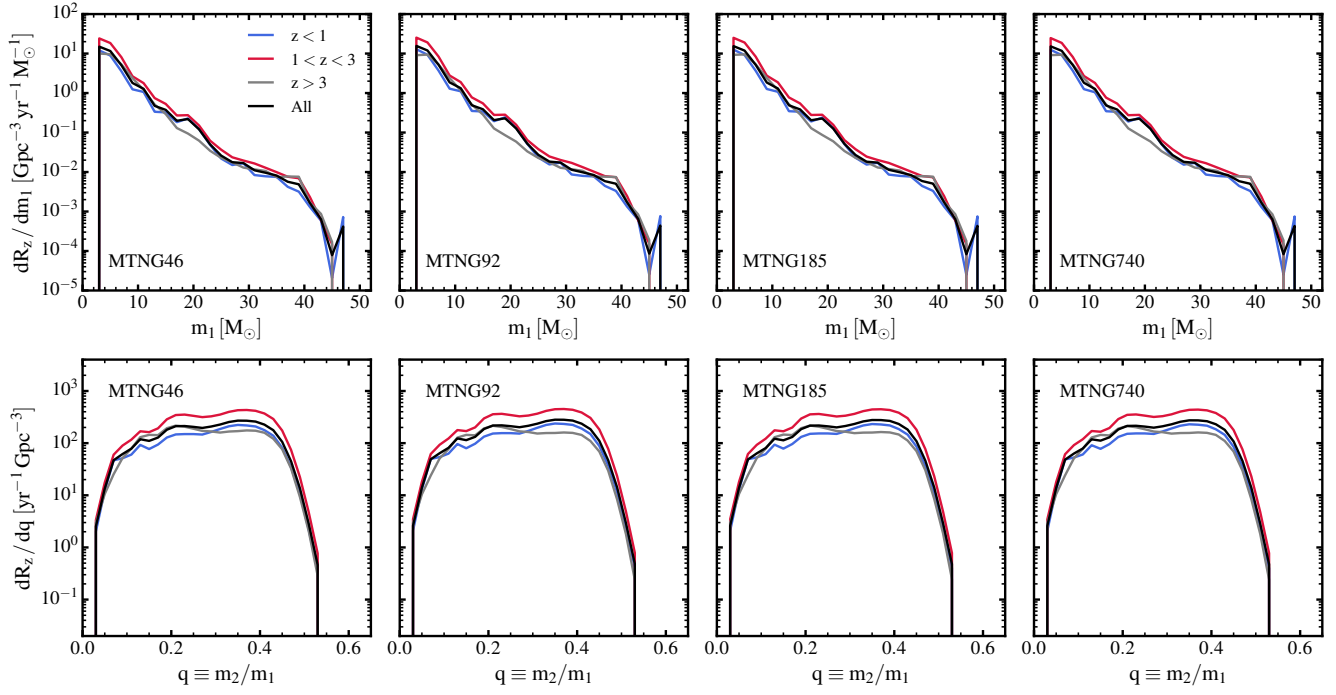


Fig. B.8. Same as Fig. B.7, but for BHNS merger GW progenitor events.

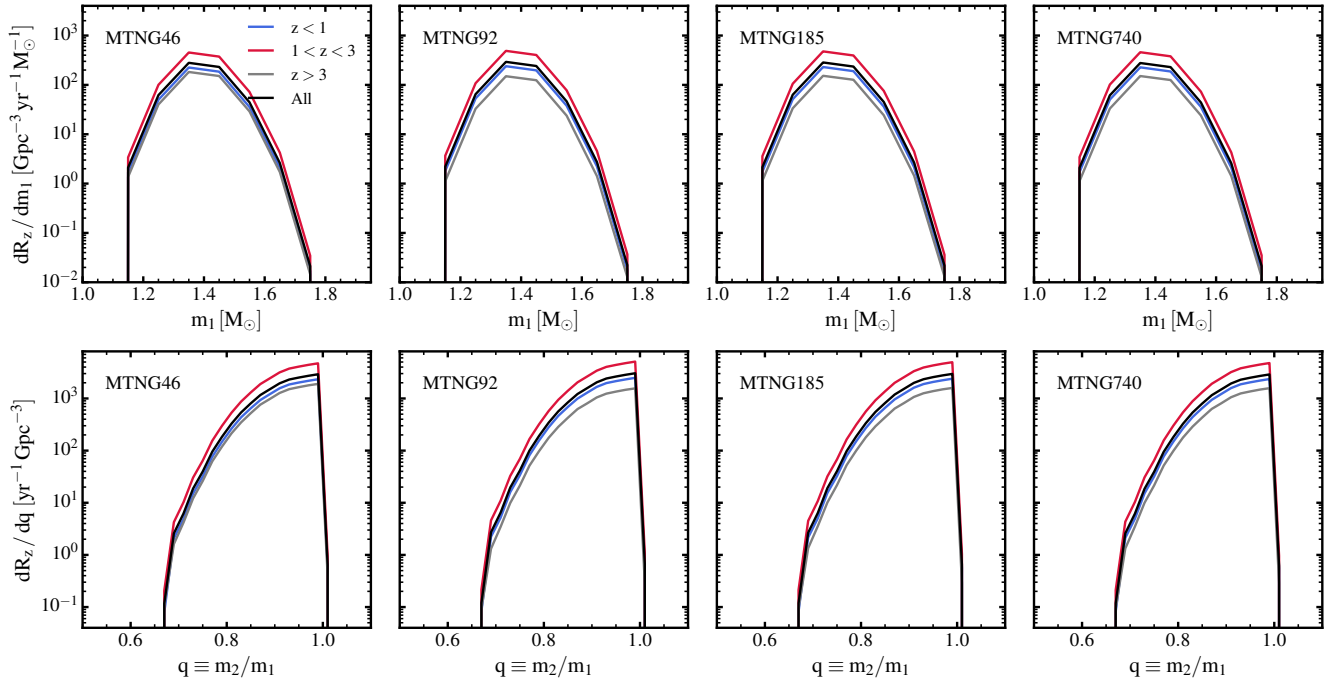


Fig. B.9. Same as Fig. B.7, but for BNS merger GW progenitor events.