Catalogues of voids as antihalos in the local Universe

Harry Desmond⋆1, Maxwell L. Hutt1, Julien Devriendt1 and Adrianne Slyz1
1Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

26 January 2022

ABSTRACT
A recently-proposed algorithm identifies voids in simulations as the regions associated with halos when the initial overdensity field is negated. We apply this method to the real Universe by running a suite of constrained simulations of the 2M++ volume with initial conditions inferred by the BORG algorithm, along with the corresponding inverted set. Our 101 inverted and uninverted simulations, spanning the BORG posterior, each identify ≈150,000 “voids as antihalos” with mass exceeding $4.38 \times 10^{11} M_{\odot}$ (100 particles) at $z = 0$ in a full-sky sphere of radius 155 Mpc/$h$ around the Milky Way. We calculate the size function, volume filling fraction, ellipticity, central density, specific angular momentum, clustering and stacked density profile of the voids, and cross-correlate them with those produced by VIDE on the same simulations. We make our antihalo and VIDE catalogues publicly available.

Key words: large-scale structure of the Universe – dark matter – galaxies: halos – software: simulations – catalogues

1 INTRODUCTION
Cosmic voids are the lowest density regions of the Universe, occupying most of its volume. While most effort to constrain structure formation and cosmology has focused on the regions of greatest density—galaxies and halos—voids are emerging as powerful probes of the constituents and physics of the Universe (see Pisani et al. 2019 and references therein). A large part of their constraining power stems from the fact that they evolve linearly for far longer than filaments, walls and halos, making them more amenable to analytic modelling. However, traditional watershed voidfinders do not identify purely linearly evolving regions, making their products difficult to connect to theoretical expectations (Nadathur & Hotchkiss 2015). In addition, voids are normally estimated from the observed distribution of galaxies, which are sparse and biased tracers of the underlying dark matter.

Recently, Pontzen et al. (2016); Stopyra et al. (2021) and Shim et al. (2021) have developed a new voidfinding algorithm for $N$-body simulations in which voids in one simulation are comprised of particles belonging to halos in a corresponding inverted simulation where the initial overdensity, displacement and velocity fields are negated. This more physically-motivated, dynamical definition yields voids with sensible 1-point statistics and density profiles, and which furthermore are well described by the Zel’dovich approximation down to $M = 10^{13} M_{\odot}/h$. However, until now this voidfinder has not been applied to real data due to the requirement that the simulations on which it operates be constrained to match the local Universe. Here we produce a suite of catalogues of “voids as antihalos” out to $\sim 155$ Mpc/$h$ from the Milky Way by running $N$-body simulations with the phases of the initial density modes tuned to match the observed large-scale structure in the 2M++ catalogue (Lavaux & Hudson 2011). In addition to securing the benefits of the antihalo definition, this suite allows full marginalisation over the uncertainties in the initial conditions for Bayesian post-processing applications. We make it publicly available.1

Sec. 2 describes our simulations and methods for making the void catalogues. Sec. 3 displays the properties of the voids, including 1- and 2-point functions and stacked density profiles. Sec. 4 summarises, discusses remaining open questions for the antihalo voidfinder, and proposes applications.

2 METHOD

2.1 BORG – CSiBORG – CSiBORG−1
Our method begins with the BORG algorithm, a probabilistic Bayesian inference machine for forward modelling the galaxy number density field and comparing to observations to constrain the phases of the dark matter density modes along with nuisance parameters including galaxy bias (Jasche & Wandelt 2012, 2013). We use the 2M++ reconstruction (Lavaux & Jasche 2016), which employs a box of side length 677.77 Mpc/$h$ (centred on the Milky Way) split into $256^3$ voxels, and follows the evolution of the density field from $z = 69$ to the present using a particle-mesh gravity solver (Jasche & Lavaux 2019). The parameters are inferred with a Hamiltonian Monte Carlo sampler, producing a chain of $\sim 100$ autocorrelation lengths in the voxel densities. Only in the region

1 https://zenodo.org/record/5503610 (Desmond et al. 2021)
in which the 2M++ survey has high completeness (within \(\sim\) 155 Mpc/h), however, is the density field well constrained.

We use the initial conditions (ICs) in the converged part of the BORG chain to run a suite of high-resolution N-body simulations, dubbed CSiBORG (“Constrained Simulations in BORG”), using the RAMSES adaptive mesh refinement code (Teyssier 2002). These simulations perform a zoom-in on a sphere of radius 155 Mpc/h around the Milky Way within which the BORG ICs are augmented with white noise to a resolution of 2048\(^3\), giving a minimum particle mass \(m_p = 4.38 \times 10^9 M_\odot\). Refinement is performed only in this central sphere, although the full BORG box is retained for accurate modelling of longer wavelength modes. To prevent numerical artifacts at the edge of the zoom-in region, a spherical buffer region of width \(\sim 10\) Mpc/h is placed around the sphere in which the resolution degrades gradually to the background BORG value. CSiBORG contains 101 simulations in all with ICs spanning the full range of the BORG posterior, allowing the uncertainties in the BORG inference to be marginalised over in post-processing. Halos in the high-resolution region are found on the fly using the mergerTree patch to RAMSES, which implements the watershed halofinder PHaw (Bleuler et al. 2015). CSiBORG is a general-purpose constrained simulation suite and has previously been used in Bartlett et al. (2021).

BORG and CSiBORG use equatorial coordinates and the cosmology \(T_{CMB} = 2.728\) K, \(\Omega_m = 0.307\), \(\Omega_c = 0.693\), \(\Omega_k = 0.04825\), \(H_0 = 70.5\) km s\(^{-1}\) Mpc\(^{-1}\), \(\sigma_8 = 0.8288\) and \(n = 0.9611\).

We have now run a corresponding set of simulations—CSiBORG\(^{-1}\)—in which the overdensities, particle displacements and velocities are the negative of their CSiBORG counterparts. We modify RAMSES so that particles with the same undispaced positions in CSiBORG and CSiBORG\(^{-1}\) are assigned the same IDs. This provides 101 realisations of the local Universe in which particles are assigned both halo and antihalo flags. Here we work solely with the \(z = 0\) snapshots. Voids have previously been identified by applying VIDE to the SDSS BORG reconstruction in Leclercq et al. (2015).

2.2 Calculating void properties

To identify the antihalos in CSiBORG we perform a Voronoi decomposition using Voro++ (Rycroft 2009). We first cut a cube out of each simulation just wide enough to enclose the full buffer region and hence containing all particles of mass less than the maximum. Although we are ultimately interested only in the \(4.41 \times 10^8\) lowest-mass particles, including their neighbours is necessary for the Voronoi cells to be computed correctly at the edge of the zoom-in region.

A void in CSiBORG comprises the set of particles indexed to a particular halo in CSiBORG\(^{-1}\). Although PHaw is able to identify substructure, here we consider only main halos in CSiBORG\(^{-1}\), mapping each particle to the highest-level (most massive) halo to which it belongs. We exclude voids with fewer than 100 particles which may be imperfectly resolved and unlikely correspond to true underdensities. For each of the remaining voids we calculate the following properties:

- \(\bar{x}_{\text{VWB}}\): The position of the volume-weighted barycentre (VWB) \(\frac{\sum_i x_i V_i}{\sum_i V_i}\), where \(V_i\) is the Voronoi volume of particle \(i\) belonging to the void.
- \(M\): The summed mass of the particles in the void. By construction this equals the halo mass in CSiBORG\(^{-1}\).

- \(V\): The summed Voronoi volumes of the void.
- \(\rho_0\): The central density, defined as the average volume-weighted density within the sphere around the VWB that encloses 64 particles: \(64 m_p / \sum_{k=1}^{64} V_k\).
- \(X, Y\): The two 3D ellipsoidal axis ratios. These are ratios of the square roots of the eigenvalues of the inertia tensor \(I_{ij} = \sum_k (x_{ij,k} - \bar{x}_{\text{VWB},j,k})(x_{ij,k} - \bar{x}_{\text{VWB},j,k})/N_{\text{part}}\), where \(N_{\text{part}}\) is the number of particles in the void.

From these we can calculate the effective radius, \(R_{\text{eff}} = (\frac{2V}{4\pi})^{1/3}\), and the lengths of the three principal axes defined so that their product is \(R_{\text{eff}}^3\): \(R_x = R_{\text{eff}}(XY)^{-1/3}\), \(R_y = R_{\text{eff}}(X^2/Y)^{1/3}\), \(R_z = R_{\text{eff}}(Y^2/X)^{1/3}\). We find the ellipsoidal approximation to be good for the majority of antihalos, although \(\sim\)15% have disconnected regions and/or strings of particles emerging from their centres.

We also calculate the density profiles of the voids in stacks. We compare four methods for this, two of which average in spherical annuli (scaled by \(R_{\text{eff}}\)) around the VWB, and the other two in ellipsoidal annuli aligned with the principal axes. In practice, the latter calculates

\[
d_{\text{ell},k} = \left(\frac{x_k - x_{\text{VWB}}}{R_x}\right)^2 + \left(\frac{y_k - y_{\text{VWB}}}{R_y}\right)^2 + \left(\frac{z_k - z_{\text{VWB}}}{R_z}\right)^2\right]^{1/2}
\]

for each particle (not only those belonging to the void) with \(d_{\text{ell},k} < 5\), where \(x', y', z'\) is a rotated coordinate system aligned with the principal axes. For both types of averaging we consider two estimators for the stacked density profile.

The first is the volume-weighted average density in each annulus with a small correction to remove bias in the estimate of the Poisson mean (Nadathur et al. 2015):

\[
\bar{\rho}_j = m_p \left(\frac{\sum_i N_{ij}}{\sum_i V_{\text{shell},ij}} + 1\right),
\]

where \(N_{ij}\) is the number of particles in shell \(j\) of void \(i\) and \(V_{\text{shell},ij}\) is the volume of the \(j\)th spherical or ellipsoidal shell of void \(i\). The second uses the Voronoi volumes of the particles instead of the volumes of the annuli:

\[
\bar{\rho}_j = m_p \frac{\sum_i N_{ij}}{\sum_i \sum_{k=1}^{N_{ij}} V_k}
\]

Eq. 2 is more robust to boundary effects than Eq. 1 and is therefore to be preferred (Nadathur et al. 2015). We only include in the stack voids with \(M > 10^{13} M_\odot/h\).

2.3 Comparison with VIDE and cross-correlations

For a first comparison of our antihalo catalogues with catalogues of voids identified by other means, we run VIDE (Sutter et al. 2015) on CSiBORG. This uses ZOBOT (Neyrinck 2008) to identify voids via a hierarchical watershed procedure, beginning with a particle-by-particle density estimate through a Voronoi tessellation. We randomly sub-sample the CSiBORG particles to a density of 0.2 (Mpc/h\(^{-3}\)) before performing the tessellation, and require \(R_{\text{eff}} > 2\) Mpc/h to remove voids smaller than the mean particle separation. We use the default density cut \(\bar{\rho}(R < R_{\text{eff}}/4) / \rho_m < 0.2\) and merging threshold 0.2 \(\bar{\rho}\). This limits the hierarchy of ZOBOT such that there are many top-level voids, and only these “parent” voids are included in our analysis. We also run VIDE on one realisation from CSiBORG with a higher (2 (Mpc/h\(^{-3}\))) and lower (4 \(\times 10^{-3}\) (Mpc/h\(^{-3}\)) sub-sampling density.
We cross-correlate antihalos, VIDE voids and halos in CSiBORG and CSiBORG$^{-1}$ in the zoom-in region between 5 and 35 Mpc$/h$ using Corrfunc (Sinha & Garrison 2020).

3 RESULTS

3.1 1-point functions of void statistics

Fig. 1 shows the median differential void size function (VSF) across all 101 realisations and the 68% confidence region between them, along with the corresponding result from VIDE. To enhance compatibility with VIDE, in all plots we restrict to antihalos with $\rho_0 \leq 0.2\,\rho_m$. We only include objects with centres within 150 Mpc$/h$ as those further out may be impacted by the buffer region. We see that our method produces both smaller and larger voids than VIDE. The VSF of VIDE is a strong function of the tracer density, with larger $\bar{n}$ producing more small voids and fewer large ones; with our default sub-sampling density of 0.2 (Mpc$/h$)$^{-3}$ there are $\geq2\times$ fewer VIDE voids than antihalos, although there are $\geq2\times$ more when $\bar{n} = 2$ (Mpc$/h$)$^{-3}$. We find a similar VIDE and antihalo VSF at the large radius end when $\bar{n} = 4 \times 10^{-3}$ (Mpc$/h$)$^{-3}$, although there are only $\sim250$ VIDE voids in total in that case. These voids are significantly smaller than those produced by conventional voidfinders applied to the galaxy field due to the higher tracer density afforded by the simulations. This is useful to boost statistics and access smaller-scale information. We verify the strong positive correlation between $R_{\text{eff}}$ and the mass of antihalos identified in Shim et al. (2021), which becomes almost monotonic for $M \gtrsim 10^{15}\,\odot/h$. This is in stark contrast to VIDE where the correlation is weak.

A reason to be wary of na"ive interpretation of $R_{\text{eff}}$ is given in Fig. 2, which shows the antihalos to be quite strongly elliptical. The modal ratio of principal axis lengths is $\sim1:2.5$. The voids are more elliptical than the corresponding halos because they are not bound structures supported by self-gravity, but rather prone to deformation by the high-density filaments and sheets that surround them. This has an important impact on their stacked density profiles (Sec. 3.2). We have checked that the antihalo major axes do not preferentially align with the line of sight, indicating that redshift space distortions are properly accounted for in BORG (Jasche & Lavaux 2019). Figs. 2 and 4 use a randomly-chosen step number of the BORG chain (8740), but all realisations produce consistent results.

Fig. 3 shows the cumulative volume filling fraction of antihalos and VIDE voids as a function of either $R_{\text{eff}}$ or $M$. Antihalos down to the minimum mass/size that we probe occupy $\sim70\%$ of the zoom-in region, while those with $M > 10^{15}\,\odot/h$ alone already occupy $\sim10\%$. In contrast, even the full set of VIDE voids occupies $<50\%$ of the available volume.

Although only included for three realisations in the public release, we also compute the antihalos’ specific angular moments of inertia around the VWB. This is typically in the range $1-10$ Mpc$^2$/s, around $100\times$ larger than for the halos in CSiBORG$^{-1}$.

3.2 Stacked density profiles

Fig. 4 shows the stacked density profiles of antihalos with $M > 10^{13}\,\odot/h$ using the four methods described in Sec. 3.2. All methods agree that the central regions are strongly underdense, with $\rho < 0.2\,\rho_m$ on average. However, using an elliptical rather than spherical averaging results in a larger central core as well as a significantly sharper rise in density near $R_{\text{eff}}$. This is because the elliptical method more accurately captures the 3D structure of the voids, while the spherical method washes out their features by combining regions that may be a very different number of principal axis lengths from the VWB. This also blurs out the “ridge” region at $\sim1-3\,R_{\text{eff}}$ caused by surrounding filaments and walls.

The other important difference is that using the volumes of the shells rather than the Voronoi volumes of the particles themselves (Eq. 1, blue and green lines) produces a marked decline in the density profiles at $R \gtrsim 2\,R_{\text{eff}}$. As already noted in Nadathur et al. (2015), this is an artifact caused by the fact that the shells can extend beyond the region in which parti-
The cumulative volume filling fraction of voids as a function of either size (blue and cyan) or mass (red and magenta).

Figure 4. Stacked antihalo density profiles as a function of distance from the VWB in units of $R_{\text{eff}}$. We consider spherical and ellipsoidal averages, and for each we consider two methods for calculating the density (Eqs. 1 and 2).

Clear positive correlation between the two void definitions and the inverse halos, and anticorrelation with the halos, affirms the expected antihalo behaviour. Applying the same mass cut to the VIDE catalogue does not, however, yield the same results. Due to the weak correlation between $M$ and $R_{\text{eff}}$ for VIDE, requiring $M > 10^{14} M_\odot/h$ leaves many small voids in overdense regions producing a positive correlation with the halos. The almost monotonic increase of mass with radius for antihalos makes their spatial distribution much more robust to the cut employed. We have also calculated the autocorrelation functions, and find that while $M \gtrsim 10^{13} M_\odot/h$ antihalos anticorrelate on all accessible scales, the VIDE voids only anticorrelate due to exclusion on scales comparable to their size. On larger scales, mass-selected samples can autocorrelate as strongly as halos. Correlation functions with any cuts may be easily explored using the public catalogues. Note that voids small compared to the $2.7 \, Mpc/h$ resolution of the BORG ICs are poorly constrained, with large variance between different CSiBORG$^{-1}$ realisations, so should be removed in studies where accurate positions are required.

3.4 Effect of the BORG ICs

To investigate the impact of the BORG constraints we have run a pair of simulations identical to CSiBORG but with random ICs. We find the antihalo results of Figs. 2, 4 and 5 to be consistent with CSiBORG, although we find somewhat more large halos and antihalos, and correspondingly fewer small halos and antihalos, compared to the unconstrained case. A similar trend is seen in the antihalo mass function, which exceeds that of Warren et al. (2006) for $\log(M/M_\odot) \gtrsim 14.35$ and lies $\sim 0.1$ dex below it at lower masses. This may indicate $\sim 3\sigma$ above average structure in the 2M++ volume compared to the cosmic mean or may be an artifact of the BORG algorithm or PHEW halofinder, and will be the subject of future work. These trends cannot be accounted for by altering $\Omega_m$ in the Warren et al. curve to match the average simulation density.
4 DISCUSSION AND CONCLUSION

We run a suite of simulations constrained to match the local Universe by using ICs inferred from the BORG algorithm applied to the 2M++ galaxy catalogue, and a corresponding suite with the ICs inverted. This enables us to apply a new voidfinder to out to 155 Mpc/h, in which voids are defined by dark matter particles belonging to halos in the inverted simulation. Running 101 inverted and uninverted simulations with ICs selected from across the BORG posterior allows us to propagate its uncertainty into our void catalogues, including the impact of density modes below the BORG grid. A key advantage of this voidfinder over traditional watershed methods is that the voids it identifies are linear to smaller scales and lower redshifts. In addition, the higher tracer density associated with the use of dark matter particles allows smaller voids to be identified and all voids to be characterised in more detail. We calculate a number of standard diagnostics of antihalos, finding a more extended size function, greater volume filling fraction, greater ellipticity and more stable correlation behaviour compared to VIDE voids. We make our antihalo and VIDE catalogues publicly available.

Our work leaves a number of important questions for the voids-as-antihalos method. The most important parameters of the voidfinder are those of the halofinder applied to the inverted simulation. While the PH EW algorithm that we use has been shown to reproduce standard statistics with the default settings (Bleuler et al. 2015), it will be important to investigate how varying these impacts the void population. It would also be useful to repeat the analysis using alternative algorithms such as friends-of-friends (Davis et al. 1985), HOP (Eisenstein & Hut 1998) or a phase-space halofinder. Note that the halofinder parameters do not contribute systematic uncertainty to the void properties but instead lead to different void definitions; the optimal parameters will be those that maximise sensitivity to whatever one wishes to test in a concrete application. The uncertainty in the void properties derives from that in the ICs (which our use of effectively the full BORG posterior allows us to marginalise over) as well as any approximations made during the simulation. One could also broaden the antihalo definition to include subhalos in the inverted simulation: it would be worth seeing if anti-subhalos have systematically different properties to anti-main-halos, and if they carry additional astrophysical or cosmological information. It would also be interesting to contrast their formation histories with those of halos, both readily available within our framework, and compare to voids found directly in the galaxy field.

Our void catalogues may find a range of applications, including studying secondary anisotropies of the Cosmic Microwave Background, characterising cosmic gas pressure, growth rate and star formation rate at low density, probing cosmological parameters and initial conditions, and searching for modified gravity (see Pisani et al. 2019 and references therein). We will explore some of these in future work. Note that while the BORG inference on which our analysis is based assumes ΛCDM, it is straightforward to incorporate additional physics within it and effort is underway to do so. This will enable direct comparison of the local voids produced under various scenarios. Similar catalogues could already be made in the SDSS-III/BOSS volume using the corresponding BORG reconstruction (Lavaux et al. 2019), and application of the algorithm to upcoming datasets from Euclid, the Roman Space Telescope or the Vera Rubin Observatory will greatly increase the accessible redshift range. Given the advantages of the antihalo definition, we believe that methods along these lines will be integral to future void science.

5 DATA AVAILABILITY

Our antihalo and VIDE catalogues are available here. Other data may be shared on request to the corresponding author.

ACKNOWLEDGEMENTS

We are indebted to D. Alonso, R. Gonzalez, N. Hamaus, M. Ivkovic, J. Jasche, G. Lavaux, C. Rycroft and R. Teyssier for input and discussions. HD is supported by St John’s College, Oxford. This work was done within the Aquila Consortium. This work used the DiRAC Complexity and DiRAC@Durham facilities, operated by the University of Leicester IT Services and Institute for Computational Cosmology, which form part of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment is funded by BIS National E-Infrastructure capital grants ST/K00373/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, STFC DiRAC Operations grant ST/K000359/1, and Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National E-Infrastructure.

REFERENCES

Bartlett D. J., Desmond H., Ferreira P. G., 2021, Phys. Rev. D, 103, 033523
Bleuler A., Teyssier R., Carassou S., Martizzi D., 2015, Computational Astrophysics and Cosmology, 2, 5
Davis M., Efthathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Desmond H., Hut M. L., Devriendt J., Slyz A., 2021, doi:10.5281/zenodo.5503610
Eisenstein D. J., Hut P., 1998, ApJ, 498, 137
Jasche J., Lavaux G., 2019, A&A, 625, A64
Jasche J., Wandelt B. D., 2012, MNRAS, 425, 1042
Jasche J., Wandelt B. D., 2013, MNRAS, 432, 894
Lavaux G., Hudson M. J., 2011, MNRAS, 416, 2840
Lavaux G., Jasche J., 2016, MNRAS, 455, 3169
Lavaux G., Jasche J., Leclercq F., 2019, arXiv e-prints, p. arXiv:1909.06396
Leclercq F., Jasche J., Sutter P. M., Hamaus N., Wandelt B., 2015, J. Cosmology Astropart. Phys., 2015, 047
Nadathur S., Hotchkiss S., 2015, MNRAS, 454, 2228
Nadathur S., Hotchkiss S., Diego J. M., Ilyev I. T., Gottlöber S., Watson W. A., Yepes G., 2015, MNRAS, 449, 3997
Neyrinck M. C., 2008, MNRAS, 386, 2101
Pisani A., et al., 2019, BAAS, 51, 40
Pontzen A., Slosar A., Roth N., Peiris H. V., 2016, Phys. Rev. D, 93, 103519
Rycroft C. H., 2009, Chaos, 19, 041111
Shim J., Park C., Kim J., Hwang H. S., 2021, ApJ, 908, 211
Sinha M., Garrison L. H., 2020, MNRAS, 491, 3022
Stopa S., Peiris H. V., Pontzen A., 2021, MNRAS, 500, 4173
Sutter P. M., et al., 2015, Astronomy and Computing, 9, 1
Teyssier R., 2002, A&A, 385, 337
Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881

MNRAS 000, 1–5 (2021)