HADES: a new numerical tool for the determination of DM over-densities

Anna Balaudo\textsuperscript{1,2}, Francesca Calore\textsuperscript{3}, Valentina De Romeri\textsuperscript{4} and Fiorenza Donato\textsuperscript{5,6} *

\textsuperscript{1}Leiden Observatory, Leiden University, PO Box 9513, Leiden 2300 RA, The Netherlands
\textsuperscript{2}Lorentz Institute for Theoretical Physics, PO Box 9506, Leiden 2300 RA, The Netherlands
\textsuperscript{3}Univ. Grenoble Alpes, USMB, CNRS, LAPTh, F-74940 Annecy, France
\textsuperscript{4}Instituto de Física Corpuscular, CSIC-Universitat de València, 46980 Paterna, Spain
\textsuperscript{5}Dipartimento di Fisica, Università di Torino, via P. Giuria, 1, I-10125 Torino, Italy
\textsuperscript{6}Istituto Nazionale di Fisica Nucleare, Sezione di Torino, via P. Giuria, 1, I-10125 Torino, Italy

22 December 2021

ABSTRACT

Cosmological simulations predict dark matter to form bound structures, i.e. main halos, hosting galaxies and eventually a population of less massive dark matter over-densities, i.e. sub-halos. The determination of the spatial dark matter distribution in halos and sub-halos is one major challenge in the analysis of galaxy formation simulations, and usually relies on halo finder algorithms coupled with approximated analytical density profiles. Such determination is crucial for deriving, among others, dark matter signatures in astroparticle observables, e.g. the flux of gamma-ray photons from dark matter particle annihilation. We present here the “Halo Accurate Density Evaluation System”, HADES, a novel numerical tool to reconstruct the dark matter density locally at any point in the simulation volume with high accuracy. We run thorough tests of the code performances on dedicated mock realisations of halos starting from analytical dark matter profile distributions. We show that on mock halos HADES can recover the dark matter density with a few % accuracy, resolving efficiently all sub-structures containing down to 1000 particles and providing conservative estimates for smaller sub-halos. We illustrate how HADES can be used to compute all-sky maps of the dark matter spatial distribution, squared and integrated along the line of sight, already accounting for the signal boosting coming from density fluctuations. We also present an application of the code to one halo in the TNG50 run from the IllustrisTNG suite, and highlight how HADES automatically maps out the asymmetries present in the dark matter spatial distribution both in halos and sub-halos, promising to become a helpful tool for a vast number of astrophysical applications.

Key words: astroparticle physics, dark matter, software:simulations, methods: numerical

1 INTRODUCTION

Probing the nature of the most elusive matter component in the universe, the dark matter (DM), is one of the major endeavours of physics, providing a bridge between particle physics and astrophysics. Astroparticle observables could be very successful in testing the DM nature for a broad range of candidates (Batista et al. 2021). One crucial ingredient for indirect DM searches is the DM distribution in astrophysical systems, from dwarf-sized galaxies to large clusters and galaxy groups. As an example, the expected photon flux from DM is proportional to the integral along the line of sight of the DM spatial distribution, eventually squared in the case of DM particle annihilation, and, therefore, the ensuing signal is very sensitive to uncertainties in the DM density.

The DM distribution in galaxies can be constrained by observational data, but its derivation is still prone to large uncertainties, especially for the inner regions of the Milky Way and ultra-faint dwarf galaxies (see e.g. Pato & Iocco (2015); Bonnivard et al. (2015); Chang & Necib (2021)).

To get an estimate of systematic uncertainties to be used for indirect signal predictions, one typically relies on the outcome of numerical simulations for galaxy formation, which provide a proxy for the DM distribution at different scales as well as a way to quantify statistical and systematic uncertainties by analysing large samples of objects (e.g. Calore et al. (2017); Coronado-Blazquez et al. (2019a); Grand & White (2021)), or, alternatively, on semi-analytical models of the halo statistics [e.g. Stref & Lavalle (2017); Ando et al. (2019)]. In the past decades, a huge effort has been put into finding a universal parametrisation to describe the density profiles of DM halos. These parametrisations are typically based on the assumption of spherical symmetry and contain a number of free parameters that need to be properly tuned to match the shape of peculiar halos. Arguably, the most famous density profiles present in the literature include central cored profiles, of which the Burkert parametrisation might be considered the main representative (Burkert 1995), along with a family of cuspy profiles, such as the Einasto one (Einasto 1965; Graham et al. 2006; Navarro et al. 2004) and the Navarro, Frenk and White parametrisation (hereafter NFW) (Navarro et al. 1996, 1997) with its many generalizations – the Moore profile (Moore et al. 1999), the Hernquist profile (Hernquist 1990), the Dehnen & McLaughlin profile (Dehnen & McLaughlin 2005) and many others. While all these functions present quite similar behaviours...
at high radii, they provide significantly different predictions on the DM density when extrapolated towards the centre of the halo. While in DM-only simulations halos tend to have a prolate shape, and the most massive ones can even present a triaxial structure (see e.g. (Despali et al. 2014)), in hydrodynamical simulations the shapes of halos, notably Milky-Way sized objects, are predicted to be more spherical (Chua et al. 2019; Prada et al. 2019). Nonetheless, a mild deviation from sphericity remains, in contrast with the common practice of describing the matter distribution by fitting the halo as a spherically averaged density profile.

Of relevance for DM indirect detection searches is also the distribution of DM over-densities in halos. The DM sub-halos or generically sub-structures offer indeed further leverages to look for DM signals. On the one hand, the presence of DM over-densities on the top of the smooth DM halos enhances the DM expected flux by the so-called “boost” factor (Silk & Stebbins 1993; Bergström et al. 1998; Diemand et al. 2007a; Pieri et al. 2008; Diemand et al. 2008; Kamionkowski et al. 2010; Bartels & Ando 2015; Stref & Lavalle 2017; Moliné et al. 2017; Hiroshima et al. 2018; Grand & White 2021). On the other hand, one can target the signal directly originating from sub-halos, for example looking for unidentified sources of X- and gamma-ray radiation (see e.g. (Bertoni et al. 2015; Schoonenberg et al. 2016; Hooper & Witte 2017; Calore et al. 2017; Coronado-Blázquez et al. 2019b; Coronado-Blázquez et al. 2019a; Coronado-Blázquez & Sánchez-Conde 2019; Di Mauro et al. 2020; Abdalla et al. 2021)), or for anisotropies in cosmic radiation backgrounds (Siegal-Gaskins 2008; Ando 2009; Fornasa et al. 2013; Ackermann et al. 2012; Calore et al. 2014; Hütten & Maier 2018).

Determining the halos and sub-halos abundance, mass and spatial distribution is typically done with the aid of halo finders, which are algorithms applied to numerical simulations of structure formation (for comprehensive reviews see for example Knebe et al. (2011); Onions et al. (2012); Knebe et al. (2013)). Halo finders are designed to uncover over-dense regions in the simulations – eventually associated to sub-halos – and to extract some of their main properties. Typically, the algorithms scan the simulation box, eventually exploiting the full 6-dimensional (spatial and velocity) information, look for local peaks in the density field, identify samples of gravitationally bounded particles, which then define the (sub-)halos. Generally speaking, halo finders can be categorised into three classes. Configuration-space finders scrutinise the spatial clustering of particles. For example, these include finders such as Friends of Friends (FoF) (Davis et al. 1985; Knebe et al. 2011), which identify halos by linking together particles closer than a specific threshold. In FoF catalogues, a given particle can be assigned uniquely to a single FoF group and sub-halos are always completely comprised within their host. A second category of halo finders includes density peak locator codes like SUBFIND (Springel et al. 2001), which estimates the local density through an adaptive kernel interpolation, selecting regions where the density exceeds a user-set lower limit. The Amiga Halo Finder (Knollmann & Knebe 2009), an improvement of the Multi-Level Adaptive Particle Mesh (MLAPM) halo finder (Gill et al. 2004), is based on an adaptive mesh applied iteratively on smaller and smaller simulation volumes, until the sub-structure is singled out. The Bound Density Maxima (BDM) halo finder (Klypin et al. 1999) detects local density maxima by means of a spherical cut-off for the halo. VOBZ (Neyrinck et al. 2005) density estimates are instead based on a Voronoi tessellation, while 6DFOF (Diemand et al. 2006), ROCKSTAR (Behroozi et al. 2013) and VELOCIRAPTOR (Elahi et al. 2019) are typical examples of phase-space halo finders considering clustering in both spatial and velocity space.

Finally, other halo finders follow a third approach, not focusing only on the last snapshot, but tracing sub-structures backwards (or forward) in time. This is the case for instance of the Hierarchical Bound-Tracing (HBT) halo finder (Han et al. 2012, 2018), which works in the time domain of each sub-halo’s evolution, and of SURV (Tormen & Moscardini 2004).

In the end, halo finders necessarily have to make some choices when defining the existence and extent of sub-structures. Then, a series of pitfalls hinders to the precise determination of DM mass and spatial profiles in (sub-)halos. First, the mass definition of bounded over-densities is not universal among different algorithms. This is a crucial issue, since whether a particle is gravitationally bound to a structure depends on the mass of the object itself. Comparisons of halo finders performances showed that the mass estimates for the same sub-halo can differ up to 20–30% for most massive structures, while differences can reach an order of magnitude or more for less massive objects (Muldrew et al. 2011; Knebe et al. 2011, 2013). Secondly, there are no universal ways of defining the edge of sub-structures and how to treat them – either as completely separated objects or contributing to the properties of the host halo. Possible improvements in this direction include the analysis of the full time evolution of the structures and not focusing only on a single snapshot. Lastly, even if different halo finders identify the same set of particles belonging to a halo, there exist different possibilities for how to derive its properties, depending on the assumed initial conditions and on the mass/edge definitions previously mentioned.

For the purpose of astrophysical applications, namely DM indirect detection, the derivation of the DM density spatial profile of the identified halos is typically imposed to follow some analytical distribution. For sub-halos, this often matches the best-fitting parametrisation of the parent halo, although past works showed, instead, that halo shapes can depend significantly on the halo mass (Gao et al. 2008). Also, sub-halos go through remarkably different evolution histories than their field host, and may suffer heavy tidal stripping which strongly affects the DM distribution, see e.g. (Diemand et al. 2007b; van den Bosch et al. 2018; van den Bosch & Ogiya 2018; Green & van den Bosch 2019).

These assumptions may strongly impact DM indirect detection signatures.

The present work builds on the following question: can we bypass the use of halo finders and derive DM indirect detection relevant quantities directly from simulation raw data? Although we fully acknowledge that halo finders provide an extremely powerful tool to analyse sub-halo populations properties and characterising their evolution through cosmic time, they may not be an indispensable tool to extract predictions of DM indirect signals. To this end, we design HADES, a “Halo Accurate Density Evaluation System”, that proposes a first step towards the accomplishment of such a delicate task.

The paper is organised as follows. We introduce the HADES algorithm in Section 2, namely how it evaluates the local mass density in cosmological simulations. In Section 3, we describe the realisation of a set of Monte Carlo simulations of galactic halos which we will first use to test HADES. We apply HADES to both the host halos and sub-halos population of our simulated catalogue in Section 4. Section 5 summarises the computation of the so-called J-factor, relevant for the application of HADES to the definition of DM indirect detection signals. Finally, we present a realistic application of HADES to an existing cosmological simulation suite in Section 6, and we draw our conclusions in Section 7.
2 THE HADES METHOD

HADES is a nearest-neighbour algorithm to estimate the local density in cosmological simulations. It is based on the simple idea that, given a discrete distribution of point-like particles with a given mass, the density around each point can be computed associating to the particles an occupation volume. In a uniform distribution, i.e. a distribution in which the density is constant everywhere, each particle is ideally associated to a same occupation volume $dV$, and the total volume of the simulation is given by the sum of all the infinitesimal volumes corresponding to different particles. In a real-life situation, where the density is not uniform, even the localisation volumes will change accordingly, being smaller in regions where the density increases and larger in lower-density zones.

We assume the infinitesimal volume $dV$ to be equivalent to a cube of side $d$, thus $d^3 = dV$, where $d$ can be regarded as an estimate of the average separation between a particle and its nearest neighbours. The determination of $dV$ assigns to each particle its localisation volume in a given region, and allows to compute the local density as $\rho = m_p/d^3$, $m_p$ being the particle mass. The main steps of the HADES algorithm are the following:

(i) Starting from the position $\bar{x}$ at which we wish to compute the density, it finds the positions $r_i^n$, $n = (1, \ldots, N)$, of the $N$ particles closest to $\bar{x}$.

(ii) For each particle at $r_i^n$, it computes the distance to the $j$ closest particles – located at $r_{ij}^n$ – and averages to find $d_j^n = \frac{1}{j} \sum_{i=1}^j |r_i^n - r_{ij}^n|$.

(iii) Step (ii) is repeated several times, varying $j$ from a minimum value $a$ to a maximum value $b$, then it averages again to obtain $d^n = \frac{1}{b-a} \sum_{j=a}^b d_j^n$.

(iv) At this point, it has an estimate of the mean distance $d^n$ between each particle and its nearest neighbour, and it averages one more time to find the mean virtual side in that simulation region: $d = \frac{1}{N} \sum_{n=1}^N d^n$.

(v) Finally it computes the density as $\rho(\bar{x}) = m_p/d^3(\bar{x})$.

Eventually, the mass density at $\bar{x}$ reads:

$$\rho(\bar{x}) = m_p \left[ \frac{1}{N} \sum_{n=1}^N \frac{1}{b-a} \sum_{j=a}^b \sum_{i=1}^j \left| r_i^n - r_{ij}^n \right| \right]^{-3}. \quad (1)$$

A few comments on this definition are in order. First of all, in writing a suitable algorithm we introduced three separate averages. These averages stabilise HADES response to small fluctuations in the particles positions, increasing the scale at which we measure density deviations. This scale is eventually fixed by the free parameter $N$ – and partially by the two parameters $a$ and $b$ – and defines the locality of the estimate: the smaller the value of $N$, the more local HADES result will be. We found that values of $N$ that keep the density estimation sufficiently stable, but local enough to allow for the resolution of structures in cosmological simulations, range in the interval $[10, 20]$. In the rest of our study we then restrict to $N = 15$.

The remaining two free parameters, the couple $(a, b)$, are instead a measure of the amount of nearest neighbours around each particle. Specifically, the couple $(a, b)$ tells us that in the region under consideration particles can have a number of nearest neighbours that ranges from $a$ to $b$. Very high values for $a$ and $b$ can eventually end up affecting the locality of the measure as well. Fixing these values $a$ priori from theoretical considerations is tricky, as the distance of a particle to the surrounding ones increases gradually, and there is no clear distinction between those particles that can be taken as nearest neighbours and those which are too far away to be considered for the density estimation. Hence, the couple $(a, b)$ needs to be fixed in a phenomenological way. An accurate description of how we made this choice is detailed in Appendix A. Here we only report that the best parameters selection for a quite general range of density distributions – including mock realisations of galactic halos – is found to be $a = 6$ and $b = 13$. These values are kept fixed in this work, unless explicitly specified. The three parameters $N, a$ and $b$ are then hyper-parameters (HP) in our method. We can highlight that this is a preliminary set-up but that we envision making the optimisation of the HP better suited for generic simulations and automatised. This will go hand in hand with speed optimisation and improvements to reach even higher accuracy. HADES was developed on test-case mock simulations that we describe in Sec. 3.

3 MONTE CARLO SIMULATION OF GALACTIC HALOS: THE DRAGON CATALOGUE

Before launching HADES on real cosmological simulations, we produced for testing purposes Monte Carlo (MC) realizations of galactic halos that mimic cosmological simulations outputs. As the procedure by which these halos were produced is important to understand some of the results of Sections 4 and 5, we provide here the essential details. We relied on results of existing N-body cosmological simulations to set the prescriptions for the mass distribution of the main host halo, together with the number, space and mass distributions of the subhalos population. We created a complete MC mock catalogue, that we dub the Dragon catalogue. We focused only on reproducing the particles’ positions (neglecting their velocities), as it is all the information that we need to compute the density with HADES. All halos in the Dragon catalogue are marked by the code ‘Dr-’ followed by a progressive number (from 1 to 9), as detailed below.

3.1 The main halo

For the sake of brevity, we have chosen to restrict our study to the Einasto (E) and the NFW density profiles, and to the cored Burkert (B) profile, all obeying a spherical mass distribution. Their parametrisations follow:

$$\rho_E(r) = \rho_{\text{iso}} \cdot \exp \left( \frac{2}{\alpha} \left( \frac{r}{r_{\text{iso}}} \right)^\alpha - 1 \right),$$

$$\rho_{\text{NFW}}(r) = \frac{4\rho_{\text{iso}}}{\left( \frac{r}{r_{\text{iso}}} \right) \left( 1 + \frac{r}{r_{\text{iso}}} \right)^2},$$

$$\rho_B(r) = \frac{4\rho_{\text{iso}}}{\left( 1 + \frac{r}{r_{\text{iso}}} \right) \left( 1 + \frac{r}{r_{\text{iso}}} \right)^2}. \quad (2c)$$

Here, $r$ is the radial distance from the centre of the halo. The chosen analytical profiles have been usefully re-parametrised as done in (Navarro et al. 2010), so that $r_{\text{iso}}, \rho_{\text{iso}}$ and $\alpha$ are free parameters representing, respectively: the radius at which the logarithmic slope of the profile equals the isothermal value, the corresponding density $\rho(r_{\text{iso}})$, and a slope parameter typical of the Einasto parametrisation that regulates the steepness of the profile. In what follows, we have kept $\alpha$ fixed at 0.17, while varying the remaining two parameters. The chosen value for $\alpha$ was taken close to those found in (Springel et al. 2008; Navarro et al. 2004, 2010).

The choice of using only spherically symmetric profiles in phase of testing is not restrictive, as we aim at very local density estimates with
Table 1. Relevant parameters for the Monte Carlo-simulated main halos. \( \rho_{-2} \) is the density profile overall normalization as computed from eq. (3). \( N_{\text{tot}} \) is the total number of particles simulated within the maximum radius \( r = 300 \) kpc, and \( M_{\text{tot}} \) the corresponding total mass.

| Name    | Profile | \( \rho_{-2} \) [M$_{\odot}$/kpc$^3$] | \( N_{\text{tot}} \) | \( M_{\text{tot}} \) [M$_{\odot}$] |
|---------|---------|-----------------------------------|-----------------|-----------------|
| Dr-1-3  | Einasto | 2.94 $\cdot$ 10$^6$               | 112435783       | 1.12 $\cdot$ 10$^{12}$ |
| Dr-4+6  | NFW     | 3.14 $\cdot$ 10$^6$               | 114054823       | 1.14 $\cdot$ 10$^{12}$ |
| Dr-7+9  | Burkert | 2.89 $\cdot$ 10$^6$               | 113674381       | 1.14 $\cdot$ 10$^{12}$ |

HADES. The overall symmetries of the system, though convenient, will not impact the results. On the contrary, we will show in Section 6 that HADES can efficiently detect and map out deviations from sphericity when applied to simulated halos.

It also typically comes more handy to use the total mass of the halo as a free parameter, from which – for a given \( r_{-2} \) – one can compute the overall normalisation \( \rho_{-2} \) by simply requiring that the volume integral of the mass profile over the whole halo yields back the total mass. This condition explicitly translates into the following prescriptions:

\[
\rho_{-2}^E = \frac{M_{200}}{4\pi \int_0^{r_{200}} dr r^2 \exp \left( \frac{r^2}{2\sigma^2} \right) - 1} ,
\]

\[
\rho_{-2}^{\text{NFW}} = \frac{M_{200}}{16\pi r_{-2}^3 \ln \left( 1 + \frac{r_{200}}{r_{-2}} \right) + \frac{1}{1+(r_{200}/r_{-2})} - 1} ,
\]

\[
\rho_{-2}^{B} = \frac{8\pi r_{-2}^3}{M_{200}} \left( \frac{1}{2} \ln \left( 1 + \frac{r_{200}}{r_{-2}} \right) ^2 - \arctan \left( \frac{r_{200}}{r_{-2}} \right) + \ln \left( 1 + \frac{r_{200}}{r_{-2}} \right) \right) ,
\]

where the integral in the denominator, in case of an Einasto profile, can only be computed analytically if \( r_{200} \to \infty \). In general, \( r_{200} \) should represent a radius containing a significant fraction of the halo mass, the corresponding \( M_{200} \). We chose to fix \( r_{200} \) at the virial radius, defined as the radius enclosing an average mass density equal to 200 times the critical density of the Universe today, \( r_{200} = (3M_{200}/800\rho_c)^{1/3} \). For all the main halos in the Dragon catalogue we fixed \( r_{-2} = 15.14 \) kpc and assumed a \( \Lambda \)CDM cosmological model with \( \rho_c \approx 148 \) M$_{\odot}$/kpc$^3$. We finally wish to stress that to assess HADES performance the specific values taken by these parameters are not fundamental, as long as the algorithm is efficient in reconstructing the correspondent density distribution. Our final Dragon catalogue contains 9 halos, three realisations for each mass distribution in eq. (2). Dr-1-3 follow an Einasto profile, Dr-4+6 are built on a NFW profile, Dr-7-9 present a Burkert profile.

We now shortly outline the procedure by which we simulated each mock halo. We performed the simulation in spherical coordinates, with the origin at the centre of the halo. For each particle we drew the coordinates \( \hat{x} = (r, \theta, \phi) \), where \( \theta \) and \( \phi \) are respectively latitude and longitude. The angular coordinates are randomly sampled from a uniform spherical distribution; \( \phi \) was extracted in the interval \([0, 2\pi] \), while the latitude was computed as \( \theta = \arccos(t) \), with \( t \in [-1, 1] \). This last trick avoids clustering of particles at the sphere poles.

Sampling the radial coordinate is more delicate. We started by computing the number of particles that should be present at any given radius to match the theoretical density profile. We dubbed this quantity \( N(r) \). More precisely,

\[
\rho(r) = \frac{N(r_{i+1} - r_{i})}{4\pi r_i^2 dr} \rightarrow \rho(r) = \lim_{dr \to 0} \frac{N(r_{i})}{4\pi r_{i}^2 dr} ,
\]

where the first equation refers to a discrete realisation of the continuous mass distribution \( \rho(r) \). \( N(r_{i+1} - r_{i}) \) is then the number of particles of mass \( m_p \) within a spherical shell of radius \( r_i \) and thickness \( dr \). Hence, we proceeded to divide the halo into nested spherical shells, ideally thin enough to resemble as much as possible a continuous distribution, but still thick enough that \( N(r_{i}) \geq 1 \) for all \( i \). The particles’ radial coordinates were then sampled from the \( N(r_{i}) \) discretised distribution. This process yields a new distribution \( N(r_{i})_{\text{MC}} \) which matches the initial one barring some small fluctuations in the form of a poissonian noise \( \Delta N(r_{i}) = \sqrt{N(r_{i})} \). We further smoothed out our simulations by re-drawing the radius of all \( N(r_{i})_{\text{MC}} \) particles in the \( i \)th shell from the interval \([r_i - dr/2, r_i + dr/2] \).

We simulated all mock halos in our catalogue up to \( r_{\text{max}} = 300 \) kpc, setting the total virial mass at \( 10^{12} \) M$_{\odot}$ and requiring that at least \( 10^8 \) particles are contained within the virial radius. These requests determine the resolution of our simulations, fixing the mass of each particle at \( m_p = 10^4 \) M$_{\odot}$. This resolution is close to those of current cosmological simulations (see e.g. Aquarius (Springel et al. (2008)), Via Lactea II (Diemand et al. (2008)), TNG50 (Pillepich et al. (2019b))) and it allows us to resolve a few thousands sub-halos, as we will discuss shortly. We point out anyway that one of the advantages of HADES is adaptability, in the sense that increasing the resolution simply allows HADES to “look closer” and make its density estimates more precise. The
relevant parameters of our MC halos are summed up in Table 1.

In Fig. 1 we plot the comparison between \( N(r_i) \) and \( N(r_{i,MC}) \) for the Dr-I halo. The two profiles are in good agreement, the biggest discrepancies occurring at the outskirts and in the centre below approximately \( r = 0.1 \) kpc, where the number of particles is smaller. We analysed the same plot for all our MC halos, and found that these differences are slightly higher for Burkert halos, as such profiles exhibit a lower central density with respect to the cuspy ones, predicting even less particles in the central regions. However, all mismatches are contained within twice the predicted poissonian noise, hence we are satisfied that our MC halos are accurate enough reproduction of the input mass distribution.

### 3.2 The sub-halos

To populate each smooth halo with its sub-halos, we begun by picking the population statistics, i.e. the two distributions describing the sub-structures spatial positions in the host halos and their masses. We model our sub-halo distributions on those found in the second run of the Via Lactea II (hereafter VLII) simulation suite (Diemand et al. 2008), exploiting the data of their public catalogue \(^1\). Halos and sub-halos in VLII were identified with the 6DFOF halo finder (Diemand et al. 2006). Of all 20048 sub-structures present in the catalogue, we selected only the 3472 that are located within a radius \( 10 \leq r \leq 250 \) kpc from the main halo centre and have masses higher than \( 10^6 \) M\(_\odot\) (i.e. containing at least 100 particles at the resolution of our MC halos).

It has been shown that the spatial distribution of sub-halos in simulations can be approximated with a radial function parametrised similarly to an Einasto profile (see e.g. (Zhu et al. 2016) and (Calore et al. 2017)):

\[
\frac{n(r_{MH})}{n_{tot}} = n_{-2} \cdot \exp \left\{ -\frac{2}{\alpha_n} \left[ \left( \frac{r_{MH}}{r_{-2}} \right)^{\alpha_n} - 1 \right] \right\},
\]

where \( r_{MH} \) is the distance from the main halo centre, \( n(r_{MH}) \) is the sub-halos number density at given radial bin and \( n_{tot} \) is the total number of structures in the sample. The parameters \( \alpha_n \), \( n_{-2} \) and \( r_{-2} \) were fitted against the reduced VLII catalogue, yielding \( \alpha_n = 1.21 \pm 0.16, n_{-2} = 2.82 \pm 0.04 \) and \( r_{-2} = (108 \pm 2) \) kpc. The results are shown in Fig. 2 (left panel).

The mass distribution of sub-structures is also well know (see for example (Zhu et al. 2016; Calore et al. 2017; Springel et al. 2008)).

\(^1\) https://www.ucolick.org/diemand/vl/index.html

---

**Figure 2.** Fits of the radial number density (left panel) and differential mass abundance (right panel) of sub-halos on the reduced catalogue from VLII. \( r_{MH} \) refers to the position of the sub-halo with respect to the main halo centre.

**Figure 3.** The right panel shows the fit of the phenomenological correlation \( r_{max} - M_{sh} \) on the sub-structures from the reduced VLII catalogue. The colour scale indicates the number of particles that would be contained in each sub-halo at our MC simulation resolution. The left panel displays the scatter plot in the \( r_{max} - M_{sh} \) plane of the Dr-I halo sub-halos’ population compared to the VLII one.

**Figure 4.** Theoretical number counts \( N(r) \) (blue line) compared to the simulated one \( N(r)_{MC} \) (black line) for sub-halos spanning the explored mass range. The shaded blue area marks twice the poissonian error.

These authors all found that the differential mass abundance of sub-halos in cosmological simulations is well reproduced by a power-law of the form

\[
\frac{dN}{dM} = M_0 \left( \frac{M}{M_\odot} \right)^{-\alpha_M}.
\]

Eq. (6) was also fitted on the reduced VLII catalogue, and the results are plotted in Fig. 2 (right panel). We obtained \( M_0 = (0.98 \pm 0.22) \cdot 10^9 \) M\(_\odot\) and \( \alpha_M = 1.9 \pm 0.1 \), this last value in particular in agreement with those found in (Springel et al. 2008) and (Zhu et al. 2016) for the halos of the Aquarius suite.

Moreover, Calore et al. (2017) found for the C-halo of Aquarius a correlation between the sub-halo masses and the radius at which the structure’s particle velocities peak, \( r_{max} \). In addition, Springel et al. (2008) showed that there is a phenomenological relation connecting \( r_{max} \) to the halo scale radius: \( r_{max} \approx 2.189 r_{-2} \).

The combination of these results gives

\[
\log_{10} \left( \frac{r_{-2}}{\text{kpc}} \right) = p_0 - \log(2.189) + p_1 \log_{10} \left( \frac{M_{sh}}{M_\odot} \right) + p_2 \left( \log_{10} \left( \frac{M_{sh}}{M_\odot} \right) \right)^2.
\]

Once again, we fitted eq. (7) on the sample of sub-halos from...
VLII and verified that this correlation still holds. We obtained $p_0 = -4.840$, $p_1 = 0.898$ and $p_2 = -0.037$. We additionally included in our model the population scatter around the best fit through the relative standard deviation $\sigma_{r_{-2}}/r_{-2} = 0.237$. The results are presented in Fig. 3 (right panel).

The left panel of Fig. 3 shows instead the comparison between the sub-halo population in the plane $r_{\text{max}} - M_{\text{tot}}$ obtained for Dr-1 and the analogous population of Via Lactea II. The agreement between the two is very high.

With all the ingredients in place, a population of sub-structures was added to every mock halo: i) the mass was sampled from the distribution (6) in the interval $10^6 \leq M_{\text{dm}} \leq 10^9 M_\odot$ and $r_{-2}$ was computed using (7) and accounting for scattering; ii) the structure was placed in the main halo by sampling eq. (5); iii) each sub-halo was simulated as in Section 3.1 to reflect the same mass profile of its host, with normalisation $r_{-2}$ determined from its total mass as in eqs. (3). We restricted to sub-halos with masses higher than $10^6 M_\odot$ so that all Dragon sub-structures contain between $10^2$ and $10^5$ particles.

In Fig. 4 we plot the particle number counts $N_{\text{MC}}(r_i)$ compared to the theoretical $N(r)$ for four examples of MC sub-halos spanning the full mass range. As sub-halos contain way less particles that their host halos, the numerical poissonian noise has a higher impact, so mismatches between $N_{\text{MC}}(r)$ and $N(r)$ are more relevant – especially, once again, in the inner regions where the number of particles is lowest. In particular for sub-halos of masses $10^8 M_\odot$, that contain a total of only 100 particles, a fine reproductions of the input $N(r)$ is evidently extremely difficult.

Finally, Fig. 5 shows the surface density map $\rho_s = M_{\text{tot}}/A_{\text{pix}}$ of three representative Dragon halos, $A_{\text{pix}}$ being the area of each pixel and $M_{\text{tot}}$ the total mass present in the corresponding column volume perpendicular to the projection plane.

4 HADES TO THE TEST: THE DRAGON CATALOGUE

With this work, we developed a numerical estimator aiming at accurate computations of the local density in cosmological simulations.
course validates our previous conclusion that the MC process is introducing a level of random scatter in the particles positions, but the mock halos are otherwise well realised: all regions of the halo seems to be equally affected by the same fluctuations.

It is however important to keep in mind that the results we reported for $\sigma_p/\rho$ are not universal, on the contrary, in general they do not apply to halos in real cosmological simulations. Indeed, simulated halos will also display some level of local fluctuations of their DM density, but there are endless physical effects that concur to their origin: deviations from spherical symmetry of the main halo mass distribution, residuals of disrupted sub-structures, structures resulting from previous mergers with other halos and not yet fully smoothed out, or an unresolved population of sub-halos, just to name the most well-known (see e.g. (Zavala & Frenk 2019) for a review). Each of these factors impacts the relative mass dispersion in various regions of the halos in different ways. Thus, in general, $\sigma_p/\rho$ can exhibit the most disparate trends, encapsulating important physical information on the halo’s formation history.

In Fig. 7 we plot the average $\rho_{\text{HADeS}}$ compared to the theoretical Einasto profile $\rho_b$ followed by the Dr-1 main halo. The right panel shows the relative difference $(\rho_{\text{HADeS}} - \rho_b)/\rho_b$, with the shaded regions marking the 1$\sigma$ and 3$\sigma$ dispersion of $\rho_{\text{HADeS}}$. Being errors over a mean value, such bands were obtained dividing the relative dispersion of the density field along the radial profile – see Fig. 6 – by $\sqrt{n}$, $n$ being the number of LOS used to compute $\rho_{\text{HADeS}}$. Deviations of the mean density from the analytical profile are centred around 0 and are included in the small numerical error, a good sign that HADeS is doing a fine job of properly reconstructing the local density.

We wish to stress here that these confidence regions are not to be mistaken as an estimate of the error made by HADeS itself in computing the density. The $\sigma$ of Figs. 6 and 7 represent a density dispersion that receives two contributions: (i) the algorithm intrinsic numerical error and (ii) local fluctuations, i.e. significant clustering/thinning of particles that change the local value of the density. Such fluctuations arise naturally from the MC process and are variations that we actually wish to detect with an accurate density estimator designed to spot sub-structures. We refer the reader to Appendix B for the full discussion on characterising HADeS numerical error. Here, we just report our main finding that HADeS uncertainty contributes only to approximately 1% of the 20% sigma dispersion that we measured in Fig. 6, which seems instead mostly dominated by local density variations.

4.2 The sub-halos

We devote this section to analyse the effects of sub-halos on HADeS estimates. Again, we randomly drew $10^5$ LOS in every smooth halo, and placed on each LOS the MC realisation of a sub-halo with fixed mass and same density profile of its host. The centre of the sub-halo falls on the LOS at $r_{\text{MH}} = 120$ kpc. As we highlighted in Section 3, the MC realisation of a sub-structure is a delicate procedure: especially when modelling the centres of the smallest halos, little variations in the number of simulated particles falling in those regions can produce very different outcomes. The purpose of re-simulating the sub-halo for each LOS is to account in our tests for this halo-to-halo scatter. Once again, we run HADeS to compute the DM density over all LOS for each Dragon halo. We will use the resulting data-set to make a number of considerations.

Firstly, Figs. 8 and 9 (left panels) show the density profiles as reconstructed by HADeS, zooming in on a selection of sub-halos with masses $10^6$ and $10^7 M_\odot$, for both Einasto and Burkert profiles. Sub-structures with an NFW profile present similar results to the Einasto case, and are not shown here. HADeS performance on high-mass sub-structures strikes as remarkably good: we found that for all sub-halos of masses $10^6 - 10^7 M_\odot$ the density in the outskirts and the central regions is reconstructed efficiently. There are only small discrepancies between HADeS computation and the theoretical profile, as can be seen in the right panels of Fig. 8, where we plot the ratio $(\rho_{\text{LOS}} - \rho_b)/\rho_b$. The blue and cyan lines mark respectively the 1$\sigma$ and 3$\sigma$ confidence regions, where $\sigma$ is the density scatter around the analytical profiles that we measured for the main halo in Fig. 6. We see clearly that fluctuations remain of the same order.
of those already observed in the main halo, hence imputable to MC effects rather than to HADES systematics. A similar result holds for masses down to \(10^7 M_\odot\). However, for sub-halos of masses lower than \(10^6 M_\odot\) (see Fig. 9) HADES overestimates the density profile in the outskirts, while dramatically underestimating the centres. At the same time, \(\delta \rho/\rho_{\text{th}}\) explodes at the position of the halo. This is due to concurrent numerical effects worsening HADES performance on these objects, as we will discuss shortly.

Moreover, note that for the plots of Fig. 9 we performed the comparison only down to \(r_{\text{sh}} \geq 0.05\) kpc, \(r_{\text{sh}}\) being the distance from the sub-structure centre. We will comment more on this choice at the end of this Section.

To better characterise the changes in the density dispersion when a substructure is present, we re-computed the profiles \(\sigma_p/\rho(r)\) with this new data-set. For bigger sub-halos (\(10^9 - 10^8 M_\odot\)), as expected, the flat behaviour of \(\sigma_p/\rho\) does not change significantly. However, a spike starts appearing in the dispersion profiles of halos+sub-halos of masses \(10^7 - 10^6 M_\odot\) and lower, as shown in the second row of Fig. 6 for the Dr-1 halo.

A thorough scrutiny of the density profiles computed over all LOS piercing a medium mass halo (\(\sim 10^7 M_\odot\)) highlighted that in a limited number of cases HADES estimates densities in the central regions of the structure that deviates significantly from the expected analytical profiles. This behaviour is likely the result of halo-to-halo scatter combined with the MC procedure pitfalls in halo centres.

To clarify this point, we recall that in the MC simulation, each sub-halo particle is randomly assigned to a shell according to \(N(r)\). This produces an effective distribution \(N_{\text{MC}}(r)\) mirroring the theoretical one. Eventually, in the \(i^{th}\) shell, \(N_{\text{MC}}(r_i)\) particles are uniformly scattered, and then the finalised sub-structure is placed in the main halo.

Now, when \(N_{\text{MC}}(r)\) is sampled from \(N(r)\), mismatches between the two appear in the form of a poissonian noise, as we showed in Section 3. This noise in turn impacts the realisation of the final halo, as at small radii \(N(r_i)\) can be very low. It follows that only a handful of particles will be designated to the corresponding shells. If, say, \(N(r_i)\) requires 5 particles, a difference of just 2 during the random extraction process already introduces an error on that shell’s average density of 40\%. Even the steepest profiles of eq. (2) predict less than 5 particles in the inner shells for low mass sub-halos, as shrinking the mass implies decreasing the overall profile normalisation, exploring lower density structures. As a result, realisations of sub-halos in the mass range \(10^7 - 10^6 M_\odot\) or lower are significantly affected by the MC uncertainties. Indeed, scanning all small sub-halos simulated for these tests we found a correlation between HADES under(over)-estimates of the central density peak and the respective discrepancy between \(N\) and \(N_{\text{MC}}\). This validates the hypothesis that the density misestimates cannot be attributed exclusively to HADES systematics, as, to begin with, the mock sub-halos themselves do not match well the theoretical mass profile due to MC uncertainties.

The spikes in the second row of Fig. 6 are now easily explained: if the MC realisations of sub-halos centres is subjected to high numerical fluctuations, this reflects automatically in a higher density scatter around the average value for these samples. But again, it does not mean that HADES is lacking efficiency in computing the densities for these halos. Indeed, the average numerical \(\rho(r)\) per se still matches the theoretical predictions extremely well – where the average is performed over all \(10^3\) LOS, and hence \(10^3\) different realisations of same-mass sub-halos. This means that the increase in variance is not accompanied by a systematic under(over)-estimate of the density, which would indeed be a sign of HADES poor performance.

The case of very low-mass halos (\(M_{\odot} \ll 10^6 M_\odot\)) is still different. For these sub-structures, the density is so low that \(N(r)\) drops below 1 for all innermost shells. As soon as this happens no particles at all are simulated in those regions. In turn, an excess of particles is deposited in the outer shells, with the double outcome of under-simulating the central density and over-simulating the peripheral one. We recall however that, at our resolution, the total number of particles for these halos is \(\sim 100\). It is not surprising then that the particles
are not enough for an accurate reproduction of a complex 3D mass distribution.

The recovery of the input analytical profile on these halos is very poor, as can be seen in Fig. 9. Nevertheless, the non-optimal reconstruction of low-mass sub-halos is not a major issue, as the main focus of HADES development is its use in producing predictions of DM indirect detection signals, for instance the annihilation gamma-ray flux in sub-structures. Given the sensitivity of current detectors, only the brightest, most massive, sub-halos are expected to be able to produce a signal high enough to be actually seen as a gamma-ray source (Calore et al. 2019; Grand & White 2021).

Lastly, we comment on the lower radial cut-off $r_{th} \geq 0.05$ kpc imposed when plotting the profiles comparisons in Figs. 8 and 9. Fig. 10 shows the profiles without cut-off for halo+sub-halo of $10^9 M_\odot$ following the same mass distribution, Einasto and Burkert respectively. While the cored profile is still overall well recovered by HADES, the cuspy analytical profile differs significantly from its numerical counterpart in the inner 50 pc, steeply reaching a peak 3 orders of magnitude higher.

Again, the limited resolution of discrete points distributions explains the observed mismatch. Indeed, it is impossible to reproduce a continuous density distribution with a finite number of particles down to arbitrary small radii. There always is a lower resolution radius at which $N(r)$ becomes lower than 1 and no particles are simulated anymore. As a result, not even in real cosmological simulations the pattern of DM particles can match any steep analytical profile down to the innermost regions of each halo, even setting aside the issues linked with numerical convergence of simulations at low radii.

Conceptually, this has important consequences. Using the profiles of eq. (2) to study sub-structures requires making an assumption on the density behaviour in the innermost regions of the halo, and then extrapolating such profiles down to $r_{th} = 0$. On the other hand though, most of these profiles are insanely steep, spanning over 3 orders of magnitude in density inside the minimum convergence radius of the halo. Hence, the cusps are not observable in cosmological simulations. Sure enough, it does not necessarily mean that they are not realistic ansatz. The robustness of their assumption is often based on the argument that the host halos, for which higher resolution allow a deeper exploration of the innermost regions, seem to be accurately described by profiles such as those of eq. (2). At the same time though, sub-halos are more concentrated than their hosts, and hence the profiles describing their density are assumed to be remarkably steeper than those of bigger halos. The accuracy of the predicted cusps then has long been a subject of debate.

In this scenario, HADES plays a very conservative role, estimating the density only based on the particles that are actually present in the halos and not relying on additional assumptions.

5 J-FACTOR COMPUTATION ON DRAGON HALOS

To show the potentialities of HADES application to DM indirect probes, we devote this Section to the computation of the so-called J-factor for our mock halos, which accounts for the information on DM morphology when calculating DM annihilation fluxes:

$$J = \int_{\text{LOS}} \rho^2(r) dI(r) = \int_0^{R_{\text{max}}} \rho^2(r) dr.$$  \hspace{1cm} (8)

Here $\psi$ is the angle between the direction of the halo centre and that of observation. We stress that, though we will refer to $J$ as ‘J-factor’, these computations are performed on mock MC halos created ad hoc for testing purposes only. Hence, the $J$ values that we present are not yet to be considered as robust predictions on real signals, as will be discussed also in section 6. Future works will be devoted to the extensive application of HADES to cosmological simulations and forecasts on annihilation fluxes. In this section in particular, we test the algorithm comparing the J-factors computed running HADES with the theoretical ones – obtained by integrating the analytical profiles of eq. (2). For clarity, we will indicate the theoretical quantities with the subscript $\text{th}$ (e.g. $\rho_{\text{th}}, J_{\text{th}}$). $\rho$ and $J$ will instead be referred to HADES’ computations, unless explicitly specified. For the definition of the LOS, here and in the rest of the paper we placed the observer at a distance $r_{\text{MH}} = 8$ kpc from the main halo centre, and referred all radial distances $r$ to this observer.

The difference between using $\rho$ and $\rho_{\text{th}}$ for the J-factor is subtle, but substantial. When computing $J$ through HADES, the density that enters eq. (8) is a local density, that will in general vary if one considers the same LOS in a different realisation of the same halo. The analytical density that gives $J_{\text{th}}$ instead is an averaged density, $\rho_{\text{th}} = \langle \rho \rangle$, where all average symbols hereafter are intended over different realisations of the same LOS.

One might be tempted to assume that the averaged $J$ should also match $J_{\text{th}}$, and in this perspective it becomes interesting to compute the difference between the two. We can write:

$$\delta J = \frac{J - J_{\text{th}}}{J_{\text{th}}} = \frac{\langle \int_{\text{LOS}} \rho^2(r, \psi) dr \rangle - \langle \int_{\text{LOS}} \rho_{\text{th}}^2(r, \psi) dr \rangle}{\langle \int_{\text{LOS}} \rho_{\text{th}}^2(r, \psi) dr \rangle}$$

$$= \frac{\int_{\text{LOS}} \rho^2(r, \psi) dr - \langle \rho^2(r, \psi) \rangle dr}{\int_{\text{LOS}} \rho_{\text{th}}^2(r, \psi) dr} = \frac{\int_{\text{LOS}} (\rho^2(r, \psi) - \langle \rho^2(r, \psi) \rangle) dr}{\int_{\text{LOS}} \rho_{\text{th}}^2(r, \psi) dr},$$  \hspace{1cm} (9)

from which we conclude that there is an offset to be expected between $(J)$ and $J_{\text{th}}$. In particular, the numerical J-factors is, on average, higher than the theoretical value. In computing $J$, we keep track of local fluctuations in the density, so the average of the squared density is higher than the square of the average density. This means that using a smooth analytical profile is substantially different from evaluating the peculiar density point-by-point along the LOS. A similar phenomenon is already known in the literature as boosting (Kuhlen et al. 2012). Boosting occurs typically when an unresolved population of sub-structures is present in the halo. This unresolved population increases the graining of the halo – and hence, the density dispersion
\[ \frac{\sigma}{\rho} - \text{and can lead to an enhancement of the 'true' J-factor with respect to } J_{th}. \]

We now use the result of Section 4 that, for smooth halos, the relative dispersion \( \sigma_{\rho}/\rho_{th} \) remains approximately the same in all regions of the halo. We can then push the computation of eq. (9) a little further and write:

\[ \delta J = J_{th} - \int_{M_{\text{LOS}}} \frac{\sigma_{\rho}^2(y, \phi)}{\rho_{th}^2} dy = \int_{M_{\text{LOS}}} \frac{\sigma_{\rho}^2}{\rho_{th}^2} dy = \frac{\sigma_{\rho}^2}{\rho_{th}^2}. \]  

As we found \( \sigma_{\rho}/\rho_{th} \approx 0.194 \), the J-factor offset can be estimated as \( \delta J_{th} = (0.194)^2 \approx 3.76\% \).

We validate this result by choosing a random LOS in each smooth halo in the Dragon catalogue, and computing the J-factor offset for \( 10^3 \) different realisations of such LOS. In practice, since simulating the whole halo \( 10^3 \) times is extremely expensive, we exploit the spherical symmetry of our halos: we observe that to all LOS identified by a given angle \( \psi \) corresponds the same averaged density profile \( \rho_{th}(r, \psi) \). They can then be considered as different realisations of the same LOS since the local density along each LOS will in general be different due to its different location inside the halo.

In the first row of Fig. 11 are shown the histograms of the \( 10^3 \) J-factor values obtained for the halos Dr-1 and Dr-7 (Einasto and Burkert density profiles respectively). The colour scale marks the single bin offset \( (J_{bin} - J_{th})/J_{th} \), while the mean value \( \langle J \rangle \) and \( J_{th} \) are marked in the plots with vertical dashed black and solid red lines respectively. The average offset value is indicated as \( \delta J/J_{th} \) and amounts to 3.44\% and 3.11\%, in reasonably good agreement with the predicted 3.76\%.

Though a difference between \( \langle J \rangle \) and \( J_{th} \) of \( \approx 3 - 4\% \) can seem a negligible effect compared to the speed and simplicity of modelling DM halos with analytical profiles, the reader should keep in mind an important result of Section 4: the values of \( \sigma_{\rho}/\rho \) presented here do not hold in general for all halos in cosmological simulations. There are numerous physical effects contributing to the graininess of simulated halos that, depending on their formation history, can easily enhance the density scatter up to much higher values. Indeed, the boosting of the J-factor signal occurring in real halos has been recently estimated to reach up to a factor of \( 2 - 3 \) (Bartels \& Ando 2015; Grand \& White 2021). If \( J \) is computed with analytical profiles, all boosting effects need to be modelled and added independently after the calculation. In this context, exploiting a tool like HADES allows to automatically account for boosting without requiring separate modelling.

In the second and third row of Fig. 11 are again reported J-factor histograms obtained computing \( J \) with HADES over 1000 realisations of LOS in the main halo. Each LOS in turns passes through the centre of a different realisation of the same sub-halo (i.e. fixed mass and profile, but re-simulated for each LOS). We repeated the computation for various sub-halo masses (\( 10^5 M_\odot \) and \( 10^7 M_\odot \) in the left and right column respectively) and for each mass distribution profile (Einasto in the second row and Burkert in the third row). We still notice the offset \( \delta J \) between the average J-factor and the theoretical value, however, such offset is now affected by the presence of the sub-structure, depending on its mass and profile.

Here, we see the consequences of HADES slightly underestimating the sub-halos central densities. This underestimate counteracts the signal boosting we described earlier: its effect is to reduce \( \langle J \rangle \) and hence \( \delta J \). However, for halos in the high mass range (\( 10^8 - 10^9 M_\odot \)) its impact is very small: comparing the values of \( \delta J \) with and without sub-structures we infer that when introducing high-mass sub-halos along the LOS \( \delta J \) is reduced of only \( \approx 0.2\% \) with respect to the smooth case, for cuspy profiles. Core profiles are more affected due to the lower number of particles that can be found in the centre of sub-structures, with a \( \delta J \) reduction of \( \approx 1.5\% \) for Burkert sub-halos.

This is consistent with the consideration of Section 4 that MC realisations of high mass sub-halos are only mildly affected by numerical uncertainties: HADES estimates on these structures are robust, even in the inner regions.

We showed in Fig. 6 that for sub-halos of masses \( 10^7 M_\odot \) the density scatter \( \sigma_{\rho}/\rho \) at the halo centres increases from \( \approx 20\% \) up to \( 60\% \). Even if variations in the density scatter are confined to a very small portion of the LOS, the numerical J-factor computed for these
sources will still feel the consequences, as in general we proved that it holds

\[
\frac{(J) - J_{th}}{J_{th}} = \frac{\int_{\text{LOS}} \sigma_{\rho}^2 dr}{\int_{\text{LOS}} \rho_{th}^2 dr}.
\]  

That is precisely what is happening in the right column of Fig. 11. Here, we see that for sub-halos with cuspy profiles \( \delta J \) values are \( \sim 2.4\% \) higher with a sub-halo than those computed in the smooth case. Again, two effects are at play: on one hand the numerical enhancement of \( \delta J \) coming from the increased density variance; on the other, the reduction of \( \delta J \) due to under-estimates of the density in the sub-halo centre. Computing the expected value of \( \delta J \) would require now a numerical integration of the relative variance point-by-point along the LOS, as we are not able anymore to tell apart the two effects efficiently. However, as \( \delta J \) is overall higher than in the smooth case, we can deduce that for cuspy profiles the augmented variance is by far the predominant effect.

On the contrary, for the Burkert halos in our sample (last panel of Fig. 11) underestimates of the central density dominate over boosting. But then again, in Fig. 6 we showed that cored profiles present a smaller increase in the density variance, while having lower central densities, making all the more difficult proper MC realisations of their flat cores. Thus, HADES estimates in their central regions, while still good, become less stable.

All J-factors presented in Fig. 11 were computed with the radial cut-off \( r_{sh} \geq 0.05 \) kpc still in place. When the cut-off is removed and contributions from \( \rho(r_{sh} \to 0) \) are included in the picture, all the tables turn. \( \delta J \) now becomes negative for every combination of halo + sub-halo, as the steep theoretical profiles predict a signal a good 10\% higher than the numerical J.

Trusting or not that DM in sub-halos would actually form steep cusps if we had enough particles available to resolve them, it looks undeniable that relying solely on analytical profiles to compute annihilation fluxes can have substantial effects.

5.1 J-factor sky maps

The next step in our analysis is to use HADES to generate HEALPix (Gorski et al. 2005) all-sky maps of the J-factor for all halos in the Dragon catalogue. Specifically, we treated each halo as a mock version of the Milky Way halo and produced celestial maps of the J-factor as seen from an observer located on the galactic plane, at a distance of \( 8 \) kpc from the halo centre. Each map has an NSIDE = 128, corresponding to 196608 pixels. We treated separately the case of smooth halos – from which sub-halos were subtracted – and full halos, i.e. containing the full sub-halos population contribution on top of the main halo.

Fig. 12 (upper row) shows the J-factor sky maps for the Dragon halos Dr-1 and Dr-7 for the full halos. These maps are strongly dominated by the main halo component, by far predominant over the sub-structures emission. The lower row of Fig. 12 instead displays the relative difference between each map and the corresponding theoretical map. The latter was again obtained computing the J-factor through eq. (8) assuming one of the analytical profiles of eq. (2). These residual maps provide a visualisation of the DM anisotropies in our galactic-sized halos. The detailed analysis of the distribution of...
DM density fluctuations in real cosmological simulations is beyond the scopes of this work. We believe, however, that it could be of great interest for characterizing the the foreground noise to gamma surveys expected from the DM galactic halo and for studies on extra-galactic background anisotropies. Mapping these fluctuations is a distinctive advantage of HADES that is not available when J-factor maps are produced with traditional techniques. Resorting to smoothed out profiles automatically implies the lost of information on the peculiar mass anisotropies.

To roughly quantify the entity of the fluctuations found in our mock halos, we point out that they follow a distribution compatible to those of Fig. 11: the difference $(J - J_{th})/J_{th}$ is lower than 25% for > 95% of the pixels in each map, and typically of the order of 10%. These results should however not be regarded as universally valid for real halos, and our own Milky Way halo specifically. We will look into this in more detail in the next Section.

6 HADES TO THE TEST: THE ILLUSTRIS-TNG CLUSTERS

In this Section, we will show a practical application of HADES. We have selected one halo from the TNG50-1-Dark simulation in the IllustrisTNG$^2$ suite (Pillepich et al. 2018; Nelson et al. 2019a,b; Pillepich et al. 2019b,a) and run HADES to compute J-factor maps on this clusterised halo. TNG50-1-Dark is a DM only realisation of a 35 $\text{Mpc}/h$ side cube, with a total of about 10 billions DM particles of mass $\sim 3.65 \times 10^8 M_\odot/h$ each. The cosmological parameters have been fixed according to Planck 2015 data (Ade et al. 2016). All halos and sub-halos have been identified with the FoF and SUBFIND (Davis et al. 1985; Springel et al. 2001) halo finders, respectively.

For our study, we selected the halo 154112 (hereafter H154112). This is the third most massive halo of the simulation, with about $1.3 \times 10^8$ DM particles, for a total mass of $4.68 \times 10^{13} M_\odot/h$. H154112 is a cluster-sized halo, with a virial radius$^3$ $\sim$ 830 kpc. As we want to keep the focus on applications to the study of galactic-like halos and subhalos, we re-scaled the chosen TNG halo to make its properties more similar to those of a typical galactic halo. Essentially, we decided to shrink all lengths of a factor of 5 and all masses of a factor of 125, so as to keep all densities unchanged. None of our conclusions are in any way affected by this manipulation. After the re-scaling, the total halo mass measured $\sim 5.5 \times 10^{11} M_\odot$. Fig. 13 shows the maps of

---

2 https://www.tng-project.org/

3 Defined as the radius enclosing an average density 200 times higher than the Universe’s critical density, $\rho_c = 3H_0^2/8\pi G$. 

---

Figure 13. Surface density maps of the TNG halo S-H154112 (smooth component). The left and right panels show projections of the halo along the x and z axis respectively (in comoving units). The gray dots show the position of the observer at 8 kpc from the centre of the halo on the x axis, while the gray circles mark a region of radius 2.5 times the virial radius, centred on the halo centre.

Figure 14. Upper panel: Fit of the smooth component density profile of the S-H154112 halo, as a function of the distance from its centre. The data were fitted with a NFW profile and an Einasto profile. Lower panel: density residuals for the two profiles.
the integrated surface density of the simulation last snapshot zoomed in on the re-scaled halo stripped of its sub-structures. We chose the reference frame so that the centre of the halo has coordinates (0, 0, 0) kpc. Panels left and right of Fig. 13 display the projection of the halo along axes x and z respectively, with the grey circle marking 2.5 times $r_{1/2} = 83.58$ kpc, defined as the radius enclosing half of the total halo mass.

We selected this particular halo because, among the most massive – and hence better resolved – ones, it also was the one whose shape looked closest to spherical symmetry. However, as it can clearly be seen in the figure, H154112 still presents a quite evident prolate shape, and its mass distribution is highly asymmetrical. Though this is more common for cluster-sized halos and asymmetries in galactic halos are typically less prominent, they are nonetheless a feature common to every cosmological simulation. We deemed it particularly interesting to show how HADES behaves in similar conditions. As a first step, we once again studied the smooth halo, as to say, the halo made up of all particles not attributed to a smaller sub-structure by FoF or SUBFIND. Hereafter, we will dub the resulting smoothed halo S-H154112.

We proceeded to fit the smooth halo with an analytical mass profile, assuming spherical symmetry to be a good description of the halo at least to a first approximation. We binned the halo in 30 logarithmic radial bins from $r_{\text{MH}} = 0.1$ kpc to $r_{\text{MH}} = 210$ kpc, leaving out the more irregular outskirts of this halo. We computed the average density in each spherical shell and fitted the resulting data points. Figure 14 shows the resulting best fit curves. We used a NFW and Einasto profiles parametrised ad in eq. (2b), with 2 and 3 free parameters respectively. We found that both curves provide a good fit to the averaged profile, with no strong preference between the two, despite the higher flexibility of the Einasto parametrisation. In all our analytical calculations, we then chose to use for the background halo an Einasto density profile with parameters $\rho_2 = 4.13 \cdot 10^3 M_\odot / \text{kpc}^3$, $r_2 = 25.96$ kpc and $\alpha = 0.165$.

We run HADES on the smooth halo to extract the J-factor maps, and simultaneously computed the theoretical map integrating the analytical background profile from $r_{\text{MH}} = 0$ kpc to $r_{\text{MH}} = 210$ kpc. The position of the observer is chosen to be ($-8.0$, 0.0, 0.0) in halo-centric coordinates. We compared the two maps computing the residual map as $(J - J_{\text{th}}) / J_{\text{th}}$, and obtained the first panel of Fig. 15. We can immediately notice a striking difference between this plot and the bottom row of Fig. 12, both in terms of intensity and morphology of the fluctuations. Instead of presenting as a statistical white background noise, the residuals have a very distinctive shape, consequence of the fact that we are attempting a fit with a spherical profile on a halo that in reality presents a much more articulate morphology. What from Fig. 14 looked like a reasonably good description of the halo, misses out the actual local distribution of DM. HADES on the contrary is not blind to these features and is capable of reproducing in detail the shape of the halo. The resulting differences between the numerical J-factors and the semi-analytical ones can reach up to $> 100\%$, depending on the LOS.

It is important to stress that the specifics of the residual mass in the first panel of Fig. 15 are by no means general. This is better illustrated in the remaining panels of the figure, where we recomputed the same J-factor maps moving the observer around in the main halo.
Specifically, we performed three more runs with HADES, placing the observer respectively at locations (8.0, 0.0, 0.0), (0.0, −8.0, 0.0) and (0.0, 8.0, 0.0) kpc. We conveniently kept the observer always at the same distance of 8 kpc from the halo centre, so that the theoretical map $J_{th}$ remains unchanged. However, the residual map transforms significantly. Observers at different locations see the mass anisotropies from different angles, and the morphology of the resulting J-factor map changes accordingly.

This has deep implications: it shows clearly that describing real halos with spherical or, more generally, smooth profiles to predict J-factor signals carries a non-negligible error. This error depends on the mass distribution of the peculiar halo, but can in principle be very high. Indeed, even if smooth profiles are proven to be great average descriptions of halos, predictions need to account for the fact that when observing objects in the sky, we do not have the chance of mediating over different points of view. The same error affects the results coming from HADES: even if the code is sensitive to the local mass distribution, simulated halos are not exact mirror images of the real ones, so in general the specific morphology of mass anisotropies will differ. Whatever the method used to compute J-factors, there is no way around this, unless one is able to gain information on the specific DM morphology of the target halo for which predictions are made.

However, the code we are developing provides an optimal tool to study the entity and impact of this uncertainty. This kind of analysis requires the systematic study of a sample of simulated halos, each looked at under different angles to quantify, statistically, how often particularly high residuals are recovered. In this regard, using hydrodynamical simulations will allow to reduce degeneracies due to the observer position, as, especially in the case of spiral galaxies, baryons tell us in which plane the disk lies. A similar endeavour goes beyond the scope of this text, but we believe the issue is worthy of investigation and we plan to devote a future work to this exploration.

We close this discussion commenting the plots of Fig. 16 where, in the top panel, we show the map obtained averaging the four maps computed through HADES with observers at different locations. The lower panel shows the residuals of this map. It is worth noticing that, though peculiar features are still evident, this map presents much milder deviations from the semi-analytical one. A validation that, as we average over several points of view, the results converge back to the expectations from the fitted spherical profile.

We now move on to the analysis of sub-structures. We added to the smooth S-H154112 all particles belonging to the sub-halos that were identified by the SUBFIND halo finder, and satisfied two criteria: (i) the total number of particles associated to the sub-structure is higher than $2 \cdot 10^3$ and (ii) the distance between the sub-halo and the main halo centres is lower than 210 kpc. The former condition is based on the results of Section 4 and guarantees that HADES will efficiently resolve the sub-structure, while the latter ensures that all selected sub-halos fall in the J-factor integration volume. We ended up with 547 sub-structures, with masses in the range $9 \cdot 10^8 - 5 \cdot 10^9 M_{\odot}$. We dubbed the resulting total halo H154112 and computed the J-factor map. We then subtracted the background map to highlight the sub-structures contribution. The results are shown in the first panel of Fig. 17, where the background halo presents indeed a prolate shape, slightly tilted towards left. The sub-halos appear to be very bright, and many of them clearly emerge above the background.

In Fig. 17 we also show the sub-halos-only contribution to the J-factor map of H154112 as computed by HADES (left column, bottom panel) and integrating the analytical profiles (right column, top panel). This latter theoretical map was also separately computed with a dedicated tool, CLUMPY (Charbonnier et al. 2012; Bonnivard et al. 2016; Hüttten et al. 2019). The results are in agreement with the map in Fig. 17, and for brevity are not reported here. The last map in the figure displays the residuals of the sub-halo maps $(J_{sh} - J_{sh}^{th}) / J_{sh}^{th}$, where we masked the pixels in which $J_{sh}^{th}$ is zero. We clearly see that HADES recovers efficiently all sub-structures in our sample. However, as already was the case for the main halo, not all sub-halos are accurately described by a spherical profile. Some halos even have very elongated shapes and overall bigger extensions than their theoretical partners, giving rise once again to peculiar features in the residual map. This point is particularly interesting for studies addressing the detection of sub-halos as extended sources in the sky (see e.g. Di Mauro et al. (2020)). Again HADES promises to be a very useful tool to deploy in these analyses, allowing for more precise estimates of sub-halo extension from simulations.

As for the intensity of the signal, both semi-analyticalcalculations and HADES estimations yield J-factors of the same order of magnitude, though in some cases the theoretical profile tends to overestimate the signal with respect to HADES calculations; in other sub-halos the opposite occurs. In the biggest sub-halos, both HADES boosting and mitigation of the signal occur in different regions of the structure. Once again, we impute this variability to the peculiar mass distribution, which deviates from the spherical model.
7 DISCUSSION AND CONCLUSIONS

The determination of dark matter (DM) halos and sub-halos abundance, mass and spatial distribution out of a cosmological simulation is usually obtained by means of halo finders. These algorithms work necessarily on some assumptions for the existence and the extent of sub-structures and hence come with a number of systematic effects. On the other hand, the identification of DM structures in cosmological simulations turns out to be relevant for many astrophysical applications, including the determination of DM indirect detection signatures. The DM density spatial profile of the identified (sub-)halos is usually assumed to follow an analytical distribution, which is determined by fitting said distribution inside the recovered structures. Whenever the statistics of particles inside the structure is not enough to allow a solid fit, the free parameters of the analytical profiles are set through phenomenological relations connecting them to the properties recovered by the halo finders, such as the halo total mass and concentration. However, such a simplified approach, while being very useful to analyse sub-halo populations properties, may not be optimal for applications where a precise estimate of the DM density distribution is required.

In order to overcome their possible intrinsic uncertainties and to assess if halo finders are really indispensable for building predictions on DM indirect detection signals, we built a new numerical tool dubbed HADES (Halo Accurate Density Evaluation System). HADES is based on a nearest-neighbour algorithm aimed at estimating the local density in cosmological simulations with three free hyper-parameters. HADES has first been developed on test-case mock simulations (the Dragon catalogue), that have been built from three different analytical DM density profiles.

The accuracy of the Dragon catalogue is assessed separately on the galactic-sized smooth halo and on the full system containing sub-halos with masses in the $10^6 - 10^9 M_{\odot}$ range. Taking advantage of our knowledge of the input parameters upon which the Dragon catalogue was built – and consequently of its halos and sub-halos properties – we tested HADES capability of correctly reconstructing the local density and the sub-structures. We found that HADES is extremely successful in recovering the local density of the main smooth halo with an accuracy of few %. As for the sub-halos, sub-structures containing more than $10^3$ particles – i.e. with masses higher than $10^7 M_{\odot}$ at the resolution of our mock halos – are reproduced just as well down to their typical convergence radius. Only on the smaller sub-halos, with merely ~ 100 particles, HADES encounters difficulties in reconstructing the density with the same level of accuracy and provides solely a conservative estimate of the density profile, especially towards the structures’ centres.

Among the possible astrophysical applications of HADES, a quantity relevant for DM indirect detection is the so called J-factor, a measure of the DM squared density along a line of sight (LOS). We computed the J-factor on the Dragon catalogue halos with HADES and compared it with the results obtained assuming analytical density profiles. With respect to the analytical predictions, HADES automatically includes in the computation the boosting of the signal due to density fluctuations. We found in Section 5 that for our mock halos...
this boosting is only a moderate effect, reaching on average less than 4% and getting as high as 25% only for a handful of LOS.

A proof-of-concept application of HADES was conducted on a simulated halo from the TNG-1-Dark run in the TNG50 simulation of the IllustrisTNG suite. We ran HADES on both the smooth component and the sub-structures of this halo extracting the J-factor maps. Simultaneously, we computed the corresponding theoretical maps integrating the analytical profiles fitted over the simulated halo and sub-halos. Both semi-analytical calculations and HADES estimations yield J-factors of the same order of magnitude. However, this halo has a much more complex 3D mass distribution than our spherical mock halos. Local anisotropies and density scatter have much higher impact and this reflects on the numerical J-factor, which deviates from its theoretical counterpart of up to 200 – 300%, depending on the LOS. The morphology of such deviations is of course dependent on the specific halo matter distribution and on the observer location in the halo.

This highlights a non-negligible level of uncertainty in using analytical profiles as good descriptors of simulated halos to compute phenomenological predictions. HADES computation does not remove such error, but a systematic study of a high number of halos with HADES would crucially allow for the first time the full characterization of its impact.

Moreover, a fine reconstruction of the DM distribution in big halos has a number of other astrophysical applications, from studies on the systematics of detecting sub-structures through weak lensing to a deeper understanding of the connection between DM anisotropies and baryon dynamics in galactic systems. We also found that HADES efficiently recovers the sub-structures in the TNG halos and is capable of telling deviations from sphericity and determine the sub-structure extension. This will be extremely relevant for statistical studies on sub-halos detectability as extended sources. A thorough analysis of HADES potential applications, related to DM indirect detection signatures and beyond is left for future works.

ACKNOWLEDGEMENTS

We thank Dylan Nelson for giving us access to the IllustrisTNG simulations data and Moritz Hütten for helping us with the installation of CLUMPY. AB is supported by a de Sitter Fellowship of the Netherlands Organization for Scientific Research (NWO). VDR acknowledges financial support by the SEI/2020/016 grant (project “Les Fosques”) funded by Generalitat Valenciana, by the Universitat de València through the sub-programme “ATRACCIÓN DE TALENTE 2019” and by the Spanish grant PID2020-113775GB-I00 (AEI / 10.13039/501100011033). The work of FD is supported by the DEPARTMENTS OF EXCELLENCE grant awarded by the Italian Ministry of Education, University and Research (MIUR) and by the Research grant TA$P$ (Theoretical Astroparticle Physics) funded by Istituto Nazionale di Fisica Nucleare.

REFERENCES

Abdalla H., et al., 2021, Astrophys. J., 918, 17
Ackermann M., et al., 2012, Phys. Rev. D, 85, 083007
Ade P. A. R., et al., 2016, Astron. Astrophys., 594, A13
Ando S., 2009, Phys. Rev. D, 80, 023520
Ando S., Ishiyama T., Hiroshima N., 2019, Galaxies, 7, 68
Bartels R., Ando S., 2015, Phys. Rev. D, 92, 123508
Batista R. A., et al., 2021, White paper of the European Consortium for Astroparticle Theory (EuCAP)

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, Astrophys. J., 762, 109
Bergstrom L., Edsjo J., Ullio P., 1998, Phys. Rev. D, 58, 083507
Bertoni B., Hooper D., Linden T., 2015, JCAP, 12, 035
Bonnivard V., et al., 2015, Mon. Not. Roy. Astron. Soc., 453, 849
Bonnivard V., Hütten M., Nezri E., Charbonnier A., Combet C., Maurin D., 2016, Comput. Phys. Commun., 200, 336
Burkert A., 1995, Astrophys. J. Lett., 447, L25
Calore F., De Romeri V., Di Mauro M., Donato F., Herpich J., Macciò A. V., Maccione L., 2014, Mon. Not. Roy. Astron. Soc., 442, 1151
Calore F., De Romeri V., Di Mauro M., Donato F., Marinacci F., 2017, Phys. Rev. D, 96, 063009
Calore F., Hütten M., Stref M., 2019, Galaxies, 7, 90
Chang J. L., Necib L., 2021, Mon. Not. Roy. Astron. Soc., 507, 4715–4733
Charbonnier A., Combet C., Maurin D., 2012, Comput. Phys. Commun., 183, 656
Chua K. T. E., Pillepich A., Vogelsberger M., Hernquist L., 2019, MNRAS, 484, 476
Coronado-Blázquez J., Sánchez-Conde M. A., 2019, Galaxies, 8, 5
Coronado-Blazquez J., Sanchez-Conde M. A., Domínguez A., Aguirre-Santaella A., Di Mauro M., Mirabal N., Nieto D., Charles E., 2019a, JCAP, 07, 020
Coronado-Blázquez J., Sánchez-Conde M. A., Di Mauro M., Aguirre-Santaella A., Ciucă I., Domínguez A., Kawata D., Mirabal N., 2019b, JCAP, 11, 045
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, Astrophys. J., 292, 371
Dehnen W., McLaughlin D., 2005, Mon. Not. Roy. Astron. Soc., 363, 1057
Despali G., Giocoli C., Tormen G., 2014, Mon. Not. Roy. Astron. Soc., 443, 3208
Di Mauro M., Stref M., Calore F., 2020, Phys. Rev. D, 102, 103010
Diemand J., Kuhlen M., Madau P., 2006, Astrophys. J., 649, 1
Diemand J., Kuhlen M., Madau P., 2007a, Astrophys. J., 657, 262
Diemand J., Kuhlen M., Madau P., 2007b, Astrophys. J., 667, 859
Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, Nature, 454, 735
Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019, Publ. Astron. Soc. Austral., 36, e021
Fornasa M., et al., 2013, Mon. Not. Roy. Astron. Soc., 429, 1529
Gao L., Navarro J. F., Cole S., Frenk C., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, Mon. Not. Roy. Astron. Soc., 387, 536
Gill S. P. D., Knebe A., Gibson B. K., 2004, Mon. Not. Roy. Astron. Soc., 351, 399
Gorski K. M., Hivon E., Banday A. J., Hansen F. K., Reinecke M., Bartelman M., 2005, Astrophys. J., 622, 759
Graham A. W., Merritt D., Moore B., Diemand J., Trenzi C., 2006, Astron. J., 132, 2685
Grand R. J. J., White S. D. M., 2021, Mon. Not. Roy. Astron. Soc., 501, 3558
Green S. B., van den Bosch F. C., 2019, Mon. Not. Roy. Astron. Soc., 490, 2091
Han J., Jing Y. P., Wang H., Wang W., 2012, Mon. Not. Roy. Astron. Soc., 427, 2437
Han J., Cole S., Frenk C. S., Benitez-Llambay A., Helly J., 2018, Mon. Not. Roy. Astron. Soc., 474, 604
Hernquist L., 1990, Astrophys. J., 356, 359
Hiroshima N., Ando S., Ishiyama T., 2018, Phys. Rev. D, 97, 123002
Hooper D., Witte S. J., 2017, JCAP, 04, 018
Hütten M., Maier G., 2018, JCAP, 08, 032
Hütten M., Combet C., Maurin D., 2019, Comput. Phys. Commun., 235, 336
Kamionkowski M., Koushiappas S. M., Kuhlen M., 2010, Phys. Rev. D, 81, 043532
Klypin A. A., Gottloeber S., Kravtsov A. V., 1999, Astrophys. J., 516, 530
Knebe A., et al., 2011, Mon. Not. Roy. Astron. Soc., 415, 2293
Knebe A., et al., 2013, Mon. Not. Roy. Astron. Soc., 435, 1618
Knollmann S. R., Knebe A., 2009, Astrophys. J. Suppl., 182, 608
Kuhlen M., Vogelsberger M., Angulo R., 2012, Phys. Dark Univ., 1, 50
Moliné A., Sánchez-Conde M. A., Palomares-Ruiz S., Prada F., 2017, Mon. Not. Roy. Astron. Soc., 466, 4974
Moore B., Quinn T. R., Governato F., Stadel J., Lake G., 1999, Mon. Not. Roy. Astron. Soc., 310, 1147
Mulderw S. I., Pearce F. R., Power C., 2011, Mon. Not. Roy. Astron. Soc., 410, 2617
Navarro J. F., Frenk C. S., White S. D. M., 1996, Astrophys. J., 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, Astrophys. J., 490, 493
Navarro J. F., et al., 2004, Mon. Not. Roy. Astron. Soc., 349, 1039
Navarro J. F., et al., 2010, Mon. Not. Roy. Astron. Soc., 402, 21
Nelson D., et al., 2019a, Comput. Astrophys., 6
Nelson D., et al., 2019b, Mon. Not. Roy. Astron. Soc., 490, 3234
Neyrinck M. C., Gnedin N. Y., Hamilton A. J. S., 2005, Mon. Not. Roy. Astron. Soc., 356, 1222
Onions J., et al., 2012, Mon. Not. Roy. Astron. Soc., 423, 1200
Pato M., Iocco F., 2015, Astrophys. J. Lett., 803, L3
Pieri L., Bertone G., Branchini E., 2008, Mon. Not. Roy. Astron. Soc., 384, 1627
Pillepich A., et al., 2018, Mon. Not. Roy. Astron. Soc., 473, 4077
Pillepich A., et al., 2019a, Mon. Not. Roy. Astron. Soc., 490, 3196
Pillepich A., et al., 2019b, Mon. Not. Roy. Astron. Soc., 490, 3196
Prada J., Forero-Romero J. E., Grand R. J. J., Pakmor R., Springel V., 2019, MNRAS, 490, 4877
Schaller M., et al., 2015, Mon. Not. Roy. Astron. Soc., 452, 343
Schoonenberg D., Gaskins J., Bertone G., Diemand J., 2016, JCAP, 05, 028
Siegal-Gaskins J. M., 2008, JCAP, 10, 040
Silk J., Stebbins A., 1993, Astrophys. J., 411, 439
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Mon. Not. Roy. Astron. Soc., 328, 726
Springel V., et al., 2008, Mon. Not. Roy. Astron. Soc., 391, 1685
Streel M., Lalavle J., 2017, Phys. Rev. D, 95, 063003
Tormen G., Moscardini L., 2004, Mon. Not. Roy. Astron. Soc., 350, 1397
Zavala J., Frenk C. S., 2019, Galaxies, 7, 81
Zhu Q., Marinacci F., Maji M., Li Y., Springel V., Hernquist L., 2016, Mon. Not. Roy. Astron. Soc., 458, 1559
van den Bosch F. C., Ogiya G., 2018, Mon. Not. Roy. Astron. Soc., 475, 4066
van den Bosch F. C., Ogiya G., Hahn O., Burkert A., 2018, Mon. Not. Roy. Astron. Soc., 474, 3043

APPENDIX A: HADES BEST PARAMETERS CHOICE

We recall from eq. (1) that HADES is fully characterised by 3 independent parameters, $N$ and the couple $(a, b)$. Here $N$ can be considered a measure of the locality of the density estimate. The higher the value of $N$, the more particles are involved in the computation, the less local the estimate will be. We found no significant difference in estimates performed on the bigger host halos when varying $N$ in the interval $[10, 20]$. However, we advise against using higher $N$ values, as while only minimally affecting estimates on the more resolved host halo, the loss of locality can have a huge impact in sub-halos density reconstruction. Throughout this work we have made the choice $N = 15$ and kept this parameter fixed.

The couple $(a, b)$, on the other hand, carries an indication of how many particles in a given random distribution can be considered as nearest neighbours, $a$ being the minimum estimate and $b$ the maximum. The issue of setting the optimal values for $(a, b)$ does not present a straightforward solution, and at least to begin with needs to be addressed phenomenologically. However, we can make some remarks on the likely interval for these parameters. We consider that, taken a particle in the simulation, there are 6 main spatial directions along which nearest neighbours can be looked for: up, down, forward, backward, right and left. Along each of these directions, it is reasonable to expect either 1 or 2 particles to be actual nearest neighbours. This empirical reasoning brought us to speculate values for $a$ and $b$ respectively of 6 and 12. We then considered the configurations $(a, b) = (5, 13), (6, 13) (6, 12), (7, 12), (7, 11), (8, 11), (8, 10), (9, 10)$, all having in common an average number of nearest neighbours of $\approx 9$. We recall that the principal constraint we imposed when generating the mock simulations of Section 3 is that the resulting halos should present a specific radial density distribution $\rho(r)$. Thus, the best parameters couple $(a, b)$ can be identified based on which choice is actually able to better reconstruct $\rho(r)$ for any given halo.

We ran HADES on each of the 9 smooth halos$^4$ in the Dragon catalogue, selecting in turn each of the 8 $(a, b)$ configurations listed above. More specifically, for each halo we evaluated the density with all chosen HADES configurations at 1000 points equally distant from the halo centre, and repeated the procedure at 10 progressive radii, for a total of 10,000 estimation spots.

Ideally, once selected a radius $r_{\text{MH}}$, each of the 1000 estimates $\rho_i$ made at fixed radius should be equal to $\rho(r_{\text{MH}})$. However, we are dealing with an MC simulation and we do expect to have local fluctuations of the density due to the randomness of the particles positions sampling process. This means that the $10^3$ $\rho_i$ will follow a Poisson distribution that – if the choice of $(a, b)$ is appropriate – will have $\rho(r_{\text{MH}})$ as the expectation value. An example of this can be observed in Fig. A1, where we show the histograms of the $\rho_i$ at fixed $r_{\text{MH}} = 90$ kpc of the Dr-1 halo, for two different choices of $a$ and $b$. In the plots, the red solid and cyan dashed lines mark respectively $\rho_{\text{th}} = \rho^E(r_{\text{MH}})$ and the mean value of the distribution $\bar{\rho}$.

To evaluate the different parameter choices for HADES we focus on the three quantities:

- $\sigma_{\bar{\rho}}/\bar{\rho}$: this is the variance of the distribution weighted over the mean value, and it is a measure of the width of the distributions in Fig. A1. If the numerical error introduced by HADES in its estimate was to be completely negligible, this ratio would quantify the entity of the density field local fluctuations. We find that for every choice of $a$ and $b$ and for each halo in our catalogue, $\sigma_{\bar{\rho}}/\bar{\rho} \approx 20\%$.

- $\sigma_{\bar{\rho}}$: this is the variance on the mean of the distribution, again weighted on the mean value. The difference with the previous ratio is that this second variance is computed as $\sigma_{\bar{\rho}} = \sqrt{\frac{1}{N} \sum_i (\rho_i - \bar{\rho})^2}$, $N$ being the total number of sample points for each histogram $\times 10^3$ in our case. We use this quantity as an approximate estimate of the error over the expectation value of the density distribution, as to say, an estimate of how $\bar{\rho}$ would differ if we were to repeat our test using a different sample of $10^3$ points, always at the same radius.

- $\delta_{\rho}/\bar{\rho}$: this ratio is defined as $(\bar{\rho} - \rho_{\text{th}})/\bar{\rho}$. To conclude that the choice we made for $(a, b)$ is acceptable, we need to require for this quantity to either be smaller or match the error on the mean value

$^4$ i.e. the MC host halos before the addition of the sub-structures.

Figure A1. Density distribution at fixed $r_{\text{MH}} = 90$ kpc for the halo Dr-1 and two different choices of $a$ and $b$. The solid red line marks the theoretical value $\rho_{\text{th}} = \rho^E(r_{\text{MH}})$, while the dashed cyan line indicates the expectation values of the distributions.
The best choice for HADES parameters will then be the one that minimises $\delta \rho / \bar{\rho}$.

In Fig. A2, we plot the ratio $\delta \rho / \bar{\rho}$ at all radii $r_{\text{MH}}$ we considered in our test, for three of the Dragon halos. The light blue area marks the $3\sigma$ relative error over the expectation value of the density distribution. The fluctuations of all these lines along the horizontal 0 value suggest that the true density distribution of the halos itself slightly differs from $\rho_0(r)$. We recall from Section 3 that the MC halos reproduce the theoretical $\rho_0(r)$ only within a given error, hence small deviations are expected. This scenario is supported by the accordant trend of the curves in Fig. A2, each one seeming to over(under)-estimate the same underlying density of a fixed amount.

From Fig. A2 we can clearly see that the couples $(5, 13)$, $(6, 12)$, $(7, 11)$ and $(8, 10)$ can be set aside, as they introduce a systematic overestimate of the density that is way higher than the $3\sigma$ uncertainty. On the contrary, the remaining 4 couples $(6, 13), (7, 12), (8, 11)$ and $(9, 10)$ seems to all be acceptable candidates, each of them slightly underestimating the average density of less than $\sim 2\%$. However, the couple $(6, 13)$ evidently yields the better result, always deviating from 0 of at most 1%. This same result holds for all the halos in the Dragon catalogue, hinting that it is generally valid for any kind of density profiles $\rho(r)$ typically found in DM halos. This is reassuring, as we wish to choose a parameter configuration to run HADES on simulations for which we have no control over the real underlying DM density, ideally introducing the smallest possible error.

To further check that the choice $(6, 13)$ is indeed the best independently of the density profile, we repeated the test performed in this Section on a number of MC simulations following very different mass distributions, from a uniform realisation to power-laws of different steepness. In each case, the couple $(6, 13)$ always resulted the best possible parameters choice, going as far as to reconstruct the average density with a precision of 0.04% in the case of a uniform realisation. This makes us reasonably confident that HADES can be applied safely to real simulations as well, with the suggested choice $(a, b) = (6, 13)$ possibly introducing only systematic reaching in the worst case scenario the $\%$ level.

**APPENDIX B: INTRINSIC HADES ACCURACY**

In this Appendix, we discuss HADES numerical error on density estimates. The main contributions to such uncertainty are due to the peculiar local particles positions that can influence HADES conclusions. These need to be distinguished from real local under(over)-densities in the particles distribution, that typically manifests on bigger scales and are not swiped out if the estimate is performed using a different set of particles. Hence, characterising the accuracy of our tool simply means identifying the lower magnitude of fluctuations to which HADES is sensitive consistently.

In practice, we are addressing the following issue: say we are using HADES to estimate the density at a specific point $\vec{x}_i$ in the halo, choosing as input parameters $(a, b)$ the couples $(6, 13)$ and $(9, 10)$. We dub the results respectively $\rho_{6,13}^{\text{Had}} = \rho(\vec{x}_i)_{|a=6,b=13}$ and $\rho_{9,10}^{\text{Had}} = \rho(\vec{x}_i)_{|a=9,b=10}$. The two estimates will not be the same: as can be deduced from the plots in Fig. A2, HADES running with the couple $(9, 10)$ uses a different number of nearest neighbours to perform the estimate and tends to systematically underestimate $\rho(\vec{x}_i)$ with respect to runs with parameters $(6, 13)$. The choice of parameters $a$ and $b$ also determines the expected density at location $\vec{x}_i$, namely $\rho_{a,b}^{\text{Had}}$. This density is connected to the analytical density $\rho_{\text{MH}}(\vec{x}_i)$ that can be computed using one of the profiles of eq. (2). $\rho_{\text{MH}}$ and $\rho_{a,b}^{\text{Had}}$ only differ because of Fig. A2: $\rho_{\text{MH}}$ is of course insensitive to the parameters $(a, b)$ while $\rho_{a,b}^{\text{Had}}$ is a consistent under-estimate of $\rho_{\text{MH}}$ that depends on the chosen parameters. As we are assuming spherical symmetry for our halos, $\rho_{\text{MH}} - \rho_{\text{Had}}$ and consequently $\rho_{a,b}^{\text{Had}}$ only depends on the distance of the point $\vec{x}_i$ from the halo centre, $r = |\vec{x}_i|$.

From the analysis of the curves in Fig. A2 we can also infer that for the parameters choices of our example, $\rho_{5,13}^{\text{Had}} \approx 1.02 \rho_{9,10}^{\text{Had}}$. Suppose now that we find $\rho_{5,13}^{\text{Had}} < \rho_{9,10}^{\text{Had}}$, i.e. the real density at $\vec{x}_i$ is lower than expected and hence we are looking at what seems to be a local under-density. We now wonder: is $\rho_{5,13}^{\text{Had}}$ also smaller than $\rho_{9,10}^{\text{Had}}$? Are the relative differences $(\rho_{5,13}^{\text{Had}} - \rho_{9,10}^{\text{Had}}) / \rho_{9,10}^{\text{Had}}$ and $(\rho_{5,13}^{\text{Had}} - \rho_{9,10}^{\text{Had}}) / \rho_{9,10}^{\text{Had}}$ comparable? In other words, does using a different set of particles to estimate the local density confirm the presence of an under-density of the same magnitude at position $\vec{x}_i$? Or is such under-density only a fake effect that will disappear changing input parameters?

Intuitively, say that HADES with input parameters $(6, 13)$ recovers
Figure B1. Example histograms of $\Delta \rho_i$ for two choices of HADES parameters. The black and blue dashed lines indicate a relative dispersion of respectively 1\% and 2\%.

at the position $\bar{x}_i$ an under-density of $\sim 18\%$, while the estimate with inputs $(9,10)$ finds a local under-density of 19\%. Obvious conclusion would be that the under-density recovered by HADES is reliable with a possible 1\% uncertainty.

Formally, we can write for the algorithm error

$$\Delta \rho_i = \frac{\rho_i^{6,13} - \rho_i^{9,10}}{\bar{\rho}_r^{6,13}} = \frac{\rho_i^{6,13} - \rho_r^{6,13} - \rho_r^{9,10} - \rho_i^{9,10}}{\bar{\rho}_r^{6,13}} \cdot 1.02 \bar{\rho}_r^{9,10},$$

which approximately leads to

$$\Delta \rho_i \approx \frac{\rho_i^{6,13} - \rho_r^{6,13}}{\bar{\rho}_r^{6,13}} = \frac{\rho_i^{9,10} - \rho_r^{9,10}}{\bar{\rho}_r^{9,10}}.$$  \hspace{1cm} \text{(B2)}$$

The same holds with only minor modifications if instead of the couple $(9,10)$ we had used any of the combinations explored in Appendix A. Hence, generically we can write

$$\Delta \rho_i \approx \frac{\rho_i^{a,b} - \rho_r^{a,b}}{\bar{\rho}_r^{a,b}} = \frac{\rho_i^{a,b} - \rho_r^{a,b}}{\bar{\rho}_r^{a,b}}. \hspace{1cm} \text{(B3)}$$

Thus, we randomly sampled 10 sets of $10^3$ points in each of our halos. All points in a set are located at the same radius, and hence have associated the same expected density $\bar{\rho}_r^{a,b}$ that can simply be computed as the average of the individual local densities at the sample points $\rho_i^{a,b}$. The radii of the 10 samples span a wide interval from 2 to 120 kpc. For all the sample points we then computed the value of $\Delta \rho_i$ for every possible choice of parameters $(a,b)$. Fig. B1 shows as examples the two resulting histograms of $\Delta \rho_i$ for the halo Dr-1 and the combinations $(6,13)$ - $(9,10)$ and $(6,13)$ - $(6,12)$. Similar plots were made for any combination $(6,13)$ - $(a,b)$, for all halos in our catalogue, yielding analogous results. We observe that for every couple $(a,b)$, $\Delta \rho_i$ falls in the interval $[-0.01; 0.01]$ for more than 95\% of the sampled points, while more than 99.9\% of points in the plots have $-0.02 < \Delta \rho_i < 0.02$. From that, and in combination with the considerations made in Appendix A, we conclude that indeed, in optimal conditions, HADES runs can be considered accurate at \% level.

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