



β -Voids:

Un Identificador de Vacíos
Cosmológicos en la Estructura de
Gran Escala para Catálogos de
Galaxias Basado en el
Grafo β -Skeleton

Felipe Leonardo Gómez Cortés

Universidad de los Andes
Facultad de Ciencias, Departamento de Física
Bogotá, Colombia
2019

β -Voids: Un identificador de vacíos cosmológicos en la estructura de gran escala para catálogos de galaxias basado en el grafo β -Skeleton

Felipe Leonardo Gómez Cortés

Trabajo de grado presentado como requisito parcial para optar al título de:
Magister en Ciencias - Física

Director:
Ph.D. Jaime Ernesto Forero Romero

Línea de Investigación:
Astrofísica Computacional
Grupo de Investigación:
AstroAndes

Universidad de Los Andes
Facultad de Ciencias, Departamento de Física
Bogotá, Colombia
2019

Para mis abuelos.

Agradecimientos

El camino de esta maestría lo he recorrido, paso a paso, siempre con el apoyo de seres humanos excepcionales. Muchos de ellos en su propia ruta académica, con experiencias similares. El apoyo mutuo ha sido fundamental para completar este camino y cerrar esta etapa.

Quiero agradecer a mis padres, Magdalena y Otto, por cultivar en mí la curiosidad que me llevó a optar por la ciencia, por todo el incondicional durante este proceso. A mi hermana, Anamaría, por enseñarme con su quehacer diario la disciplina y dedicación a su pasión: la música.

Gracias a Ángela por su comprensión y paciencia infinita para los momentos más críticos. Gracias a Jaime por el la dirección de este proyecto de investigación. Gracias a mi familia y a mis amigos que son la familia que uno escoge, por las palabras de ánimo en el camino. A Michel y Greg, Esteban, Mayra, Helena, Miguel, César. A Fredy, Diana, Stephany y Oscar.

A los amigos y compañeros de oficina de asistentes graduados, especialmente a Diego, John, Astaroth, Paula. A Rudy, Elizabeth, Nathaly, Alejandro, Angélica María y María Isabel. Y en especial a las guayabas de la i-115: Nicolás y Susana.

Gracias totales.

Resumen

Las grandes regiones de baja densidad en la Estructura de Gran Escala (EGE) del universo, también conocidos como vacíos, son características prominentes de la red cósmica que pueden ser utilizados como pruebas cosmológicas. Nosotros hemos desarrollado un nuevo algoritmo basado en el grafo β -Skeleton para encontrar vacíos en la EGE, este algoritmo puede operar sobre datos observacionales (posición de galaxias) y sobre datos simulados (posiciones de halos de materia oscura y galaxias). Métodos basados en el grafo β -Skeleton han sido utilizados ampliamente en machine learning, procesos de optimización y reconocimiento de imágenes, pero ha sido introducido recientemente en astrofísica como una herramienta para analizar la EGE. En el análisis que desarrollamos, sobre datos simulados, aproximamos los vacíos encontrados a elipsoides y estudiamos sus propiedades estadísticas como volumen y forma, encontrando resultados acordes con otros buscadores de vacíos.

Palabras clave: Cosmología: Estructura de gran escala del Universo. **Métodos:** Análisis de datos.

Abstract

Large underdense regions in the Large Scale Structure (LSS) of the Universe, also known as voids, are a prominent features of the cosmic web that can also be used as a cosmological probe. We have developed a new algorithm based on the β -Skeleton graph to find voids in the LSS, that can be applied to both observational and simulated data. Methods based on the β -Skeleton graph have been widely used on machine learning, optimization and image recognition. It has been recently introduced in astrophysics as a LSS analysis tool. The analysis we have performed, on simulated data, considers voids as ellipsoids. We study their statistical properties such as volumes and shape finding a good agreement with other void finders.

Keywords: Cosmology: Large-Scale Structure of Universe. **Methods:** Data analysis.

Contenido

| | |
|--|------------|
| Agradecimientos | VII |
| Resumen | IX |
| 1. Introducción | 2 |
| 1.1. Cosmología moderna | 2 |
| 1.2. Vacíos cosmológicos | 6 |
| 1.3. Buscadores de Voids | 8 |
| 2. El Algoritmo β-Voids | 9 |
| 2.1. El grafo β -Skeleton | 9 |
| 2.2. El Algoritmo β -Voids basado en β -Skeleton | 11 |
| 2.2.1. Catálogo de vacíos encontrados | 13 |
| 2.3. Prueba de concepto: modelos de vacíos de juguete | 13 |
| 3. Pruebas Sobre el Buscador | 15 |
| 3.1. Análisis estadístico de los Vacíos Cosmológicos | 15 |
| 3.1.1. Análisis morfológico de vacíos individuales | 15 |
| 3.1.2. Elipticidad y Prolacidad | 16 |
| 3.1.3. Análisis de poblaciones de vacíos | 16 |
| 3.1.4. Orientación isotrópica | 17 |
| 3.1.5. Función de tamaño de Vacíos | 17 |
| 3.1.6. Función de Correlación de Dos Puntos | 17 |
| 3.2. Efectos de la resolución del catálogo de entrada | 19 |
| 3.2.1. Catálogo de Entrada | 19 |
| 3.2.2. Función de abundancia de vacíos | 21 |
| 3.2.3. Morfología de los vacíos | 22 |
| 3.3. Efectos de la variación de los parámetros β y r | 24 |
| 3.3.1. Vacío más grande | 24 |
| 3.3.2. Función de tamaño de vacíos | 25 |
| 3.3.3. Fracción de ocupación de los vacíos | 26 |
| 3.3.4. Morfología de los vacíos | 27 |
| 3.3.5. Orientación isotrópica | 29 |
| 3.3.6. Función de correlación de dos puntos | 30 |

| | |
|---|-----------|
| 4. Conclusiones | 32 |
| A. Anexo: Gráficas de la variación de parámetros del algoritmo | 34 |
| Bibliografía | 39 |

1. Introducción

La cosmología física actual considera al Universo como un ente dinámico. Hay dos aspectos principales en esta evolución: la expansión que hace que la densidad global de materia disminuya con el tiempo cosmológico y, el segundo, la formación de estructuras localmente cada vez más densas debido al colapso gravitacional. Las observaciones del fondo de radiación cósmica de microondas [12] y de la distribución de galaxias a gran escala [2] apuntan a que esta evolución puede ser descrita por un puñado de parámetros cosmológicos, donde los más importantes son la densidad de materia y la densidad de energía oscura. El reto de la cosmología actual es aumentar la precisión de las mediciones de estos parámetros cosmológicos. Esto no sólo se logra con mediciones más precisas sino con métodos independientes para acotar los parámetros cosmológicos. Aunque un método independiente pueda tener una incertidumbre grande, considerar las cotas impuestas por varios métodos simultáneamente reduce la incertidumbre sobre los parámetros cosmológicos. Una de las pruebas cosmológicas que ha emergido en la última década es la caracterización de los vacíos cosmológicos; grandes regiones del espacio que cuentan con una baja densidad de galaxias.

1.1. Cosmología moderna

La cosmología pasó de ser algo especulativo a una ciencia exacta a comienzos del siglo XX. A finales del siglo XIX se consideraba que el universo estaba limitado a la Vía Láctea, lleno de estrellas y con algunos cuerpos nebulosos como Andrómeda dentro de ella. En esa transición se pasó de un universo limitado a la vía láctea a un universo mucho más vasto, lleno de universos isla extra-galácticos.

La formulación de la teoría general de la relatividad por Einstein permitió describir los primeros modelos del universo con soluciones como la propuesta por de Sitter [15], con un universo en expansión donde la luz de objetos distantes tendría un corrimiento al rojo, efecto encontrado experimentalmente por Hubble [25]. El trabajo de Friedmann [21] y Lemaître [31] permitió concluir que si el universo se encuentra actualmente en expansión, dando marcha hacia atrás en el tiempo, en algún momento la energía y la materia del universo se encontrarían en una singularidad con temperatura y densidad muy altas, llamada por los detractores de este modelo, el Big-Bang. Uno de los fundamentos para el desarrollo de esta teoría es la homogeneidad y la isotropía en la distribución de materia a gran escala en el universo. Esto mostraban las observaciones de Hubble [26], con un universo poblado

uniformemente por galaxias, y aparentemente sin fronteras hasta donde lo permitía el límite de resolución y magnitud de los telescopios y las placas fotográficas.

A la par, la física subatómicas hacían su aparición. Hacia 1946 Gamow y Alpher, y luego trabajando junto con posteriormente Herman [4], establecen la formación de núcleos atómicos a partir de el caldo primigenio de neutrones, protones y electrones a altas temperaturas, permitiendo reacciones nucleares en un entorno bañado por rayos gamma. Al disminuir la temperatura lo suficiente, el universo quedaría completamente transparente a la radiación y esta quedaría viajando libremente hasta la fecha. En 1964 Dicke y Peebles trabajaban en la formulación teórica de esta radiación reliquia del universo temprano, cuando Wilson y Penzias detectaron, sin proponérselo, la radiación cósmica de fondo. El resultado fue publicado por Dicke et al. [16] y Penzias and Wilson [37].

La observación del cielo cada vez más profundo permitió conocer los cúmulos de galaxias. En la década de los 50, con observaciones del Palomar Observatory Sky Survey (POSS) mostró evidencia clara de una estructura en la distribución de galaxias a gran escala, agrupándose en nodos, filamentos y presentando regiones de baja densidad. En los 80's, con mapas tridimensionales del universo observable, utilizando redshift espectroscópico como el Harvard-Smithsonian Center for Astrophysics, se hizo evidente esta estructura de gran escala con supercúmulos de galaxias, estructuras filamentarias, paredes y vacíos que podían llegar a los cientos de Megapársecs,

En 1933 Zwicky [51] utiliza el teorema del virial para calcular la masa contenida en el cúmulo de Coma a partir de mediciones en las velocidades de las galaxias del cúmulo. Encontró que la masa calculada a partir de la luminosidad era mucho menor a la calculada a partir del teorema del virial. En 1970 el trabajo de Rubin and Ford [42], encontraron con espectroscopía óptica que la masa estelar no era suficiente para explicar las curvas de rotación de galaxias. Observaciones en radio dieron soporte a estos trabajos, y más adelante observación de lentes gravitacionales en clústers de galaxias y nubes de gas con fuerte emisión de rayos X dieron la evidencia fulminante para aceptar esta materia faltante dentro. A la materia no luminosa faltante se le acuñó el término de Materia Oscura.

A finales de los 80's, con la sonda espacial COBE se tienen las primeras mediciones precisas de la radiación cósmica de fondo, y esta precisión ha venido mejorando con los datos de las misiones WMAP (2001-2010) y Planck (2009-2013). Esta medición muestra, por un lado, que el universo tiene curvatura nula, o muy cercana a cero. Por otro lado, las anisotropías en esa radiación de fondo establecen regiones distantes (a una distancia angular de 2°) se encontraban conectadas causalmente en un universo temprano[9].

El aumento del volumen de datos manejado en los catálogos de galaxias ha exigido nuevas herramientas estadísticas para interpretar la física detrás de la estructura del universo. Las funciones de correlación y el espectro de potencias[36] permiten encontrar información oculta en la distribución de galaxias como las oscilaciones acústicas bariónicas (BAOs). Las BAOs son reliquias de las perturbaciones iniciales: Las sobredensidades iniciales al colapsar aumentan su temperatura y presión dada por la interacción materia-energía empuja hacia

afuera en un proceso oscilatorio que se propaga como ondas de sonido, hasta que se desacoplan materia y energía por la expansión y enfriamiento del universo, dejando una huella del mismo tamaño en todo el universo. Estas fueron detectadas en catálogos de redshift de galaxias (SDSS y 2dFGRS) hacia 2005.

Las supernovas tipo Ia tienen un brillo estándar, en 1998 dos grupos que estudiaban encontraron evidencia de un universo en expansión acelerada, idea contraintuitiva si se considera solo la gravedad como el componente dominante en grandes escalas que desaceleraría la expansión, es decir, nos encontramos en una era del universo en expansión acelerada con una constante cosmológica $\Lambda > 0$.

El modelo cosmológico estándar Λ CDM se ha construido para explicar las observaciones de la radiación cósmica de fondo, la abundancia de elementos químicos, la estructura de gran escala del universo y la expansión acelerada del mismo [39]. Está basado en la métrica de Robertson-Lemaître-Robertson-Walker para un universo plano en expansión que tuvo su origen en el Big-Bang, considera la existencia de la constante cosmológica Λ introducida por Einstein, materia bariónica (visible) y materia oscura fría (con dispersión de su velocidad mucho menor a la velocidad de la luz) no colisional y con una interacción débil o nula con la luz. Con seis parámetros independientes se pueden reproducir las propiedades estadísticas del universo observable: la constante de Hubble, la ecuación de estado, la densidad de materia bariónica, la densidad de materia oscura fría, la amplitud del espectro de potencias en escala de $h^{-1}\text{Mpc}$ y el índice espectral[23].

El trabajo de interpretación de estos mapas en términos de parámetros cosmológicos requiere la realización de simulaciones que sigan la formación de estructuras en un Universo en expansión. Las simulaciones de N-cuerpos interactuando tienen su origen en los años 60's con los trabajos de von Hoerner [49] que inicialmente tenían entre 16 y 25 partículas, diez años después Aarseth [1] habían alcanzado las 500 partículas. Para 1979 se empezaron a hacer experimentos numéricos para procesos de agrupamiento de galaxias con hasta 4000 partículas [24]. Una de las simulaciones más emblemáticas es *Millennium Run*[45] hecha en el Instituto Max Planck de Astrofísica, que ya cumple 15 años. En ella tienen 2160^3 partículas de masa $8.6^8 M_\odot$, dentro de una caja cúbica de lado $500h^{-1}\text{Mpc}$ con condiciones de frontera periódicas y resolución espacial de $5h^{-1}\text{kpc}$. La simulación corre desde $z = 127$ hasta $z = 0$, logrando definir con suficiente detalle la estructura de gran escala para aplicar modelos semi analíticos de formación de galaxias dentro de los halos de materia oscura.

Entre los problemas abiertos que deja el modelo Λ CDM están el problema del horizonte y el problema de la curvatura. El problema del horizonte yace en que regiones distantes en el universo en la CMB están en equilibrio térmico, pero a una escala de distancias que no permitiría causalidad en el Big-Bang. El problema de la curvatura que se ha medido cercana a cero requiere ajuste fino en las condiciones iniciales para no dispararse con la expansión del universo. Allan Guth propone en los 80's la teoría de inflación, que explica cómo el universo en un tiempo comparable a la escala de Planck (desde $10^{-36}s$ hasta $10^{-34}s$ después del Big-Bang) se expande por un factor de 3×10^{43} . (Carroll & Ostlie 2007). Se esperaría ver huellas

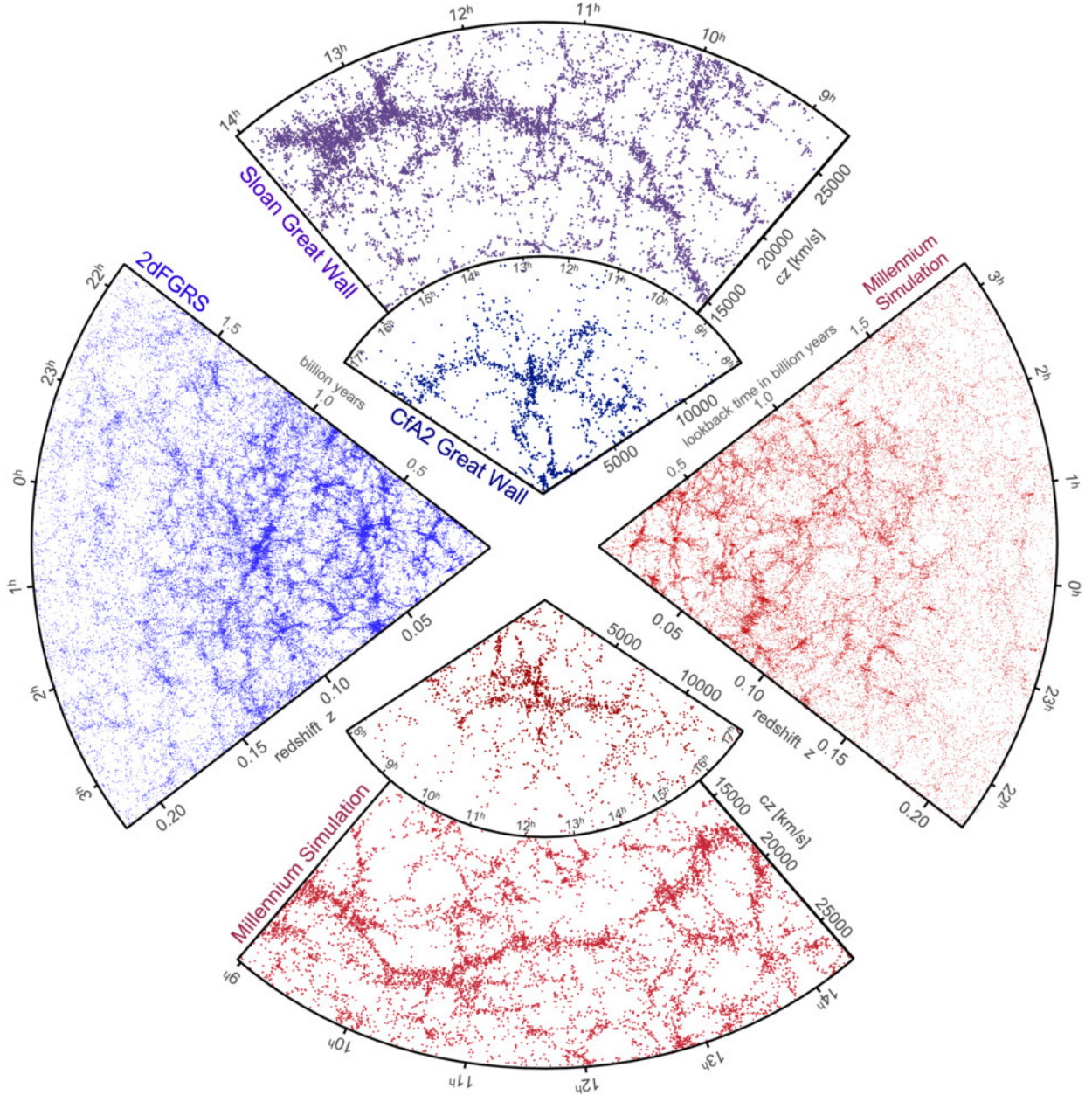


Figura 1-1.: Distribución espacial de galaxias observada en mapeos como el SDSS y el 2dFGRS (en azul) comparadas con resultados de la simulación Millennium (rojo). Cada punto representa una galaxia. La Vía Láctea se encuentra en el centro de los mapas, que cubren diferentes regiones del cielo en escalas distintas. Crédito: Max-Planck-Institute für Astrophysik <https://wwwmpa.mpa-garching.mpg.de/millennium/>

de este periodo de inflación en la polarización en modos B de la CMB (en curso, BICEP). Otro problema es la asimetría entre materia y antimateria. Hasta donde sabemos, el universo está compuesto de materia, pero previo a la reionización, el plasma primigenio tenía creación por igual de pares partícula antipartícula. A esto se añade la naturaleza misma de la materia oscura y la energía oscura. Desde la física de partículas se hace una búsqueda intensa de candidatos de partículas de materia oscura teórica y experimentalmente. Se tienen proyectos para crear mapas tridimensionales de mejor resolución (DESI y LSST), estos permitirán aclarar la tensión entre las mediciones de la constante de Hubble en universo temprano (Planck) y universo cercano. Es aquí donde los vacíos cosmológicos entran a jugar un rol fundamental por ser escenarios de evolución dominados por la energía oscura.

1.2. Vacíos cosmológicos

El ojo humano puede identificar de inmediato los vacíos cosmológicos en mapas de la distribución tridimensional de galaxias [2]. Sin embargo, a la fecha no existe una definición única de vacío cosmológico [30] entre los algoritmos buscadores de vacíos en la estructura de gran escala. Esto no es impedimento para continuar en la búsqueda continua de mapas del universo más profundos y detallados para comprender mejor la formación y evolución de la estructura de gran escala. El proyecto de mapeo más emblemático de la primera década del siglo XXI fue el *Sloan Digital Sky Survey (SDSS)* el cuál con el experimento BOSS ha estimado la posición en el espacio de más de 1.5 millones de galaxias de redshift $0 < z < 1.0$.

La Figura 1-1 compara los mapas de la distribución de galaxias obtenidos por simulaciones (rojo) y por observaciones (azul). Cada punto representa una galaxia. En estas imágenes los grandes filamentos donde se aglomeran las galaxias se complementan con las regiones donde hay menos galaxias que conforman los vacíos que nos interesan en esta tesis. En el segmento de la izquierda muestra la estructura a gran escala compuesta por filamentos donde se agrupan las galaxias y grandes vacíos donde es baja la densidad de galaxias, del catálogo 2dFGRS. Hacia la periferia disminuye la densidad de galaxias: solo se observan las galaxias más brillantes. A la derecha, un segmento de la simulación Millennium del mismo tamaño con las misma magnitud límite. En los segmentos superiores se muestran estructuras de varias decenas de Mpc encontradas en observaciones en ambos catálogos (SDSS y 2dFGRS). En los dos segmentos inferiores, vemos estructuras con tamaños similares generadas por la misma simulación Millennium.

La estructura de gran escala muestra que la mayor parte del volumen del Universo está dominado por los vacíos de la red cósmica. Allí la densidad de materia es inferior al 20 % de la densidad media del universo [48]. El universo inicialmente homogéneo tiene una evolución con dinámica no lineal en regiones con sobredensidad donde se forman estructuras de materia ligadas gravitacionalmente cuyos marcadores son las galaxias (agrupadas en filamentos, paredes y nodos). Se han desarrollado formalismos para obtener aproximaciones por teoría de perturbaciones para la evolución de las sobredensidades partiendo del modelo Top-Hat que

conducen al colapso de estas estructuras y la posterior formación de halos de materia oscura. Entre los pioneros en estos modelos analíticos están Zel'Dovich [50] y Press and Schechter [41].

Los modelos anteriores pueden describir hasta cierto punto el colapso antes de entrar en un régimen no lineal donde la densidad es mayor a 200 veces la densidad crítica en el redshift de observación [43]. Así que se puede aplicar un análisis similar a los vacíos, que regiones donde la densidad tiene una cota inferior, cero, a la cual tiende asintóticamente, trabajando en todo momento en el régimen lineal.

El modelo SvdW Sheth and van de Weygaert [44] estudia vacíos aislados en un universo homogéneo, describe los vacíos como: 1) regiones esféricas en expansión, 2) a medida que se expanden la materia está siendo evacuada de ellos, reduciendo cada vez más la densidad, 3) el perfil de densidad se asemeja a un perfil de Top-Hat invertido que se hace cada vez más profundo debido al potencia repulsivo efectivo, 4) en la expansión y evacuación de materia se presentan velocidades mayores a la velocidad de Hubble (“*Super-Hubble velocity field*”), 4) dada la baja densidad y el potencial repulsivo se suprime la formación de estructuras, 5) la materia expulsada termina acumulándose en los bordes de los vacíos y finalmente, 6) en la expulsión de material a distintas velocidades se dan fenómenos de capas internas avanzando hacia afuera más rápido que capas externas, eventualmente se sobreponen las capas en los bordes del vacío. A este modelo se le añaden un par de condiciones que rompen la simetría con las sobredensidades: los vacíos formados dentro de otros vacíos terminan siendo asimilados por el vacío más grande (void-in-void), y los vacíos formados dentro de sobredensidades terminan diluyéndose (void-in-cloud). Con esto se completa el *excursion set formalism*, que permite extender las propiedades de vacíos individuales (perfiles de densidad, velocidad de evacuación de materia) extendiéndolo hasta estadísticas de población, (función de abundancia y fracción de ocupación del volumen total del universo). Si bien, los vacíos pueden ser identificados como estructuras irregulares, se ha mostrado que una descripción de los vacíos cosmológicos como elipsoides triaxiales resulta mejor que la descripción en primera aproximación como esferoides.

En estas regiones de baja densidad la evolución está dominada por la energía oscura. Aquí yace el interés por estudiar los vacíos en la red cósmica para explorar parámetros cosmológicos. A partir del análisis estadístico de la morfología y la dinámica de los vacíos se pueden estudiar los parámetros cosmológicos. Las condiciones iniciales de distribución de materia en el Universo se toman como el modelo analítico de Zel'dovich. En Park and Lee [35] demuestran cómo la elipticidad de los vacíos es sensible a los parámetros cosmológicos, como un balance entre efectos de marea producidos por la componente de materia del Universo y la expansión misma del Universo: las perturbaciones iniciales tienden a crecer esféricamente por la expansión del Universo, aquí entra el parámetro Ω_Λ , en tanto las fuerzas de marea (la contribución de Ω_m) tienden a deformar los vacíos, alejándolos de la forma esférica. Desarrollan modelos analíticos sobre la distribución de elipticidad para la población de vacíos según los parámetros cosmológicos. En Bos et al. [6] y Pisani et al. [38] se compa-

ran incluso cosmologías distintas al modelo estándar Λ CDM, de nuevo, por medio del análisis morfológico de los vacíos.

1.3. Buscadores de Voids

En Lavaux and Wandelt [30] se proponen tres categorías principales para clasificar los algoritmos de buscadores de vacíos: basados en densidad, geometría y criterios dinámicos. La mayoría de los algoritmos presentados en Colberg et al. [11] caen en la primera y segunda categoría. En la primera categoría se identifican los vacíos como regiones donde la densidad es inferior a una densidad crítica, típicamente del 20 % de la densidad media del universo. Dentro de esta categoría se encuentra el algoritmo de Padilla et al. [34], que identifica los vacíos dentro de catálogos de halos o galaxias (semi-analíticas) como regiones esféricas donde la densidad media es inferior al 10 % de la densidad del universo, este umbral, sin embargo, puede llegar hasta el 40 % sin afectar significativamente el tamaño de los vacíos encontrados. El centro de dichas esferas se genera aleatoriamente y la tasa de éxito aumenta al incrementar el número de puntos aleatorios.

En Cannarozzo [8] se destacan los algoritmos basados en criterios geométricos de Neyrinck [33] ZOBOV (ZOnes Bordering On Voidness) y Platen et al. [40] WST (WaterShed Transform), basados ambos en teselación de Voronoi para el cálculo y reconstrucción del campo de densidad, para luego utilizar la transformación de cuenca (Watershed Transform), ambos códigos terminan siendo precursores del VIDE Sutter et al. [46]. De otro lado, está el algoritmo de Foster and Nelson [20] que utiliza como criterio la distancia al tercer vecino más cercano en cada galaxia, para definir si una galaxia hace parte de una pared o si hace parte del campo (field galaxy) y le está permitido permanecer dentro de un vacío, luego se discretiza el espacio en celdas, se conectan las celdas donde hayan galaxias de pared y se identifican los vacíos como las celdas de campo conectadas.

En la categoría de algoritmos basados en criterios dinámicos se encuentran: Forero-Romero et al. [19] que a partir del Hessiano del potencial gravitacional distingue nodos, filamentos, paredes y vacíos; DIVA presentado por Lavaux and Wandelt [30] que identifica los vacíos por medio de reconstrucción de órbitas.

2. El Algoritmo β -Voids

2.1. El grafo β -Skeleton

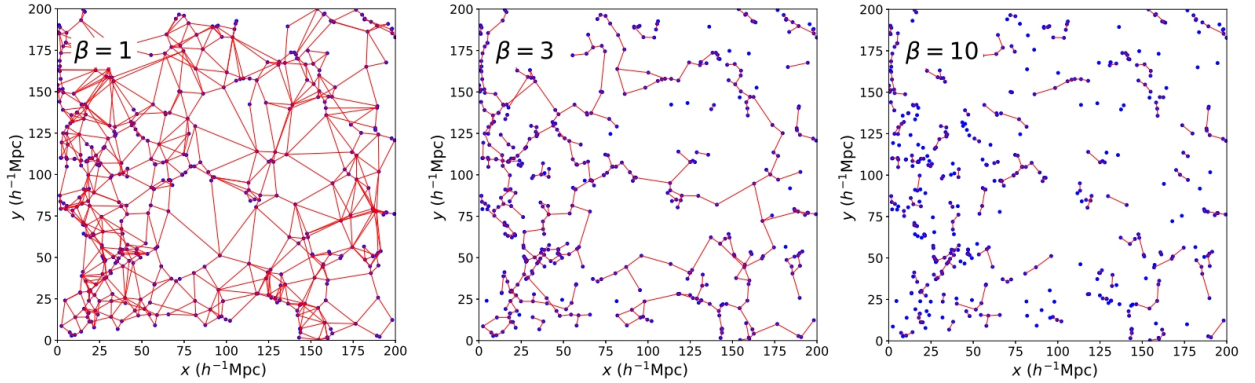


Figura 2-1.: Grafo β -Skeleton calculado sobre la estructura de galaxias a gran escala con tres valores diferentes $\beta = 1, 3, 10$. Crédito: Fang et al. [18]

El método β -Skeleton ha sido introducido en astrofísica recientemente por Fang et al. [18] para encontrar el “esqueleto” de la EGE en catálogos de simulaciones y datos observacionales. Ha sido usado desde hace ~ 40 años para analizar imágenes en búsqueda de contornos e identificación de objetos, tiene aplicaciones en *machine learning*, problemas de minimización en conexión de redes inalámbricas. Este método es computacionalmente económico y rápido.

Depende del parámetro continuo $\beta \geq 0$ para establecer las conexiones entre conjuntos de puntos en el espacio. Cuando este parámetro tiende a cero, el método calcula todas las conexiones entre todos los puntos de la red. Se ha encontrado que valores de este parámetro cercanos a 3 generan las estructuras filamentosales en la EGE. Valores cercanos a 10 reducen demasiado las conexiones y se pierde la estructura filamental (figura 2-1).

El β -Skeleton es un grafo no direccional que describe relaciones de proximidad entre un conjunto de puntos de dos o tres dimensiones mediante vértices que los conectan según un criterio basado en el parámetro continuo β .

A principios de los 80's la línea de investigación en reconocimiento de patrones empezó a explorar variaciones de los grafos bien conocidos para conjuntos de puntos en el plano euclidiano. En Edelsbrunner et al. [17] definen por primera vez el cascarón- α (α -hull) de un conjunto de puntos, con el parámetro $-\infty < \alpha < \infty$ que encierra el conjunto completo

de puntos. Se define también la α -forma como la colección de segmentos de arco de radio $1/\alpha$ que conectan puntos de frontera. Con valores positivos se tienen regiones convexas, con valores negativos se tienen regiones cóncavas que determinan los contornos.

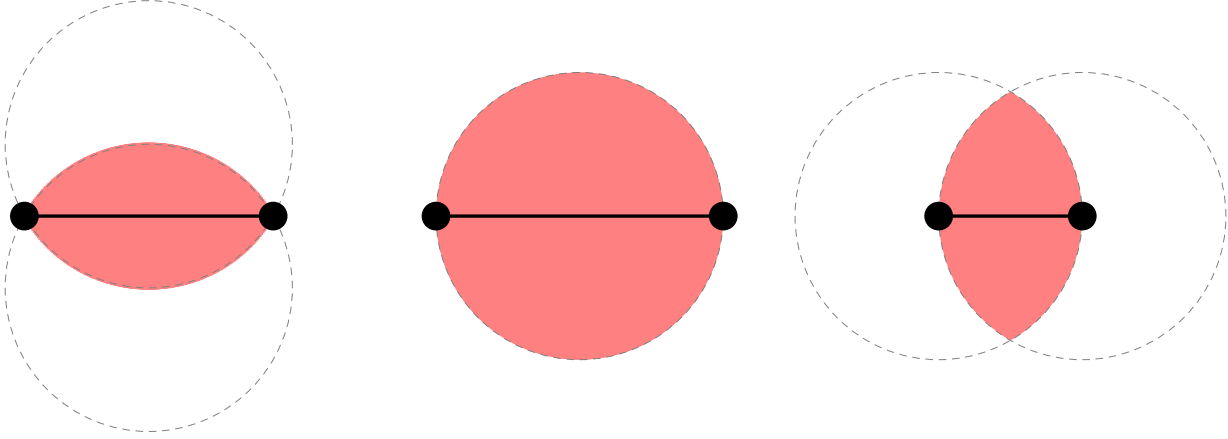


Figura 2-2.: Zona de exclusión (sombreada) en el grafo β -Skeleton. Izquierda: $\beta < 1$. Centro: $\beta = 1$, se tiene el Grafo de Gabriel. Derecha: $\beta > 1$ basado en la definición de luna. Crédito: Fang et al. [18]

Kirkpatrick and Radke [27] trabajan en términos del parámetro β , definiendo si dos puntos están conectados entre sí, si y solo si dentro de una zona de exclusión entre los dos no hay un tercer punto. El dominio de este parámetro será $\beta \in (0, \infty)$. Si la distancia entre los dos puntos p y q es d_{pq} la zona de exclusión se define en la figura 2-2 como:

- $\beta < 1$: La intersección entre dos círculos de radio $R = d_{pq}/2\beta$. Sus centros se encuentran sobre la recta perpendicular a la recta que pasa por p y q , y contienen a p y q sobre su frontera.
- $\beta = 1$: El círculo de radio $R = d_{pq}/2$ centrado en el punto medio entre p y q .
- $\beta > 1$: La intersección entre dos círculos de radio $R = \beta d_{pq}/2$ centrados en $p + \beta(q - p)/2$ y $q + \beta(p - q)/2$.

La anterior corresponde a la definición de lunas para $\beta > 1$. Existe también una definición basada en círculos, similar a los círculos para $\beta < 1$, pero ahora tomando la unión de los dos conjuntos en lugar de su intersección. De aquí en adelante se trabajará solamente con la definición de lunas.

Al tomar el límite $\beta \rightarrow 0$, las zonas de exclusión se reducen a vecindades infinitesimales alrededor del segmento de recta \overline{pq} . A menos que se tengan puntos colineales, el grafo que se construye con $\beta \rightarrow 0$ conecta cada nodo con todos los demás nodos, para una red de n nodos se tiene un número de conexiones del orden $\mathcal{O}(n^2)$. El grafo para $\beta = 1$ corresponde al

Grafo de Gabriel [22], y el grafo $\beta = 2$ corresponde al Grafo de Vecinos Relativos, Relative Neighbour Graph (RNG) en la definición de Toussaint [47]. Al tomar el límite $\beta \rightarrow \infty$ los círculos que definen la zona de exclusión se hacen muy grandes, resultando en una franja infinita comprendida entre las rectas perpendiculares al segmento \overline{pq} pasando por p y q , reduciendo seriamente el número de conexiones incluso hasta cero. Se tiene entonces la siguiente propiedad de los β -Skeleton: si $\beta_1 < \beta_2$, todas las conexiones del grafo construido con β_2 están contenidas dentro del grafo β_1 .

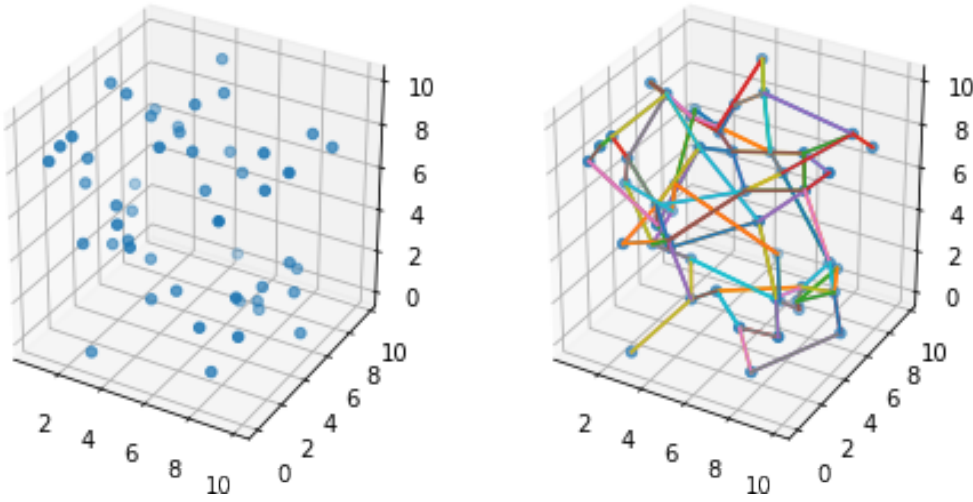


Figura 2-3.: Conjunto tridimensional de puntos (izquierda). 1-Skeleton asociado (derecha)

La extensión de dos a tres dimensiones es directa. Las zonas de exclusión se pueden construir como sólidos de revolución rotando las zonas de exclusión bidimensionales sobre la recta que contiene a p y q . [14]

La primera implementación del β -Skeleton en astrofísica la desarrollan Fang et al. [18], estudiando el número de conexiones, la longitud y la orientación de las mismas.

2.2. El Algoritmo β -Voids basado en β -Skeleton

Queremos encontrar los vacíos en la Estructura Cosmológica de Gran Escala llenándolos con puntos aleatorios y luego usando el β -Skeleton para identificarlos y agruparlos.

El conjunto de datos de entrada **DC** consiste en puntos tridimensionales (figura 2-4 panel izquierdo), que bien pueden provenir de catálogos de corrimiento al rojo de galaxias como el SDSS [2], o posiciones de halos de materia oscura provenientes de simulaciones como [23]. Este catálogo de entrada tiene N_d datos.

Poblamos el mismo volumen de los datos de entrada con un catálogo de N_r puntos **RC** aleatorios con distribución uniforme (figura 2-4, panel central). La idea es llenar los

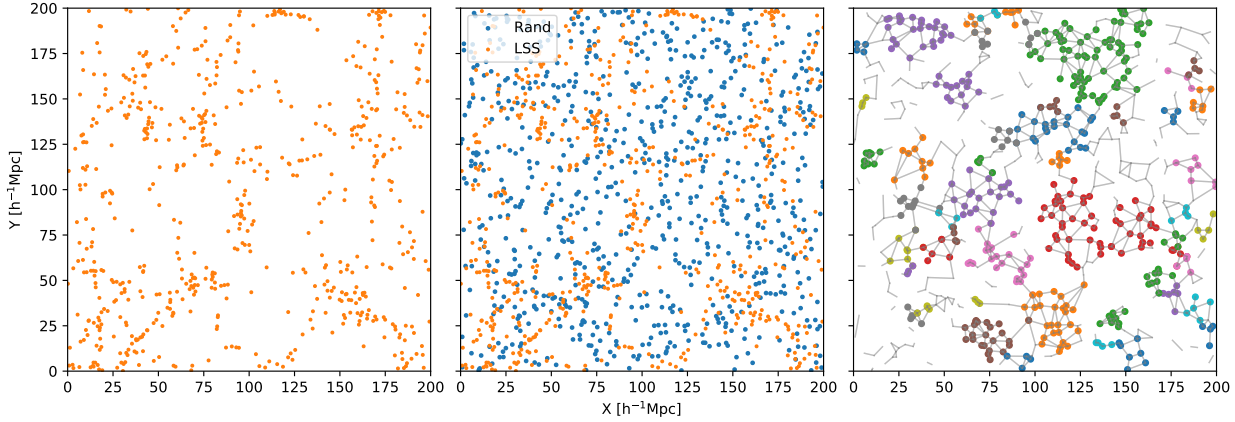


Figura 2-4.: Conjunto de datos de entrada correspondiente a marcadores de la Estructura Cosmológica de Gran Escala (izquierda). El volumen es poblado con puntos obedeciendo a una distribución aleatoria uniforme, en azul (centro). El β -Skeleton es calculado y se quedan solamente puntos aleatorios conectados con otros puntos aleatorios, estos definen los vacíos de la ECGE (derecha)

vacíos cosmológicos con una cantidad representativa de puntos aleatorios para que estos tomen la forma de los vacíos de la mejor forma posible, así como un gas toma la forma de su contenedor. La proporción de puntos aleatorios sobre datos $r = N_r/N_d$ es el primer parámetro del algoritmo. Tener una proporción $r < 1$ resulta en una descripción pobre de los vacíos. Parte de estos puntos aleatorios caerán dentro de la estructura de gran escala, estos no serán tenidos en cuenta para describir los vacíos.

Construimos el catálogo completo **FC** como la unión de los catálogos RC y DC. Sobre este catálogo se calculamos el β -Skeleton. Para esta construcción no se tiene distinción alguna entre puntos aleatorios y puntos de datos. β es el segundo parámetro libre del algoritmo β -Voids. El β -Skeleton del FC contiene toda la información de tamaño y forma de los vacíos.

El β -Skeleton se calcula utilizando la librería desarrollada en Fortran por [18], modificada para correr en el Clúster de Cómputo de Alto Rendimiento (HPC) de la Universidad de Los Andes. El archivo de salida contiene la información de todos los vértices del grafo: es un arreglo de dos columnas, la primera contiene el nodo de origen y la segunda el nodo final.

Ahora sigue definir cuáles puntos aleatorios caen dentro de los vacíos y cuáles caen en la EGE. Definimos entonces que un punto del catálogo aleatorio hace parte de un vacío si y solo si todas sus conexiones son con otros puntos aleatorios. Aplicar este criterio de exclusión es sencillo: como el FC se construye apilando primero el RC y luego el DC, al eliminar de la primera columna del β -Skeleton los nodos mayores a N_r se van todos los puntos del DC, y luego se descartan todos los nodos que en la segunda columna tengan conexiones con puntos del DC. Finalmente, por medio de una función recurrente, se inicia buscador en el primer punto que pertenece a algún vacío cuáles son sus vecinos, se crea guardan en una lista y la función recurrente repite el proceso recorriendo sobre cada elemento de la lista que solo

tenga conexiones con otros puntos aleatorios. Se tiene un arreglo paralelo de control para evitar ciclos infinitos de recurrencia. Cuando se terminan de encontrar todos los puntos que pertenecen a este primer vacío, el algoritmo continúa la búsqueda en el siguiente punto que pertenezca a algún vacío aún no reconocido.

El algoritmo entrega en esta etapa una lista maestra donde se especifica para cada partícula identificada como perteneciente a un vacío su posición y el identificador del vacío al cuál pertenece.

2.2.1. Catálogo de vacíos encontrados

Las propiedades estadísticas más relevantes de los vacíos se definen en el capítulo 3. Para esto calculamos un elipsoide asociado a la distribución de puntos aleatorios de cada vacío encontrados por el algoritmo β -Voids y tenemos las siguientes columnas de salida para cada void:

- Identificación: tipo entero.
- Número de partículas: tipo entero. Indica la cantidad de partículas aleatorias que lo conforman.
- Centro del vacío: expresado en las mismas coordenadas X,Y,Z del catálogo original, tres columnas tipo float.
- longitud de los semiejes a, b, c , tres columnas tipo float
- vectores unitarios de los semiejes: nueve columnas tipo float.

2.3. Prueba de concepto: modelos de vacíos de juguete

Antes de probar el algoritmo buscando vacíos con formas desconocidas en datos de la LSS, realizamos pruebas con modelos de vacíos de juguete como control para comprobar si el algoritmo encontraba la forma esperada o no.

Las primeras pruebas se realizaron en una caja cúbica de 100 unidades de lado poblada con 5000 puntos aleatoriamente distribuidos. El primer vacío de control fue una esfera de radio 20 unidades situada en el centro de la caja. El segundo control fue la identificación de varias esferas en distintas posiciones dentro de la misma caja. El tercer control fue la identificación de vacíos con forma elipsoidal con distintas elipticidades.

En la última prueba colocamos algunos vacíos elipsoidales en superposición con otros formando un vacío irregular. El algoritmo logró identificar exitosamente los vacíos de control (identificando como uno solo el vacío irregular) junto con pequeñas estructuras que surgieron en la creación del DC aleatorio como se puede ver en la figura 2-5.

Esto nos dio luz verde para iniciar pruebas con catálogos de puntos de la Estructura de Gran Escala.

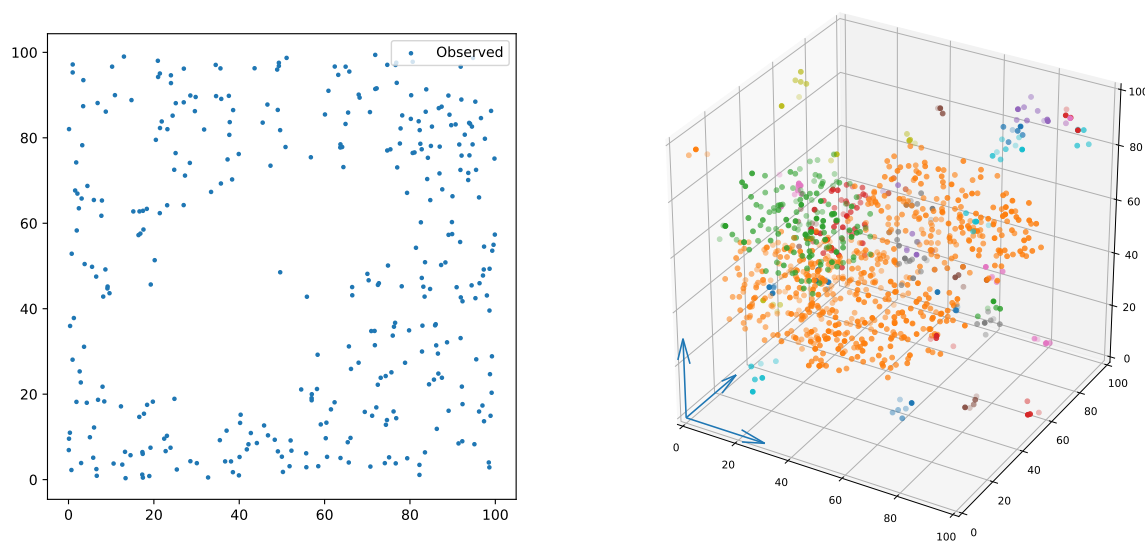


Figura 2-5.: Prueba de concepto. Izquierda: Una tajada de la caja cúbica de 100 unidades de lado donde el catálogo de datos DC que preparamos como una distribución de puntos aleatorios con cuatro vacíos elipsoidales, algunos intersectando para formar un vacío irregular. Derecha: El algoritmo identificó los vacíos de control incluyendo el vacío irregular (amarillo).

3. Pruebas Sobre el Buscador

3.1. Análisis estadístico de los Vacíos Cosmológicos

3.1.1. Análisis morfológico de vacíos individuales

El algoritmo identifica los vacíos como conjuntos tridimensionales de puntos aleatorios conectados por el grafo β -Skeleton. Con estos puntos podemos calcular el centro y la forma por medio de un pseudo-tensor de momentos de inercia.

La posición de la n -ésima partícula que hace parte de un vacío viene dada por $\mathbf{x}_n = [x_1, x_2, x_3]$ en el sistema de coordenadas cartesianas del catálogo de entrada. La posición del centro de un vacío se determina a partir de la posición de las N partículas que lo conforman según

$$\mathbf{x}_c = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad (3-1)$$

Si bien la morfología de los vacíos es irregular, el análisis estadístico puede abordarse considerándolos en primera aproximación como esferoides ([34]), o bien como elipsoides ([35, 3]). Estableciendo el origen del sistema de coordenadas, se puede calcular un pseudo-tensor de momentos de inercia como

$$I_{ij} = \frac{1}{N} \sum_{n=1}^N |\mathbf{x}_n|^2 \delta_{ij} - x_{n,i} x_{n,j}. \quad (3-2)$$

La diagonalización de I_{ij} arroja tres autovectores $I_a \leq I_b \leq I_c$ relacionados con el elipsoide descrito por:

$$\left(\frac{x'}{a}\right)^2 + \left(\frac{y'}{b}\right)^2 + \left(\frac{z'}{c}\right)^2 = 1 \quad (3-3)$$

con x', y', z' definidos en el sistema de coordenadas ortogonal dado por las componentes unitarias de los autovectores del pseudo-tensor de momentos (ec. 3-2).

Recordando que para un elipsoide sólido con densidad uniforme y masa $M = 1$, con sus semiejes alineados con el sistema de coordenadas, su momento de inercia es:

$$I = \frac{1}{5} \begin{bmatrix} b^2 + c^2 & 0 & 0 \\ 0 & c^2 + a^2 & 0 \\ 0 & 0 & a^2 + b^2 \end{bmatrix} = \begin{bmatrix} I_a & 0 & 0 \\ 0 & I_b & 0 \\ 0 & 0 & I_c \end{bmatrix} \quad (3-4)$$

Los semiejes del elipsoide asociado el vacío en cuestión están dados por:

$$a = \sqrt{\frac{5}{2}(-I_a + I_b + I_c)} \quad b = \sqrt{\frac{5}{2}(I_a - I_b + I_c)} \quad c = \sqrt{\frac{5}{2}(I_a + I_b - I_c)}, \quad (3-5)$$

con esto tenemos el volumen del elipsoide $V = \frac{4}{3}\pi abc$ y definimos el radio efectivo $R_{eff} = \sqrt[3]{abc}$ como el radio de una esfera con el mismo volumen del elipsoide asociado.

3.1.2. Elipticidad y Prolacidad

Dentro de la marco de consideración de los vacíos como elipsoides, es útil tener una medida numérica que cuantifique qué tan cercanos o no son a una esfera. El trabajo de Park and Lee [35] estudia el cambio de morfología de los vacíos en diferentes escenarios cosmológicos utilizando la definición de elipticidad:

$$\epsilon = 1 - \frac{c}{a}, \quad (3-6)$$

para una esfera los semiejes son iguales y la elipticidad es cero.

Sin embargo, dentro de los elipsoides existen dos formas notables: prolotos y oblatos. Un elipsoide se define prolato si uno de sus semiejes domina sobre los otros dos, resultando en elipsoides con forma alargada como un cigarro. Un elipsoide se define oblato si dos semiejes son mucho mayores que el tercero, resultando en una forma achatada como un tipo pancake u oblea. Trabajos recientes, tanto en vacíos cosmológicos como en halos de materia oscura, se ciñen a la definición de prolacidad y elipticidad para elipsoides triaxiales propuesta por Bardeen et al. [5]. Nadathur [32] realiza esta cuantificación en términos de los momentos de inercia en tanto [3] realizan sigue la misma definición pero en términos de los semiejes calculados, quedando la elipticidad definida como:

$$e = \frac{1}{4} \frac{a^2 - c^2}{a^2 + b^2 + c^2} \quad (3-7)$$

y la prolacidad:

$$p = \frac{1}{4} \frac{(b^2 - a^2) + (b^2 - c^2)}{a^2 + b^2 + c^2}, \quad (3-8)$$

quedando los valores de la prolacidad restringidos a $-e \leq p \leq e$. Elipsoides oblatos tendrán valores de $p > 0$ y elipsoides prolotos tendrán $p < 0$.

3.1.3. Análisis de poblaciones de vacíos

Una vez hemos caracterizado cada vacío en el catálogo, procedemos a estudiar la morfología de la población de vacíos estadísticamente desde la distribución de elipticidad y prolacidad. Respecto al tamaño de los vacíos, estudiamos la función de abundancia de vacíos dependiente del radio efectivo y la función de correlación de dos puntos con los centros de los vacíos. Teniendo además la orientación de los semiejes se puede hacer un chequeo adicional de la orientación isotrópica de los vacíos.

3.1.4. Orientación isotrópica

El principio cosmológico establece que a escalas lo suficientemente grandes, todos los observadores detectarán las mismas propiedades del Universo. Como consecuencia, se tendrá homogeneidad e isotropía en la distribución de materia, sin direcciones preferenciales. Esta propiedad se extiende a los vacíos cosmológicos, así que se espera encontrarla en los catálogos de vacíos arrojados por el algoritmo.

Para esto se elige uno de los tres semiejes y una dirección con la cual se realiza el producto punto.

$$\cos \theta_x = \hat{a} \cdot \hat{i} \quad (3-9)$$

3.1.5. Función de tamaño de Vacíos

Uno de las principales medidas estadísticas en cosmología son las funciones de abundancia. Estas dan cuenta de la estructura jerárquica donde los pequeños componentes que son los más abundantes se fusionan para dar lugar a componentes más grandes y menos abundantes, como en la función de luminosidad de galaxias y la función de tamaño de halos de materia oscura. Para vacíos cosmológicos se realiza el mismo estudio de abundancia, que puede incluso ser utilizado como prueba cosmológica [13].

La función de abundancia puede parametrizarse para estudiar múltiples escenarios. Se puede describir como una función del radio efectivo del vacío R con dos potencias:

$$f(R; n_1, n_2, R_0, N_0) = \frac{N_0}{\left(\frac{R}{R_0}\right)^{n_1} + \left(\frac{R}{R_0}\right)^{-n_2}}. \quad (3-10)$$

con R_0 como el radio de transición entre el régimen de vacíos grandes dominado por la potencia n_1 y el régimen de vacíos pequeños donde domina la potencia n_2 , y un factor de normalización N_0 .

3.1.6. Función de Correlación de Dos Puntos

Dentro de un conjunto de puntos distribuidos en un espacio tridimensional, Según Peebles [36], la probabilidad de encontrar en el volumen δV_1 un punto y simultáneamente encontrar otro en el volumen δV_2 a una distancia r viene dada por:

$$dP = n^2 [1 + \xi(r)] \delta V_1 \delta V_2 \quad (3-11)$$

con n la densidad volumétrica de datos y $\xi(r)$ es la función de correlación de dos puntos (**TPCF**). Esta función puede entenderse también como el exceso de probabilidad de encontrar puntos respecto a un campo aleatorio de puntos, funciona como indicador de información subyacente en estructuras, como por ejemplo las oscilaciones bariónicas acústicas.

El estimador de Landy and Szalay [29] introduce correlación cruzada entre el campo de puntos aleatorio y el campo de datos para reducir la varianza en la función de correlación.

$$\xi(r) = \frac{N_r(N_r - 1)}{N_d(N_d - 1)} \frac{DD(r)}{RR(r)} - \frac{2(N_r - 1)}{N} \frac{DR(r)}{RR(r)} + 1 \quad (3-12)$$

3.2. Efectos de la resolución del catálogo de entrada

3.2.1. Catálogo de Entrada

Como catálogo de entrada elegimos un catálogo de halos de materia oscura de la suite de simulaciones cosmológicas de N-cuerpos Abacus Cosmos Project [23]. Este proyecto cuenta con simulaciones de gran escala en distintas resoluciones y en 40 diferentes escenarios cosmológicos, los resultados son de libre acceso y tienen tanto información de los halos de materia oscura como de las partículas que los conforman.

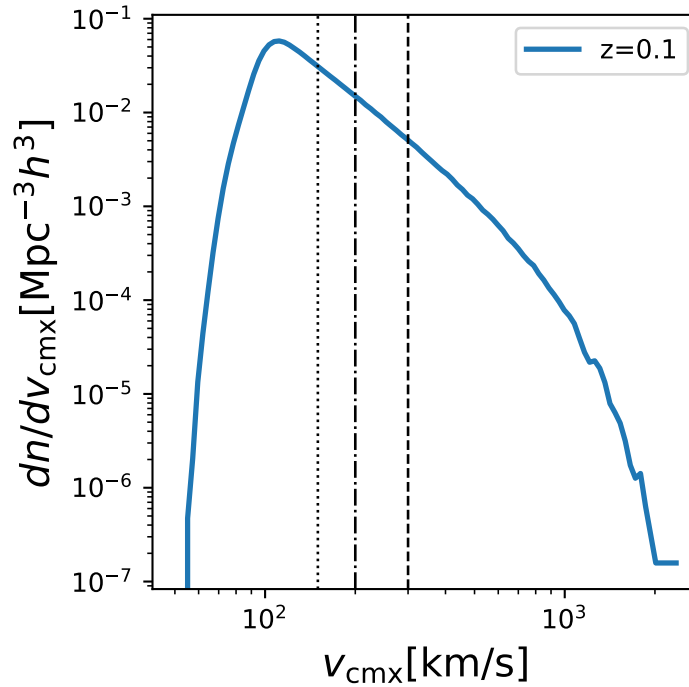


Figura 3-1.: Función de densidad de halos para el catálogo de halos extraído de Abacus Cosmos. Se extraen tres subconjuntos definidos por la velocidad circular máxima $v_{cmx} = 150, 200$ y 300 km/s.

Elegimos el catálogo de halos de materia oscura tipo Friend Of Friends (**FOF**) de la simulación dentro de la caja cúbica de $720h^{-1}\text{Mpc}$, con 1440^3 partículas de $1 \times 10^{10}h^{-1}M_{\odot}$. Elegimos la captura de la simulación a redshift $z = 0,1$, el más tardío disponible, también equivalente al redshift típico de los surveys actuales. Este catálogo original contiene cerca de $\sim 6 \times 10^6$ halos.

En los catálogos FOF no aparece directamente la masa contenida en los halos de materia oscura como si sucede en los catálogos encontrados con ROCKSTAR, sin embargo el criterio de velocidad circular máxima $v_{\text{circ max}} = \sqrt{GM(<r)/R}$ ([28]) cumple con esta función y

resulta particularmente útil cuando se está trabajando con halos incrustados dentro de halos más grandes.

Del catálogo original seleccionamos los halos haciendo cortes de velocidad circular máxima a $v_{cmx} < 150, 200$ y 300 km/s. Esto resultó en catálogos con un número de halos $N_d \approx 1,9 \times 10^6, 8,7 \times 10^5$ y $2,7 \times 10^5$ respectivamente.

La simulación original trabaja con condiciones de frontera periódicas. Sin embargo, el algoritmo que proponemos en el presenta trabajo no las contempla. Es posible que se identifiquen vacíos altamente irregulares debido al truncamiento en las fronteras del catálogo de entrada.

3.2.2. Función de abundancia de vacíos

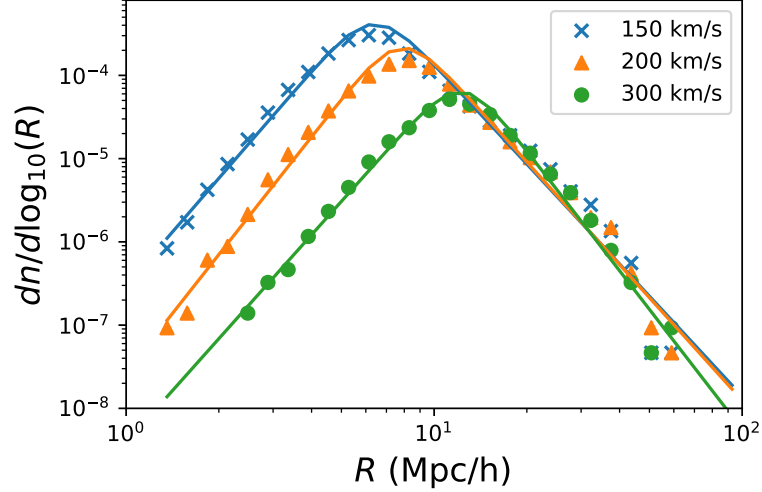


Figura 3-2.: Función de abundancia de tamaño de vacíos bajo distinta resolución del catálogo de entrada.

El β -Voids fue alimentado con uno de los tres catálogos de entrada definidos por $v_{\text{cmx}} > 150, 200$ y 300 km/s con parámetros fijos $r = 1,0$ y $\beta = 1,0$. En la figura **3-2** se muestran las funciones de abundancia de tamaño de los vacíos. Las tres funciones de abundancia presentan comportamientos similares que puede ser descrito por la función de ajuste dada por la ecuación 3-10 con diferentes parámetros. El parámetro de ajuste para la altura del pico de abundancia N_0 crece con la resolución del catálogo de entrada, el parámetro R_0 que indica el radio efectivo de transición disminuye cuando aumenta la resolución del catálogo de entrada. Esto es congruente con el hecho que al aumentar la cantidad de puntos en el catálogo de entrada, se da lugar a la aparición de más vacíos pequeños que en un catálogo de baja resolución.

En el régimen de grandes vacíos, $R_{\text{eff}} > R_0$ las tres funciones de abundancia se superponen, en los tres casos se está detectando la misma cantidad de vacíos grandes, como es de esperarse, los trazadores de la LSS (galaxias/halos) de baja masa son escasos en los vacíos más grandes. La pendiente n_1 para este régimen es similar en los tres casos.

En el régimen de vacíos pequeños, $R_{\text{eff}} < R_0$ se observan pendientes similares para las tres funciones de abundancia. Si bien una estructura de gran escala con mayor resolución da lugar a la aparición de vacíos más pequeños, se debe tener en cuenta que el parámetro fijo en $r = 1$ aumenta proporcionalmente la cantidad de puntos aleatorios que pueden estar llenando los espacios previamente existentes en la LSS de baja resolución. Esto se analizará con mayor detalle en la sección 3.3

Como se mencionó en la sección 3.2.1, el algoritmo actualmente no considera condiciones de frontera periódicas. Esto puede llevar a la identificación de vacíos truncados por la pared

de la caja que contiene el catálogo de datos. Para reducir este efecto se decidió descartar los halos cuyo centro estuviera a una distancia menor a $30 h^{-1}\text{Mpc}$ de las paredes de la caja. Esto reduce el volumen efectivo de la simulación, pasando de una caja de lado $(720 h^{-1}\text{Mpc})^3$ a una caja de lado $(660 h^{-1}\text{Mpc})^3$.

3.2.3. Morfología de los vacíos

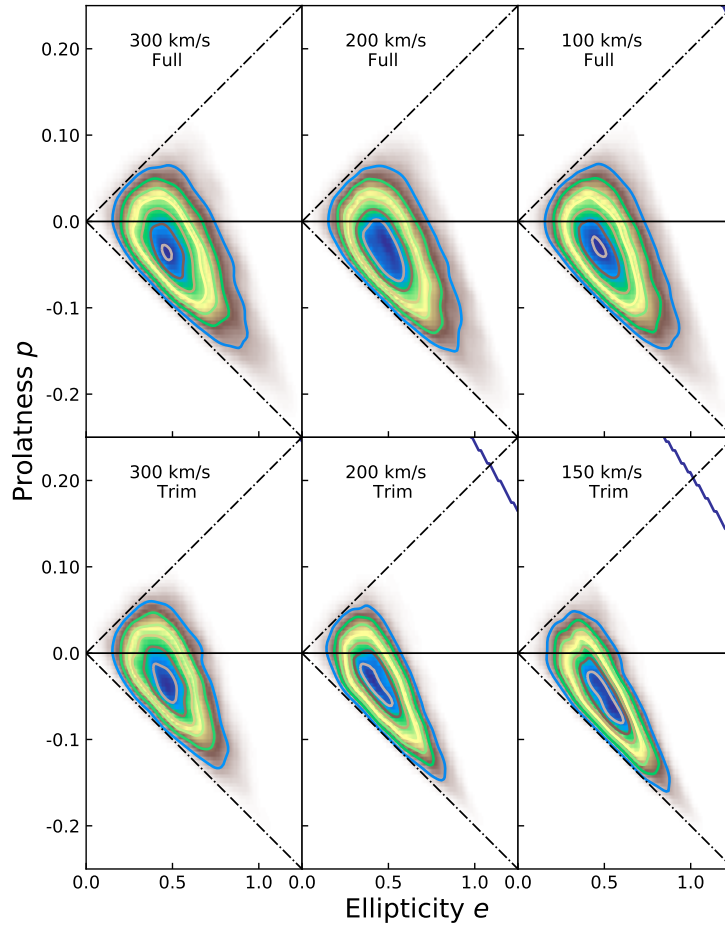


Figura 3-3.: Diagramas de dispersión de los vacíos para las tres diferentes resoluciones de entrada, fila superior, junto con los diagramas para la selección de vacíos descartando todos los vacíos formándose cerca de las paredes de la caja, cuyos centros estén a menos de $30 h^{-1}\text{Mpc}$, fila inferior.

En la figura 3-3 tenemos los diagramas de dispersión Prolatidad-Elipticidad. Al recortar del catálogo de vacíos todos aquellos cercanos a las paredes de la caja se reduce la varianza de la elipticidad y la prolatidad, tabla 3.2.3. Estos son consistentes con los valores $e = 0,10^{+0,05}_{-0,02}$ y $p = -0,05^{+0,03}_{-0,04}$ reportados por Adermann et al. [3] utilizando un buscador de

| Catálogo v_{cmx} (km/s) | Completo | | Reducido | |
|-------------------------------------|------------------------|-------------------------|------------------------|-------------------------|
| | e | p | e | p |
| 300 | $0,10^{+0,03}_{-0,02}$ | $-0,03^{+0,03}_{-0,03}$ | $0,10^{+0,02}_{-0,02}$ | $-0,03^{+0,03}_{-0,03}$ |
| 200 | $0,10^{+0,03}_{-0,02}$ | $-0,03^{+0,03}_{-0,04}$ | $0,09^{+0,02}_{-0,02}$ | $-0,03^{+0,03}_{-0,03}$ |
| 150 | $0,10^{+0,03}_{-0,02}$ | $-0,03^{+0,03}_{-0,04}$ | $0,10^{+0,02}_{-0,02}$ | $-0,04^{+0,03}_{-0,03}$ |

Tabla 3-1.: Prolacidad y elipticidad de los vacíos encontrados bajo variación de la resolución del catálogo de entrada. Catálogo completo y Catálogo reducido descartando vacíos cercanos a las paredes de la caja. Se presentan el valor medio de cada distribución junto con los límites superior e inferior para percentiles 75 y 25.

vacíos directamente sobre datos crudos de la posición de las partículas de materia oscura y gas en una simulación de baja resolución en una caja cúbica de $500 \text{ h}^{-1}\text{Mpc}$ mediante la discretización del espacio, y los resultados de Nadathur [32] utilizando el buscador ZOBOV.

3.3. Efectos de la variación de los parámetros β y r

El algoritmo de β -Voids tiene dos parámetros libres: β en el calculador del β -Skeleton y r la proporción entre el tamaño del catálogo de números aleatorios RC y el número de objetos en el catálogo de entrada DC. Para estudiar la influencia de estos parámetros en el catálogo de vacíos encontrados, decidimos trabajar con un único catálogo de entrada con velocidad circular máxima $v_{\text{cmx}} \geq 300$ km/s de aquí en adelante. Elegimos variar los parámetros en una grilla de seis valores diferentes para $\beta = 1,0, 1,2, 1,4, 1,6, 1,8, 2,0$ y cuatro diferentes para $r = 1,0, 1,2, 1,4, 1,6$, en total de 24 puntos de muestreo.

3.3.1. Vacío más grande

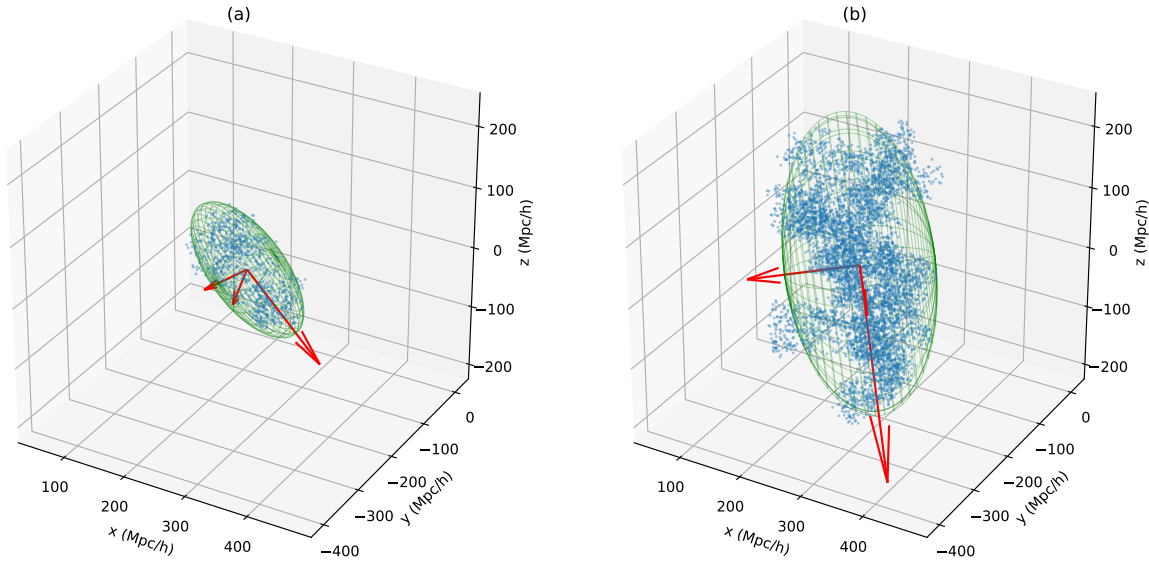


Figura 3-4.: Influencia de la variación de los parámetros β y r del algoritmo, calculando sobre el mismo catálogo de entrada $v_{\text{cmx}} \geq 300$ km/s. A la izquierda se muestran el vacío más grande encontrado con parámetros $(\beta, r) = (1,0, 1,0)$, a la derecha el vacío más grande encontrado con $(1,8, 1,6)$. Ambos gráficos están centrados en el mismo punto y comparten la escala. El elipsoide asociado a cada vacío se representa como la malla en verde. Los semiejes se resaltan en rojo, escalados por un factor de 1.5 para facilitar su visualización.

El indicador más claro del efecto de los parámetros en los vacíos hallados es el tamaño del vacío más grande detectado. En la figura 3-4 se tienen los vacíos más grandes encontrados con los parámetros $p_1 = (1,0, 1,0)$ y $p_2 = (1,8, 1,6)$. El radio efectivo para los vacíos es $R_{\text{eff}} = 75,8$ y $151,4 \text{ h}^{-1}\text{Mpc}$ respectivamente. Se puede ver que el vacío p_1 es una estructura

de puntos compacta, bien descrita por el elipsoide asociado, este tipo de vacíos es un vacío elipsoidal regular. En tanto que el vacío más grande para p_2 tiene espacios donde pasa la estructura de gran escala, alejando, su forma es bastante irregular, aunque de todos modos es posible calcular su pseudo-tensor de inercia y un elipsoide asociado. Esto sugiere que el calcular una tasa de ocupación o una densidad de puntos dentro del elipsoide es una medida de qué tan regular o irregular es un vacío. Cabe notar que se puede entender este vacío irregular como un conjunto de vacíos pequeños que a r bajo se encuentran desconectados, pero al aumentar r empiezan a conectarse por medio de un efecto percolación [7] dada la naturaleza aleatoria de la posición de los puntos aleatorios que definen los vacíos en este algoritmo.

El radio efectivo del vacío más grande encontrado con cada conjunto de parámetros se muestra en la figura 3-5. En un corte lateral de la superficie generada, se puede observar que el rango de variación más grande del radio efectivo corresponde a la variación del parámetro r .

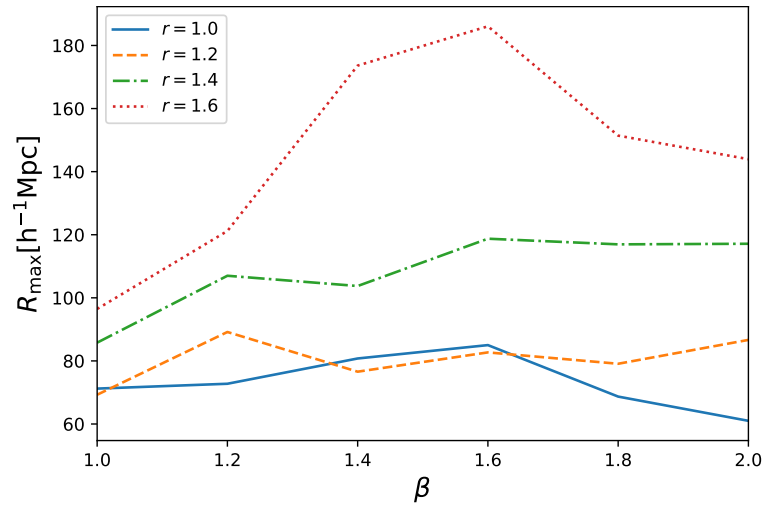


Figura 3-5.: Radio efectivo máximo encontrado en la variación de parámetros β y r . Se observa una fuerte dependencia de r

3.3.2. Función de tamaño de vacíos

En la sección 3.2.2 estudiamos cómo la resolución del catálogo de entrada cambia la función de tamaño de vacíos y sugiere además la revisión de la dependencia de la función de tamaño con los parámetros del algoritmo. En la figura A-1 se muestran las distintas funciones de tamaño de vacíos con los diferentes parámetros cosmológicos ajustadas con la función descrita en la ecuación 3-10 que simplifica el análisis. Este ajuste se realizó en el espacio log-log para darle igualdad de peso a todos los valores de R_{eff} y a todo el dominio de la función de tamaño de vacíos.

En la figura 3-6 se puede observar que dos de los parámetros de ajuste (R_0 y n_1) tienen correlación con β y r , en tanto N_0 y n_2 se no presentan dicha correlación. Para N_0 este comportamiento representa que el pico de transición entre el régimen dominado por n_1 hacia el régimen dominado por n_2 tiene una cantidad similar de vacíos, sin grandes variaciones. Para n_2 tampoco se tiene una correlación clara con β y r . Los parámetros de ajuste R_0 y n_1 si presentan una relación con clara con β y r , ambas con exponente negativo (no necesariamente $\propto 1/\beta$ o $\propto 1/r$). La reducción de R_0 indica que se están detectando estructuras más pequeñas, que equivale a un aumento de la resolución del buscador. La reducción de n_1 indica directamente un aumento en la pendiente en la gráfica log-log de la función de tamaño de vacíos para el régimen de grandes vacíos $R > R_0$, se detecta una población mayor de vacíos grandes, por lo tanto el radio efectivo del vacío más grande encontrado crece con valores altos de r , figura 3-5.

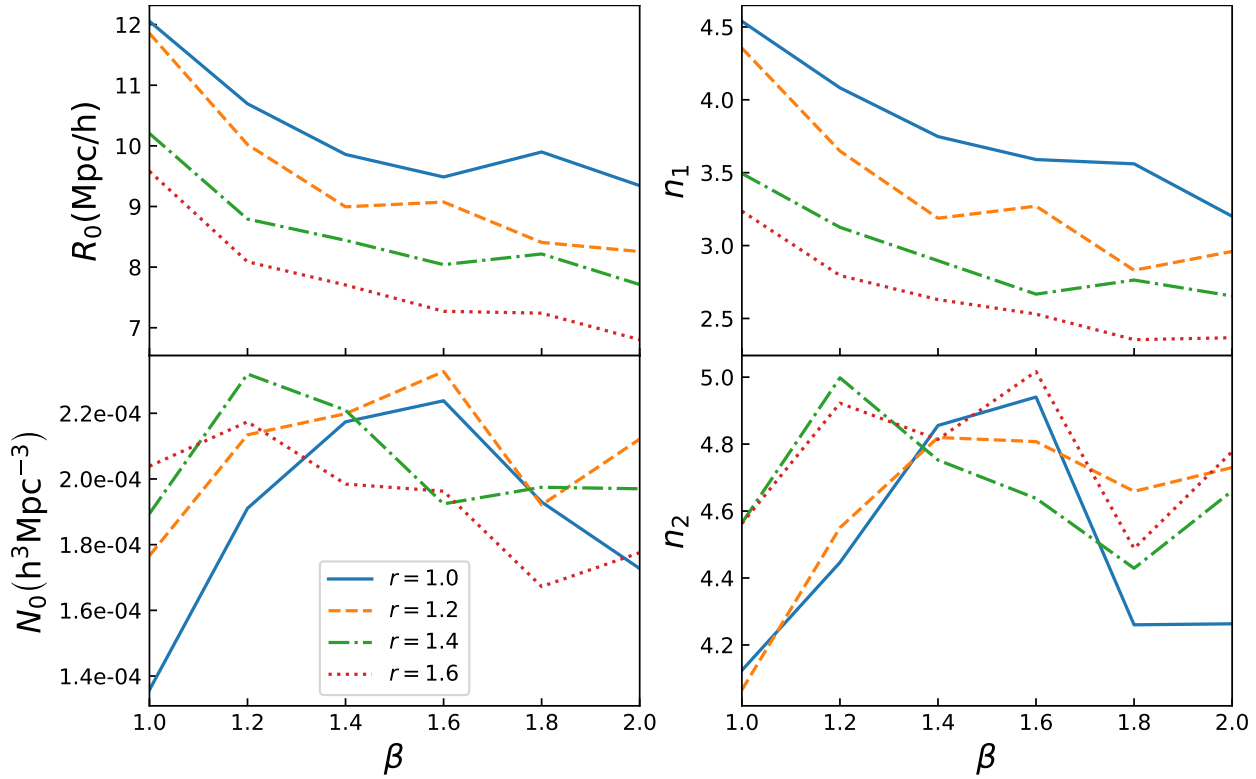


Figura 3-6.: Parámetros de ajuste para la función de tamaño de vacíos explorando sobre la grilla β - r

3.3.3. Fracción de ocupación de los vacíos

Según la estimación de Cautun et al. [10] utilizando el NEXUS+ los vacíos ocupan un 77% del volumen del universo, conteniendo tan solo un 15% de la materia total del mismo.

en la figura 3-7 Al evaluar la fracción del volumen ocupada por los vacíos encontrados se tiene que r domina sobre β la fracción de ocupación. El rango operacional queda como $1,0 \leq r \leq 1,4$ y $1,0 \leq \beta \leq 2,0$.

El algoritmo β -Voids trabaja llenando el espacio con puntos aleatorios y conectándolos mediante el β -Skeleton. Con un r bajo se describen los vacíos con un conjunto de puntos aleatorios pequeño, dejando espacio entre el conjunto de puntos definido como vacío y los puntos de la LSS. Al aumentar la cantidad de puntos aleatorios con los que se trabaja, más puntos estarán poblando los vacíos, llenando dichos intersticios entre vacíos definidos con bajo r y la estructura de gran escala. Para valores de $r \geq 1,6$ empiezan a caer suficientes puntos aleatorios dentro de los filamentos y las hojas de la LSS, suficientes para que el grafo β -Skeleton conecte vacíos adyacentes y eventualmente, con valores de r lo suficientemente grandes, empiezan a conectarse más y más vacíos desbordando en un gran vacío que ocupa todo el volumen de datos y tiene islas que vienen a ser los cúmulos de galaxias/halos.

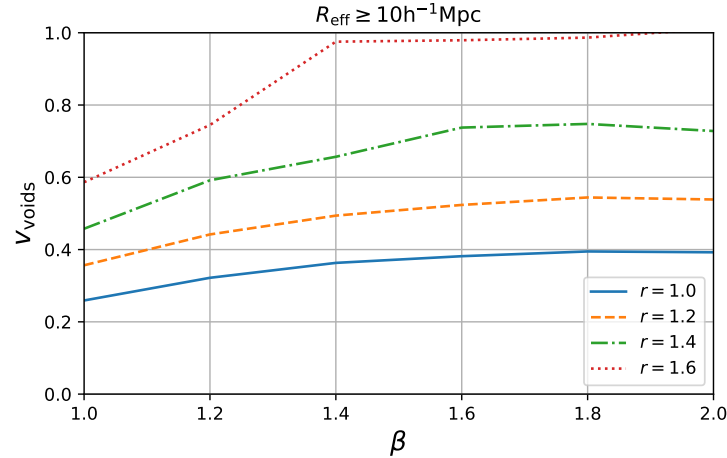


Figura 3-7.: Fracción de ocupación en función de los parámetros β - r del algoritmo β -Voids, para vacíos con $R_{\text{eff}} \geq 10 \text{ h}^{-1} \text{Mpc}$. La selección de un radio efectivo menor incrementa menos de un 10 % los valores actuales.

3.3.4. Morfología de los vacíos

Tomamos como referencia la distribución de Prolacidad-Elipticidad de la figura 3-3, calculada con $(\beta, r) = (1,0, 1,0)$. Al calcular de nuevo variando β y r se observa muy poca influencia de r la prolacidad y la elipticidad, como se había observado en el tamaño de los voids. La influencia de β es mayor.

Sin embargo, los vacíos más grandes que se forman con cada par de parámetros (β, r) , empiezan a formarse como uniones de vacíos contiguos, como se puede ver en la figura 3-4.

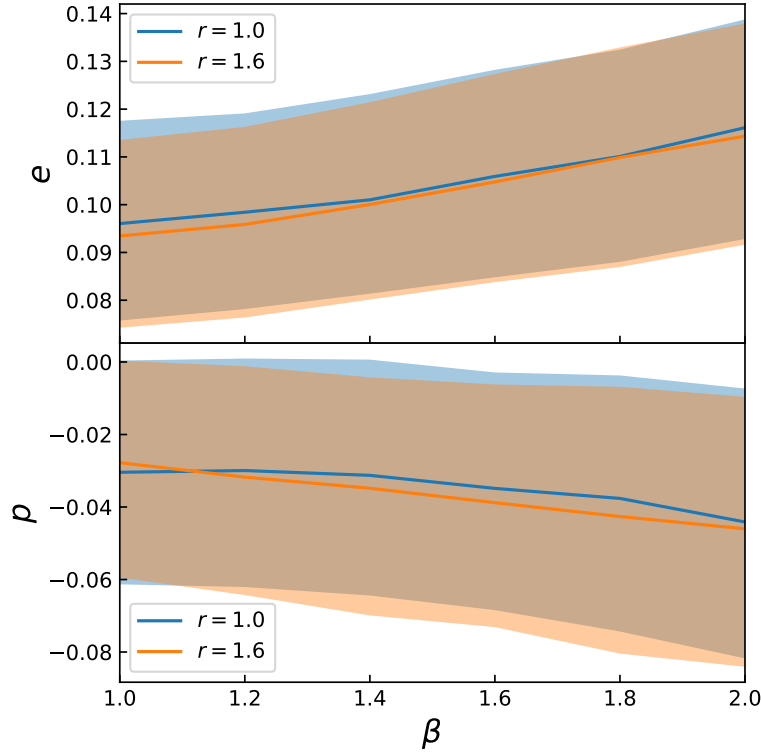


Figura 3-8.: Dependencia de β y r en la forma de los vacíos. Para r es prácticamente constante. Para cambios en β hay una pequeña variación que en p se queda dentro de la franja de los percentiles 25-75, y en e llega al percentil 75.

Definimos la densidad de datos (galaxias/halos) Δ entre el volumen total del catálogo de datos. Recordando que la cantidad de puntos aleatorios en el volumen total viene dada por:

$$N_r = r N_d \quad (3-13)$$

la densidad de puntos aleatorios es igual a la densidad de datos por la proporción r

$$\Delta_r = \frac{N_r}{V} = r \frac{N_r}{N_d} \frac{N_d}{V} = r \Delta \quad (3-14)$$

$$\frac{\Delta_r}{r} = \Delta \quad (3-15)$$

sin importar r .

Se define la densidad normalizada de partículas para un vacío como el número de puntos aleatorios que lo conforman dividido entre el producto de la proporción r por el volumen del elipsoide asociado al vacío.

$$\delta_i = \frac{N_i}{r V_i} \quad (3-16)$$

En vacíos pequeños el elipsoide asociado encierra los puntos aleatorios que definen el vacío de un modo ajustado, con puntos muy cerca o sobre la superficie del elipsoide, a menos de la mitad de la distancia media entre puntos aleatorios. Esto se traduce en una densidad normalizada mayor a la densidad Δ . En los vacíos más grandes, formados como fusión de dos o más vacíos contiguos, el elipsoide asociado resulta atravesado por la estructura de gran escala, dejando espacio libre entre el conjunto de partículas de vacío y el elipsoide asociado, resultando en una densidad normalizada inferior Δ , como en el panel derecho de la figura 3-4, este vacío tiene $\delta < 0,5\Delta$.

Se pueden definir los vacíos regulares como aquellos con $\delta \geq \Delta$, generalmente vacíos pequeños. Y vacíos irregulares como aquellos con $\delta < 1$.

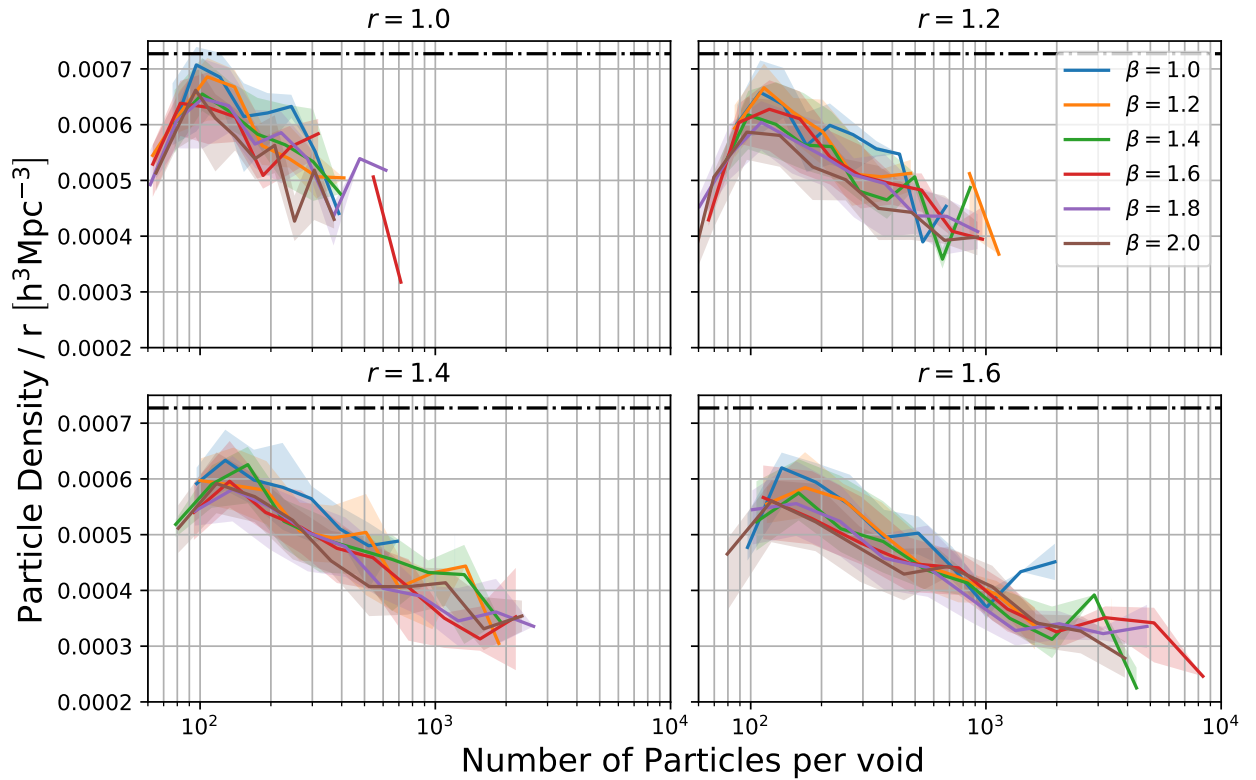


Figura 3-9.: Densidad de partículas aleatorias de vacío normalizada por r para vacíos grandes ($R_{\text{eff}} \geq 30h^{-1}\text{Mpc}$). La línea horizontal representa la densidad de puntos de datos en el volumen total del catálogo de entrada (Δ), se muestra como referencia.

3.3.5. Orientación isotrópica

Revisamos uno por uno los semiejes de los elipsoides asociados a los vacíos. Calculamos el ángulo que forma con una dirección determinada, por simplicidad los ejes x, y, z del sistema

de coordenadas. Calculamos la distribución de probabilidad. Luego realizamos una prueba de Kolmogorov-Smirnov comparando con la hipótesis nula de una distribución uniforme. Eligiendo el valor de significancia $\alpha = 0,01$ se encontró que en las tres direcciones el p-value es mayor que α , salvo el caso atípico de el semieje a con el eje x con un $P = 0,005$. En general, aceptamos la hipótesis nula: la orientación de los vacíos encontrados tienen una distribución aleatoria uniforme.

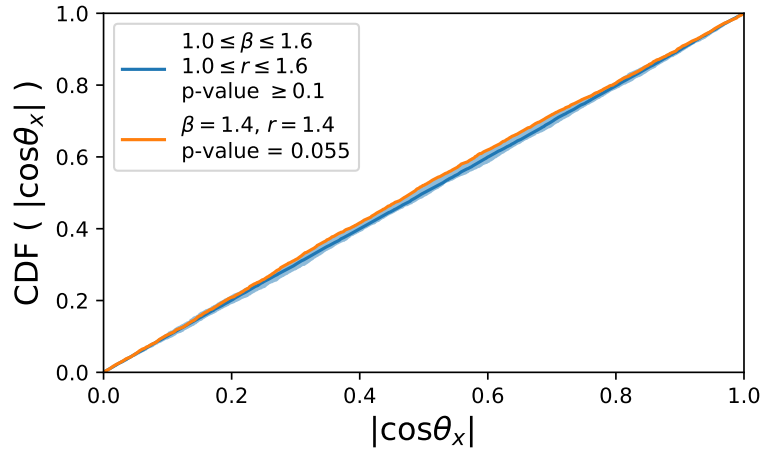


Figura 3-10.: Función de probabilidad acumulada para un componente del semieje mayor con el eje x del sistema de coordenadas para los vacíos encontrados en todo el rango de parámetros del algoritmo explorados.

3.3.6. Función de correlación de dos puntos

Calculamos la función de correlación de dos puntos entre los centros de los vacíos encontrados para todos los catálogos en los parámetros β, r explorados. Encontramos una anticorrelación para el tamaño de los vacíos para distancias menores a $16h^{-1}\text{Mpc}$.

Si el radio efectivo medio de los vacíos encontrado en las funciones de abundancia es cercano a los $8h^{-1}\text{Mpc}$, entonces la distancia típica entre dos vacíos será $16h^{-1}\text{Mpc}$ (donde se hace cero la función de correlación), la probabilidad de encontrar dos vacíos a menos de dicha distancia será baja.

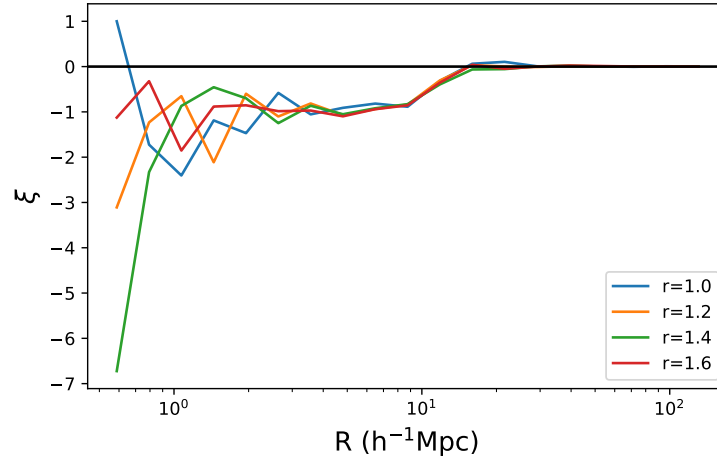


Figura 3-11.: Función de correlación para todos los vacíos del catálogo. Se puede ver un régimen de anticorrelación para distancias menores a $16h^{-1}\text{Mpc}$. Se calculó para $\beta = 1,0$.

4. Conclusiones

En este trabajo presentamos un nuevo algoritmo geométrico para identificar vacíos en la Estructura de Gran Escala de Galaxias a partir de catálogos tridimensionales de trazadores de la LSS, bien sean galaxias de observaciones o halos de materia oscura de simulaciones. Este algoritmo está basado en el grafo β -Skeleton, tiene dos parámetros libres, encuentra los vacíos de la LSS poblándolos con puntos ubicados aleatoriamente conectados por medio del β -Skeleton.

Realizamos pruebas del algoritmo β -Voids sobre un catálogo FOF de halos de materia oscura de la simulación Abacus Cosmos, una caja cúbica de $720 \text{ h}^{-1}\text{Mpc}$, haciendo selección de tamaño de los halos por medio de la velocidad circular máxima con cortes en $v_{\text{cmx}} \geq 150, 200, 300 \text{ km/s}$. El análisis estadístico de los vacíos encontrados lo realizamos desde la aproximación estándar de la morfología de los vacíos a elipsoides triaxiales, permitiendo calcular fácilmente los semiejes, orientación y volumen de los vacíos individualmente.

Primero estudiamos la influencia de la resolución del catálogo de entrada. Esta sigue una estructura jerárquica donde los vacíos más grandes son los menos abundantes. Encontramos un máximo de abundancia para vacíos con un radio efectivo $R_0 \sim 6 - 10 \text{ h}^{-1}\text{Mpc}$. Por debajo de este valor la cantidad de vacíos encontrados disminuye, como establece el modelo SvdW con void in cloud [44]. Sin embargo, la cantidad de vacíos pequeños aumenta con la resolución del catálogo de entrada; esto obedece principalmente a que un catálogo con mayor resolución tiene una topología más rica en estructura fina. Este radio de transición R_0 disminuye a medida que aumenta la cantidad de puntos del catálogo de entrada. Para el régimen de vacíos grandes hay concordancia en la cantidad de vacíos encontrados con las diferentes resoluciones de entrada. El análisis de elipticidad y prolacidad en el catálogo completo de vacíos encontrados y el catálogo omitiendo vacíos cercanos a los bordes, arroja valores $e = 0,10^{+0,03}_{-0,02}$ y $p = -0,03^{+0,03}_{-0,03}$, con una reducción de $1/3$ en la varianza de la elipticidad para el catálogo omitiendo vacíos de borde, en general los vacíos que este método encuentra tienen forma prolata, con una pequeña fracción en la cola, de vacíos oblatos. Los valores encontrados de e y p son cercanos a los reportados por Adermann et al. [3].

La segunda parte del análisis la concentramos en los efectos de la variación de los parámetros libres del algoritmo β -Voids. Se trabajó sobre una grilla de 24 parejas de parámetros con $\beta = 1,0, 1,2, 1,4, 1,6, 1,8, 2,0$, que definen la zona de exclusión con la que se construye el β -Skeleton, y $r = 1,0, 1,4, 1,6, 1,8$, que define la proporción entre número de puntos aleatorios sobre el número de puntos de datos.

Encontramos que la función de tamaño de vacíos es altamente sensible a variaciones

en r . Con valores de $r > 1,6$ la cantidad de puntos aleatorios puede ser suficiente para presentar un fenómeno de percolación, conectando vacíos adyacentes por medio de puntos aleatorios, propios del algoritmo, dentro de filamentos u hojas de la LSS. Eventualmente, con una proporción de puntos aleatorios lo suficientemente grande, todo el volumen del catálogo puede llegar a ser detectado como un gran vacío con los nodos de la LSS como islas embebidas dentro del gran vacío.

Encontramos que la morfología de los vacíos es sensible a variaciones en β aumentando la varianza de la proclividad, principalmente en el régimen de vacíos pequeños, $R_{\text{eff}} < 10 \text{ h}^{-1}\text{Mpc}$, cuando $\beta > 1,6$. En los vacíos más grandes el efecto es menor. Al reducir la cantidad de conexiones los vacíos definidos por un conjunto pequeño de puntos aleatorios terminan siendo más irregulares.

Este algoritmo provee una forma adicional para cuantificar la irregularidad de los vacíos, particularmente útil en los vacíos más grandes $R_{\text{eff}} < 30 \text{ h}^{-1}\text{Mpc}$. Definimos una densidad normalizada de partículas para cada vacío, tomando el número de partículas de cada vacío sobre el producto de r por el volumen del elipsoide asociado. En vacíos cercanos al elipsoide asociado, la densidad de partículas es similar a la densidad de puntos en el catálogo de datos. Para vacíos irregulares tenemos espacios por donde pasa la red cósmica, el volumen ocupado por las partículas que lo conforman es menor al volumen del vacío asociado y por ende, tienen una densidad de partículas inferior a la densidad media de puntos del catálogo de datos.

El criterio principal para acotar el rango de operación de los parámetros del algoritmo es la fracción de ocupación de vacíos. Según [10] los vacíos ocupan un 77 % del volumen total del universo. Con los parámetros dentro de los rangos $1,0 \leq \beta \leq 2,0$ y $1,0 \leq r \leq 1,4$ calculamos fracciones de ocupación entre el rango 30 – 80 % con poblaciones de vacíos dadas por $R_{\text{eff}} < 10 \text{ h}^{-1}\text{Mpc}$, teniendo en cuenta que no existe una definición estándar de vacío entregamos estos como los parámetros de operación para el método.

Finalmente se ejecutamos dos pruebas robustas sobre el catálogo de vacíos encontrado por el algoritmo β -Voids. Del principio cosmológico se espera que no haya una orientación preferencial para los vacíos. Tomamos esto como la hipótesis nula, mediante una prueba Kolmogorov-Smirnov con un nivel de significancia $\alpha = 0,01$ encontramos p-values sobre α para la función de distribución de probabilidad de orientación de los componentes de los semi-ejes de los elipsoides asociados respecto a una dirección fija. La segunda prueba aplicada fue la función de correlación de dos puntos. Como trabajo futuro podemos aplicar este buscador sobre diferentes escenarios cosmológicos del proyecto Abacus-Cosmos y sobre datos observacionales utilizando el mismo análisis estadístico desarrollado hasta este punto: función de tamaño de vacíos, análisis morfológico con elipticidad y proclividad, isotropía de la orientación de los vacíos y función de correlación.

A. Anexo: Gráficas de la variación de parámetros del algoritmo

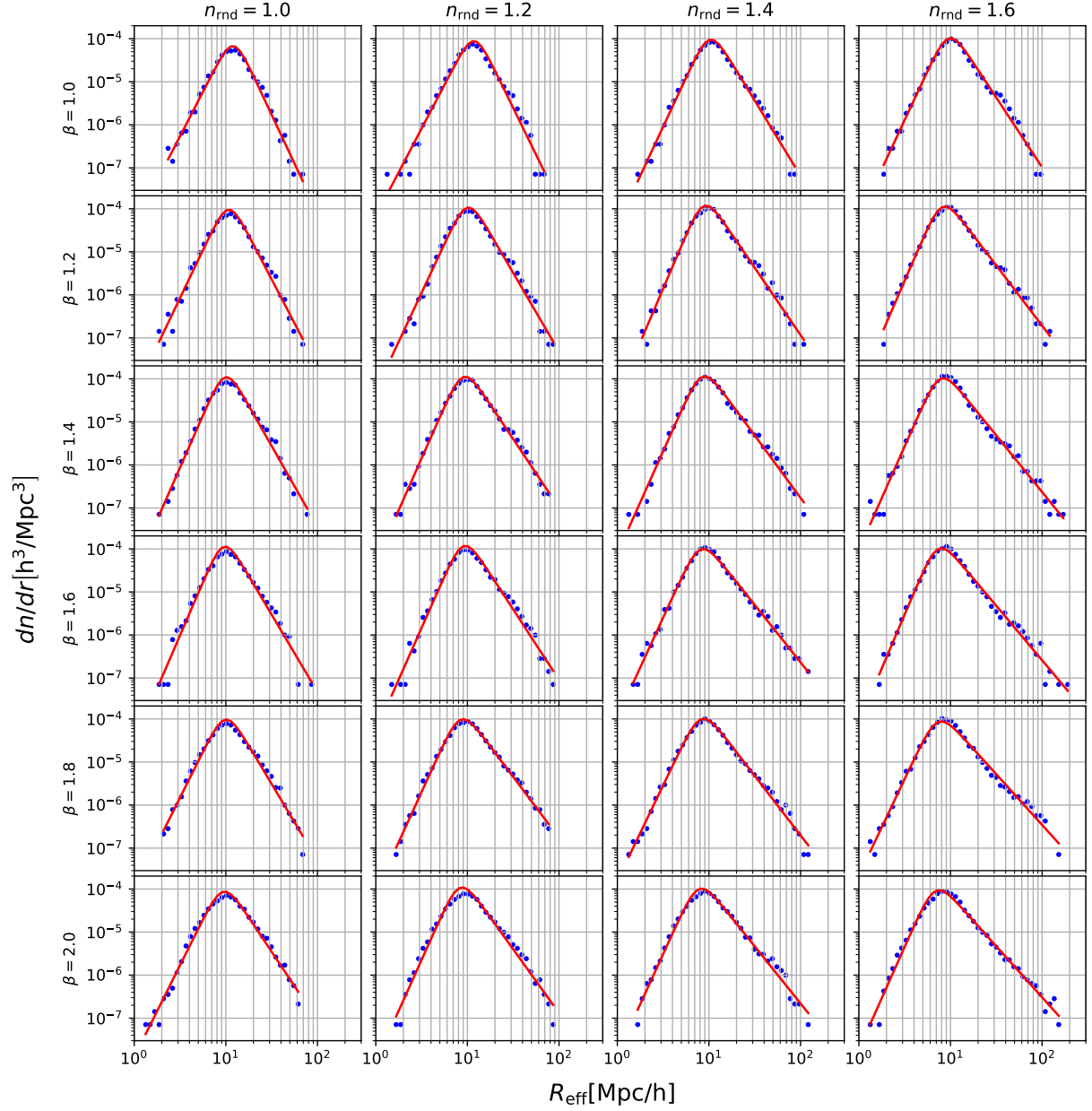


Figura A-1.: Efecto de la variaci3n de par3metros β y r sobre las funciones de tama1o de vac3os en azul, junto sus respectivas funciones de ajuste en rojo, seg3n la ecuaci3n 3-10

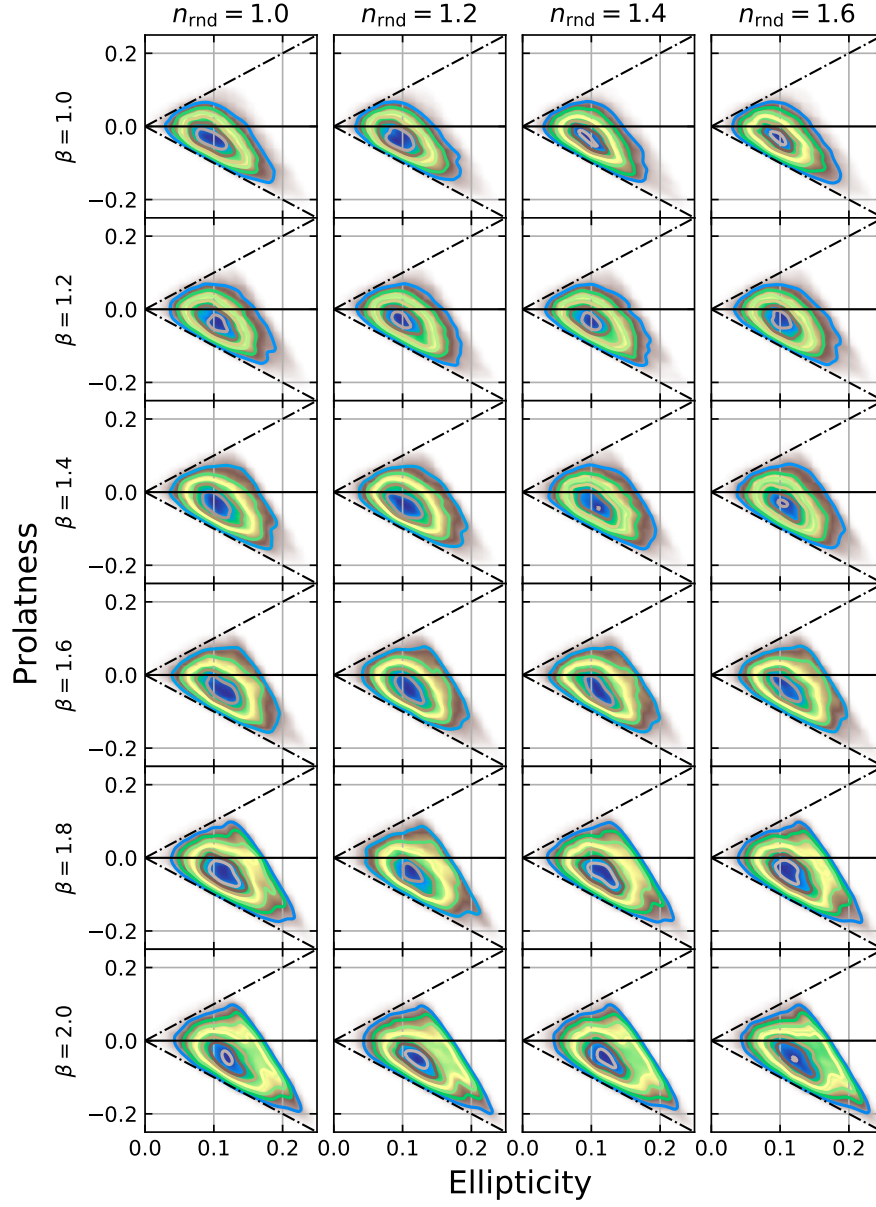


Figura A-2.: Efecto de la variación de parámetros del buscador sobre la morfología de vacíos con $R_{\text{eff}} \geq 1 \text{ h}^{-1}\text{Mpc}$.

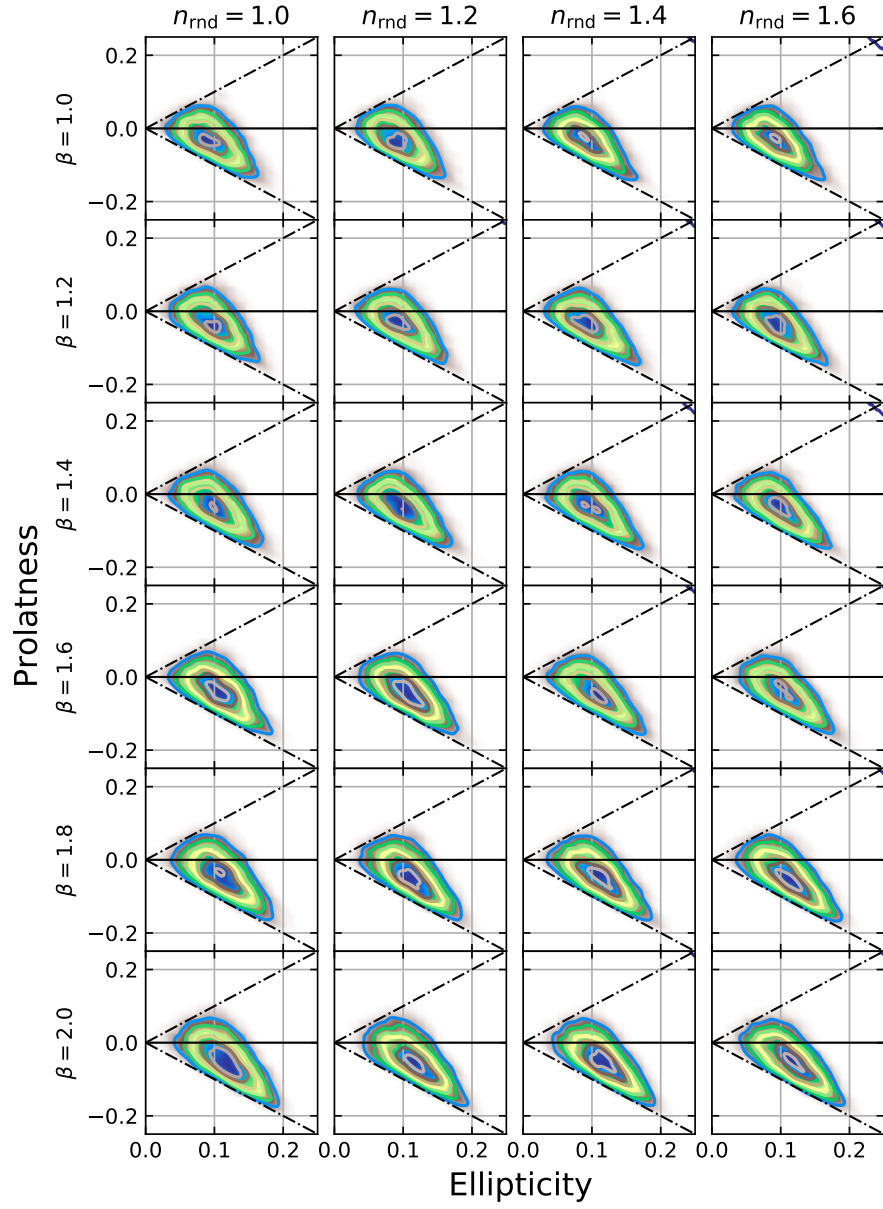


Figura A-3.: Efecto de la variación de parámetros del buscador sobre la morfología de vacíos con $R_{\text{eff}} \geq 10 \text{ h}^{-1}\text{Mpc}$.

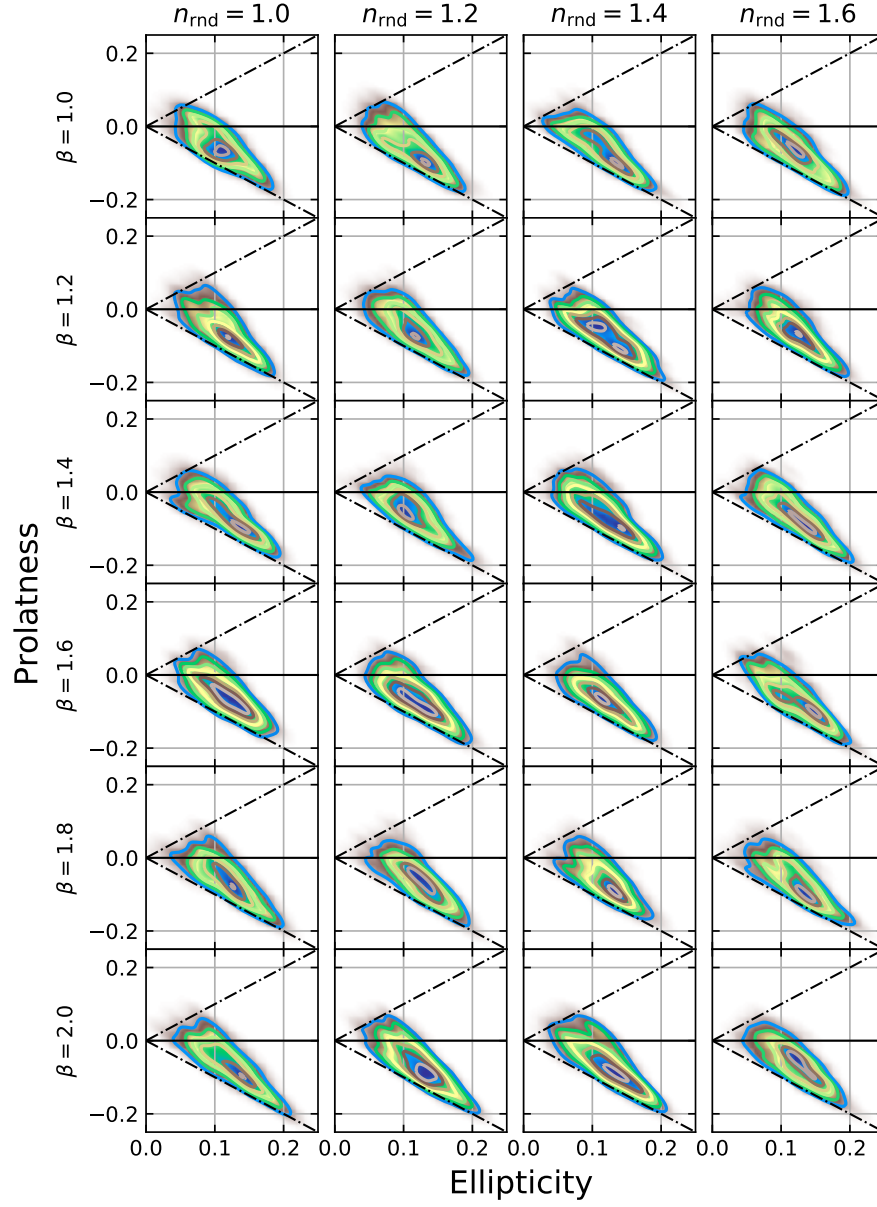


Figura A-4.: Efecto de la variación de parámetros del buscador sobre la morfología de vacíos con $R_{\text{eff}} \geq 30 \text{ h}^{-1}\text{Mpc}$.

Bibliografía

- [1] S. J. Aarseth. Numerical Experiments on the N-Body Problem (Papers appear in the Proceedings of IAU Colloquium No. 10 Gravitational N-Body Problem (ed. by Myron Lecar), R. Reidel Publ. Co. , Dordrecht-Holland.). *Astrophysics and Space Science*, 14 (1):20–34, Nov 1971. doi: 10.1007/BF00649191.
- [2] B. Abolfathi and D. S. e. a. Aguado. The Fourteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the extended Baryon Oscillation Spectroscopic Survey and from the second phase of the Apache Point Observatory Galactic Evolution Experiment. *The Astrophysical Journal Supplement Series*, 235(2):42, Apr. 2018. ISSN 1538-4365. doi: 10.3847/1538-4365/aa9e8a. URL <http://arxiv.org/abs/1707.09322>. arXiv: 1707.09322.
- [3] E. Adermann, P. J. Elahi, G. F. Lewis, and C. Power. Cosmic voids in evolving dark sector cosmologies: the low-redshift universe. *Mon Not R Astron Soc*, 468(3):3381–3394, July 2017. ISSN 0035-8711. doi: 10.1093/mnras/stx657. URL <https://academic.oup.com/mnras/article/468/3/3381/3077183>.
- [4] R. A. Alpher, R. Herman, and G. A. Gamow. Thermonuclear Reactions in the Expanding Universe. *Physical Review*, 74(9):1198–1199, Nov 1948. doi: 10.1103/PhysRev.74.1198.2.
- [5] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The statistics of peaks of Gaussian random fields. *The Astrophysical Journal*, 304:15–61, May 1986. ISSN 0004-637X. doi: 10.1086/164143. URL <http://adsabs.harvard.edu/abs/1986ApJ...304...15B>.
- [6] E. G. P. Bos, R. van de Weygaert, K. Dolag, and V. Pettorino. The darkness that shaped the void: dark energy and cosmic voids. *Monthly Notices of the Royal Astronomical Society*, 426(1):440–461, Oct. 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21478.x. URL <http://arxiv.org/abs/1205.4238>. arXiv: 1205.4238.
- [7] S. R. Broadbent and J. M. Hammersley. Percolation processes: I. Crystals and mazes. *Mathematical Proceedings of the Cambridge Philosophical Society*, 53(3):629–641, July 1957. ISSN 1469-8064, 0305-0041. doi: 10.1017/S0305004100032680. URL <https://www.cambridge.org/core/journals/>

mathematical-proceedings-of-the-cambridge-philosophical-society/
article/percolation-processes/C00CC4943F48228F8AC8031092FE84EC.

- [8] C. Cannarozzo. A new void finder based on cosmic dynamics. Master's thesis, Scuola di Scienze, Dipartimento di Física e Astronomia, Università di Bologna, Via Zamboni, 33 - 40126 Bologna, Mar 2017.
- [9] B. W. Carroll and D. A. Ostlie. *An Introduction to Modern Astrophysics*. Addison-Wesley, San Francisco: Pearson, 2nd (international) edition, 2007.
- [10] M. Cautun, R. van de Weygaert, B. J. T. Jones, and C. S. Frenk. Evolution of the cosmic web. *Mon Not R Astron Soc*, 441(4):2923–2973, July 2014. ISSN 0035-8711. doi: 10.1093/mnras/stu768. URL <https://academic.oup.com/mnras/article/441/4/2923/1213214>.
- [11] J. M. Colberg, F. Pearce, C. Foster, E. Platen, R. Brunino, M. Neyrinck, S. Basilekos, A. Fairall, H. Feldman, S. Gottloeber, O. Hahn, F. Hoyle, V. Mueller, L. Nelson, M. Plionis, C. Porciani, S. Shandarin, M. S. Vogeley, and R. van de Weygaert. The Aspen–Amsterdam Void Finder Comparison Project. *Monthly Notices of the Royal Astronomical Society*, 387(2):933–944, June 2008. ISSN 0035-8711, 1365-2966. doi: 10.1111/j.1365-2966.2008.13307.x. URL <http://arxiv.org/abs/0803.0918>. arXiv: 0803.0918.
- [12] P. Collaboration et al. Planck 2015 results. XIII. Cosmological parameters. *Astronomy & Astrophysics*, 594:A13, Oct. 2016. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201525830. URL <http://arxiv.org/abs/1502.01589>. arXiv: 1502.01589.
- [13] S. Contarini, T. Ronconi, F. Marulli, L. Moscardini, A. Veropalumbo, and M. Baldi. Cosmological exploitation of the size function of cosmic voids identified in the distribution of biased tracers. *Monthly Notices of the Royal Astronomical Society*, 488(3): 3526–3540, Sept. 2019. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stz1989. URL <http://arxiv.org/abs/1904.01022>. arXiv: 1904.01022.
- [14] C. D. Correa and P. Lindstrom. Towards Robust Topology of Sparsely Sampled Data. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1852–1861, Dec. 2011. ISSN 1077-2626. doi: 10.1109/TVCG.2011.245. URL <http://ieeexplore.ieee.org/document/6064948/>.
- [15] W. de Sitter. Einstein's theory of gravitation and its astronomical consequences. Third paper. *Monthly Notices of the Royal Astronomical Society*, 78:3–28, Nov 1917. doi: 10.1093/mnras/78.1.3.
- [16] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. *Astrophysical Journal*, 142:414–419, Jul 1965. doi: 10.1086/148306.

- [17] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29(4):551–559, July 1983. ISSN 0018-9448. doi: 10.1109/TIT.1983.1056714. URL <http://ieeexplore.ieee.org/document/1056714/>.
- [18] F. Fang, J. Forero-Romero, G. Rossi, X.-D. Li, and L.-L. Feng. beta-Skeleton analysis of the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 485(4):5276, June 2019. doi: 10.1093/mnras/stz773. URL <https://ui.adsabs.harvard.edu/abs/2019MNRAS.485.5276F/abstract>.
- [19] J. E. Forero-Romero, Y. Hoffman, S. Gottlöber, A. Klypin, and G. Yepes. A dynamical classification of the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 396:1815, July 2009. doi: 10.1111/j.1365-2966.2009.14885.x. URL <https://ui.adsabs.harvard.edu/2009MNRAS.396.1815F/abstract>.
- [20] C. Foster and L. A. Nelson. The Size, Shape, and Orientation of Cosmological Voids in the Sloan Digital Sky Survey. *The Astrophysical Journal*, 699(2):1252–1260, Jul 2009. doi: 10.1088/0004-637X/699/2/1252.
- [21] A. Friedmann. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 10:377–386, Jan 1922. doi: 10.1007/BF01332580.
- [22] K. R. Gabriel and R. R. Sokal. A New Statistical Approach to Geographic Variation Analysis. *Systematic Zoology*, 18(3):259, Sept. 1969. ISSN 00397989. doi: 10.2307/2412323. URL <https://academic.oup.com/sysbio/article-lookup/doi/10.2307/2412323>.
- [23] L. H. Garrison, D. J. Eisenstein, D. Ferrer, J. L. Tinker, P. A. Pinto, and D. H. Weinberg. The Abacus Cosmos: A Suite of Cosmological N-body Simulations. *ApJS*, 236(2):43, June 2018. ISSN 0067-0049. doi: 10.3847/1538-4365/aabfd3. URL <https://doi.org/10.3847/1538-4365/aabfd3>.
- [24] I. Gott, J. R., E. L. Turner, and S. J. Aarseth. N-body simulations of galaxy clustering. III. The covariance function. *Astrophysical Journal*, 234:13–26, Nov 1979. doi: 10.1086/157468.
- [25] E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15(3):168–173, Mar 1929. doi: 10.1073/pnas.15.3.168.
- [26] E. Hubble. The Distribution of Extra-Galactic Nebulae. *Astrophysical Journal*, 79:8, Jan 1934. doi: 10.1086/143517.

- [27] D. G. Kirkpatrick and J. D. Radke. A Framework for Computational Morphology. In *Machine Intelligence and Pattern Recognition*, volume 2, pages 217–248. Elsevier, 1985. ISBN 978-0-444-87806-9. doi: 10.1016/B978-0-444-87806-9.50013-X. URL <https://linkinghub.elsevier.com/retrieve/pii/B978044487806950013X>.
- [28] A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, A. A. Klypin, S. Gottloeber, B. Allgood, and J. R. Primack. The Dark Side of the Halo Occupation Distribution. *ApJ*, 609(1):35–49, July 2004. ISSN 0004-637X, 1538-4357. doi: 10.1086/420959. URL <http://arxiv.org/abs/astro-ph/0308519>. arXiv: astro-ph/0308519.
- [29] S. D. Landy and A. S. Szalay. Bias and Variance of Angular Correlation Functions. *Astrophysical Journal*, 412:64, Jul 1993. doi: 10.1086/172900.
- [30] G. Lavaux and B. D. Wandelt. Precision cosmology with voids: definition, methods, dynamics. *Monthly Notices of the Royal Astronomical Society*, 403(3):1392–1408, Apr. 2010. ISSN 00358711, 13652966. doi: 10.1111/j.1365-2966.2010.16197.x. URL <http://arxiv.org/abs/0906.4101>. arXiv: 0906.4101.
- [31] G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Société Scientifique de Bruxelles*, 47:49–59, Jan 1927.
- [32] S. Nadathur. Testing cosmology with a catalogue of voids in the BOSS galaxy surveys. *Mon. Not. R. Astron. Soc.*, 461(1):358–370, Sept. 2016. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stw1340. URL <http://arxiv.org/abs/1602.04752>. arXiv: 1602.04752.
- [33] M. C. Neyrinck. ZOBOV: a parameter-free void-finding algorithm. *Monthly Notices of the Royal Astronomical Society*, 386(4):2101–2109, Jun 2008. doi: 10.1111/j.1365-2966.2008.13180.x.
- [34] N. D. Padilla, L. Ceccarelli, and D. G. Lambas. Spatial and Dynamical Properties of Voids in a LCDM Universe. *Monthly Notices of the Royal Astronomical Society*, 363(3): 977–990, Nov. 2005. ISSN 00358711, 13652966. doi: 10.1111/j.1365-2966.2005.09500.x. URL <http://arxiv.org/abs/astro-ph/0508297>. arXiv: astro-ph/0508297.
- [35] D. Park and J. Lee. Void Ellipticity Distribution as a Probe of Cosmology. *Physical Review Letters*, 98(8):081301, Feb. 2007. doi: 10.1103/PhysRevLett.98.081301. URL <https://ui.adsabs.harvard.edu/abs/2007PhRvL..98h1301P/abstract>.
- [36] P. J. E. Peebles. *The Large-Scale Structure of the Universe*. Princenton Series in Physics. Princenton University Press, The address, 1 edition, 1980. ISBN 0-691-08239-1.

- [37] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 142:419–421, Jul 1965. doi: 10.1086/148307.
- [38] A. Pisani, E. Massara, D. N. Spergel, D. Alonso, T. Baker, Y.-C. Cai, M. Cautun, C. Davies, V. Demchenko, O. Doré, A. Goulding, M. Habouzit, N. Hamaus, A. Hawken, C. M. Hirata, S. Ho, B. Jain, C. D. Kreisch, F. Marulli, N. Padilla, G. Pollina, M. Sahlén, R. K. Sheth, R. Somerville, I. Szapudi, R. van de Weygaert, F. Villaescusa-Navarro, B. D. Wandelt, and Y. Wang. Cosmic voids: a novel probe to shed light on our Universe. *arXiv:1903.05161 [astro-ph]*, Mar. 2019. URL <http://arxiv.org/abs/1903.05161>. arXiv: 1903.05161.
- [39] Planck Collaboration, P. A. R. Ade, and e. a. Aghanim. Planck 2013 results. XVI. Cosmological parameters. *Astronomy and Astrophysics*, 571:A16, Nov 2014. doi: 10.1051/0004-6361/201321591.
- [40] E. Platen, R. van de Weygaert, and B. J. T. Jones. A cosmic watershed: the WVF void detection technique. *Monthly Notices of the Royal Astronomical Society*, 380: 551–570, Sept. 2007. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2007.12125.x. URL <http://adsabs.harvard.edu/abs/2007MNRAS.380..551P>.
- [41] W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *Astrophysical Journal*, 187:425–438, Feb 1974. doi: 10.1086/152650.
- [42] V. C. Rubin and J. Ford, W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophysical Journal*, 159:379, Feb 1970. doi: 10.1086/150317.
- [43] P. Schneider. Cosmology II: Inhomogeneities in the Universe. In P. Schneider, editor, *Extragalactic Astronomy and Cosmology: An Introduction*, pages 277–307. Springer, Berlin, Heidelberg, 2006. ISBN 978-3-540-33175-9. doi: 10.1007/978-3-540-33175-9_7. URL https://doi.org/10.1007/978-3-540-33175-9_7.
- [44] R. K. Sheth and R. van de Weygaert. A hierarchy of voids: much ado about nothing. *Monthly Notices of the Royal Astronomical Society*, 350(2):517, May 2004. doi: 10.1111/j.1365-2966.2004.07661.x. URL <https://ui.adsabs.harvard.edu/abs/2004MNRAS.350..517S/abstract>.
- [45] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, June 2005. ISSN 0028-0836. doi: 10.1038/nature03597. URL <http://adsabs.harvard.edu/abs/2005Natur.435..629S>.

- [46] P. M. Sutter, G. Lavaux, N. Hamaus, A. Pisani, B. D. Wandelt, M. Warren, F. Villaescusa-Navarro, P. Zivick, Q. Mao, and B. B. Thompson. VIDE: The Void IDentification and Examination toolkit. *Astronomy and Computing*, 9:1–9, Mar. 2015. ISSN 2213-1337. doi: 10.1016/j.ascom.2014.10.002. URL <http://adsabs.harvard.edu/abs/2015A%26C.....9....1S>.
- [47] G. T. Toussaint. Computational Geometry: Recent Results Relevant to Pattern Recognition. In P. A. Devijver and J. Kittler, editors, *Pattern Recognition Theory and Applications*, NATO ASI Series, pages 295–305, Berlin, Heidelberg, 1987. Springer. ISBN 978-3-642-83069-3. doi: 10.1007/978-3-642-83069-3_24.
- [48] R. van de Weygaert. Voids and the Cosmic Web: cosmic depression & spatial complexity. *Proceedings of the International Astronomical Union*, 11 (S308):493–523, June 2014. ISSN 1743-9213, 1743-9221. doi: 10.1017/S1743921316010504. URL https://www.cambridge.org/core/product/identifier/S1743921316010504/type/journal_article.
- [49] S. von Hoerner. Die numerische Integration des n-Körper-Problemes für Sternhaufen. I. *Z. Astrophys*, 50:184–214, Jan 1960.
- [50] Y. B. Zel'Dovich. Gravitational instability: an approximate theory for large density perturbations. *Astronomy and Astrophysics*, 5:84–89, Mar 1970.
- [51] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, Jan 1933.