

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

Dipartimento di Fisica e Astronomia “Augusto Righi”

Corso di Laurea in Fisica

**Validazione del nuovo metodo dinamico per  
la ricerca di vuoti cosmici BitVF su  
cataloghi di traccianti simulati**

**Relatore:**

**Prof. Federico Marulli**

**Presentata da:**

**Francesco Giunti**

**Correlatori:**

**Dott.ssa Sofia Contarini**

**Dott. Simone Sartori**

Anno Accademico 2025/2026

# Abstract

I vuoti cosmici, vaste regioni sottodense che costituiscono la maggior parte del volume dell'Universo, stanno emergendo come sonde cosmologiche di rilievo, poiché le loro proprietà statistiche dipendono fortemente dai valori dei parametri del modello cosmologico.

Tuttavia, l'identificazione dei vuoti cosmici è complessa: non esiste una definizione univoca e gli algoritmi esistenti sfruttano metodologie identificative eterogenee. Questa tesi testa e valida un nuovo algoritmo, il *Back-in-time Void Finder*, o **BitVF**, sviluppato dal Dott. S. Sartori e dal gruppo di ricerca del Prof. F. Marulli del Dipartimento di Fisica e Astronomia 'Augusto Righi' dell'Università di Bologna. Invece di cercare minimi locali nel campo di densità, esso ricostruisce a ritroso le traiettorie dei traccianti tramite il principio di minima azione e l'approssimazione di Zel'dovich, identificando i vuoti come i massimi della divergenza del campo di spostamento.

L'analisi, condotta su simulazioni N-corpi di aloni di materia oscura, confronta i cataloghi di vuoti ottenuti con **BitVF** con quelli ottenuti con *Void IDentification and Examination toolkit*, o **VIDE**, un algoritmo geometrico di riferimento e consolidato in letteratura. La validazione si articola nell'analisi delle proprietà statistiche e nell'applicazione di tre metriche di accoppiamento vuoto-a-vuoto, per valutare la riproducibilità di **BitVF** e la sua robustezza nel tracciare i vuoti tra spazio reale e spazio dei redshift.

**BitVF** produce cataloghi stabili e riproducibili (completezza  $> 83\%$ ) e, rispetto a **VIDE**, risulta più robusto nel tracciare i vuoti tra i due spazi: le proprietà statistiche sono più stabili e le metriche di accoppiamento mostrano una maggiore corrispondenza spaziale. La mappatura dei volumi conferma una maggiore coerenza nei singoli accoppiamenti per **BitVF**, mentre **VIDE** riproduce meglio il volume medio.

Questa tesi valida **BitVF** in simulazioni e, seppur con delle limitazioni, ne supporta l'uso futuro su dati osservativi reali.

# Indice

|                                                                            |           |
|----------------------------------------------------------------------------|-----------|
| <b>Introduzione</b>                                                        | <b>5</b>  |
| <b>1 Cosmologia</b>                                                        | <b>8</b>  |
| 1.1 Cenni di relatività generale . . . . .                                 | 8         |
| 1.2 Principio cosmologico e metrica di Friedmann-Lemaître-Robertson-Walker | 9         |
| 1.2.1 Costante cosmologica ed espansione dell'Universo . . . . .           | 11        |
| 1.3 Legge di Hubble-Lemaître . . . . .                                     | 11        |
| 1.3.1 Redshift cosmologico . . . . .                                       | 14        |
| 1.4 Equazioni di Friedmann . . . . .                                       | 15        |
| 1.4.1 Modello $\Lambda$ CDM . . . . .                                      | 16        |
| 1.4.2 Storia termica dell'universo . . . . .                               | 17        |
| 1.5 Materia ed energia oscura . . . . .                                    | 19        |
| <b>2 Formazione delle strutture cosmiche</b>                               | <b>22</b> |
| 2.1 Teoria lineare delle perturbazioni . . . . .                           | 23        |
| 2.1.1 Contrasto di densità . . . . .                                       | 23        |
| 2.1.2 Descrizione fluidodinamica . . . . .                                 | 23        |
| 2.1.3 Instabilità di Jeans . . . . .                                       | 24        |
| 2.2 Crescita delle perturbazioni . . . . .                                 | 25        |
| 2.2.1 Fattore di crescita . . . . .                                        | 25        |
| 2.3 Descrizione statistica del campo di densità . . . . .                  | 26        |
| 2.3.1 Funzione di correlazione a due punti . . . . .                       | 26        |
| 2.3.2 Spettro di potenza e parametro $\sigma_8$ . . . . .                  | 27        |
| 2.4 Regime non lineare . . . . .                                           | 28        |
| 2.4.1 Limiti della teoria lineare . . . . .                                | 28        |

|          |                                                                                                            |           |
|----------|------------------------------------------------------------------------------------------------------------|-----------|
| 2.4.2    | Collasso sferico . . . . .                                                                                 | 28        |
| 2.4.3    | Approssimazione di Zel'dovich . . . . .                                                                    | 29        |
| 2.4.4    | Evoluzione delle sottodensità . . . . .                                                                    | 30        |
| 2.4.5    | Bias di aloni oscuri e traccianti luminosi . . . . .                                                       | 30        |
| 2.5      | Simulazioni cosmologiche N-corpi . . . . .                                                                 | 31        |
| <b>3</b> | <b>Back-in-time Void Finder</b>                                                                            | <b>34</b> |
| 3.1      | Generalità sulla procedura di <i>void finding</i> . . . . .                                                | 35        |
| 3.2      | Proprietà statistiche dei vuoti cosmici . . . . .                                                          | 37        |
| 3.2.1    | <i>Void size function</i> . . . . .                                                                        | 39        |
| 3.2.2    | Profili di densità dei vuoti . . . . .                                                                     | 41        |
| 3.3      | L'algoritmo BitVF . . . . .                                                                                | 42        |
| 3.3.1    | Ricostruzione del campo di spostamento . . . . .                                                           | 43        |
| 3.3.2    | Calcolo del campo di densità . . . . .                                                                     | 44        |
| 3.3.3    | Identificazione dei vuoti e definizione dei raggi . . . . .                                                | 44        |
| 3.3.4    | Parametri modificabili di BitVF . . . . .                                                                  | 45        |
| 3.4      | Procedura di <i>void cleaning</i> . . . . .                                                                | 46        |
| 3.5      | CosmoBolognaLib . . . . .                                                                                  | 47        |
| <b>4</b> | <b>Analisi preliminare di BitVF: proprietà statistiche e confronto con VIDE</b>                            | <b>50</b> |
| 4.1      | Esecuzione preliminare di BitVF . . . . .                                                                  | 51        |
| 4.2      | Confronto tra BitVF e VIDE . . . . .                                                                       | 53        |
| 4.3      | Analisi del parametro <code>merging_threshold</code> per cataloghi VIDE . . . . .                          | 60        |
| 4.3.1    | <i>Cleaning</i> dei cataloghi BitVF . . . . .                                                              | 65        |
| <b>5</b> | <b>Validazione di BitVF: analisi di riproducibilità e confronto tra spazio reale e spazio dei redshift</b> | <b>68</b> |
| 5.1      | Analisi 1: metrica del primo vicino . . . . .                                                              | 71        |
| 5.1.1    | Esecuzione1 vs Esecuzione2: riproducibilità di BitVF . . . . .                                             | 72        |
| 5.1.2    | BitVF e VIDE: compatibilità tra spazio reale e dei redshift . . . . .                                      | 75        |
| 5.2      | Analisi 2: metrica della somiglianza di raggio con filtro spaziale . . . . .                               | 79        |
| 5.2.1    | Esecuzione1 vs Esecuzione2: riproducibilità di BitVF . . . . .                                             | 80        |
| 5.2.2    | BitVF e VIDE: compatibilità tra spazio reale e dei redshift . . . . .                                      | 84        |
| 5.3      | Analisi 3: mappatura dei volumi . . . . .                                                                  | 86        |

|          |                                                                                              |            |
|----------|----------------------------------------------------------------------------------------------|------------|
| 5.3.1    | Esecuzione1 vs Esecuzione2: coerenza della mappatura dei volumi<br>in spazio reale . . . . . | 87         |
| 5.3.2    | BitVF: mappatura dei volumi fra spazio reale e dei redshift . . . .                          | 89         |
| 5.3.3    | VIDE: mappatura dei volumi fra spazio reale e dei redshift . . . .                           | 90         |
| 5.3.4    | Confronto tra le metriche e prospettive per un'analisi combinata .                           | 92         |
| <b>6</b> | <b>Conclusioni e prospettive future</b>                                                      | <b>94</b>  |
| 6.1      | Riassunto dei risultati principali . . . . .                                                 | 94         |
| 6.2      | Implicazioni per l'uso dei vuoti cosmici come sonde cosmologiche . . . .                     | 97         |
| 6.3      | Limitazioni del presente lavoro . . . . .                                                    | 98         |
| 6.4      | Prospettive future e considerazioni conclusive . . . . .                                     | 99         |
| <b>A</b> | <b>Statistiche utilizzate nell'analisi</b>                                                   | <b>107</b> |
| A.1      | Residuo in sigma . . . . .                                                                   | 107        |
| A.2      | Frazione di match . . . . .                                                                  | 107        |
| A.3      | Correlazione pesata . . . . .                                                                | 108        |
| A.4      | Chi-quadro ridotto . . . . .                                                                 | 108        |
| A.5      | Errori sulle statistiche . . . . .                                                           | 108        |

# Introduzione

La cosmologia è la branca della fisica che studia l'Universo in quanto tale: le sue leggi, la sua evoluzione temporale e composizione. Si considerano quindi scale spaziali estremamente ampie, sulle quali i traccianti (ad esempio galassie e ammassi di galassie) appaiono oggi disposti in strutture che rivelano zone molto concentrate e zone a bassa densità. Queste ultime prendono il nome di vuoti cosmici e il loro studio è cruciale al fine della validazione del modello cosmologico standard. L'assenza di una definizione univoca di vuoto cosmico ha portato nel tempo alla nascita di diversi algoritmi di ricerca, basati su varie metodologie. Questa tesi si pone l'obiettivo di testare e validare uno di essi, che utilizza l'approccio dinamico nella ricerca dei vuoti e prende il nome di *Back-in-time Void Finder* (**BitVF**).

L'evoluzione di queste strutture è determinata dall'instabilità gravitazionale a partire da piccole fluttuazioni primordiali, e le loro proprietà statistiche, come la *void size function* (VSF) e i profili di densità, dipendono in modo significativo dai parametri cosmologici, in particolare: la densità di materia,  $\Omega_m$ , e l'ampiezza delle fluttuazioni di densità su scale di  $8 h^{-1}\text{Mpc}$ ,  $\sigma_8$ ; come mostrato da Sheth and van de Weygaert [2004] per la VSF e da Hamaus et al. [2014] per i profili di densità, e da eventuali parametri di teorie della gravità modificata. Inoltre, i vuoti si trovano in un regime moderatamente non lineare, il che li rende complementari ad altre sonde cosmologiche come gli ammassi di galassie o il *lensing* gravitazionale.

La crescente disponibilità di dati da *survey* galattiche di nuova generazione, come Euclid Collaboration et al. [2022], rende sempre più rilevante lo sviluppo e la validazione di algoritmi robusti per l'identificazione dei vuoti. In questo contesto, **BitVF** rappresenta una proposta innovativa: invece di cercare regioni a bassa densità nel campo di densità osservato, l'algoritmo ricostruisce le traiettorie dei traccianti indietro nel tempo, utilizzando il principio di minima azione di Hamilton e l'approssimazione di Zel'dovich

[Zel’dovich, 1970], e identifica i vuoti come le regioni di massimo della divergenza del campo di spostamento. Questo approccio, che sfrutta nello specifico la relazione tra il campo di spostamento e il contrasto di densità, è pensato per essere meno affetto dallo *shot noise* e più robusto rispetto agli effetti delle distorsioni dello spazio redshift (RSD) rispetto agli algoritmi puramente geometrici.

La tesi si articola in due parti principali. Nella prima parte (Capitoli 1-3) vengono introdotti i fondamenti teorici necessari per la comprensione del lavoro. Il Capitolo 1 presenta i principi della cosmologia moderna, dalla relatività generale al modello  $\Lambda$ CDM, con particolare attenzione alle equazioni di Friedmann e alla composizione dell’Universo. Il Capitolo 2 discute la formazione delle strutture cosmiche, dalla teoria lineare delle perturbazioni al regime non lineare, introducendo alcuni strumenti statistici per la descrizione del campo di densità e il ruolo delle simulazioni N-corpi. Il Capitolo 3 descrive le principali proprietà statistiche dei vuoti cosmici e gli algoritmi per la loro identificazione, classificandoli in tre categorie principali (basati sulla densità, sulla geometria e sulla dinamica), e presenta in dettaglio l’algoritmo **BitVF**: le sue fasi principali e i parametri che lo controllano, concludendo con una descrizione delle librerie **CosmoBolognaLib** utilizzate per le analisi.

La seconda parte della tesi (Capitoli 4-6) è dedicata alla presentazione dei risultati. Nel Capitolo 4 vengono presentate le proprietà statistiche dei cataloghi di vuoti prodotti da **BitVF**, con una validazione preliminare e un confronto sistematico con i cataloghi ottenuti da *Void IDentification and Examination toolkit*, o **VIDE**, un algoritmo geometrico consolidato in letteratura. Vengono analizzate le VSF, i monopoli e i quadrupoli, e viene studiato l’effetto del parametro **merging\_threshold** sui cataloghi **VIDE**. Il Capitolo 5 si concentra sulla validazione di **BitVF** attraverso tre diverse metriche di accoppiamento: il criterio del primo vicino, la somiglianza di raggio e la mappatura dei volumi. Queste metriche, applicate sia al confronto tra due run indipendenti di **BitVF** che al confronto tra spazio reale e spazio dei redshift per entrambi gli algoritmi, forniscono una valutazione completa della riproducibilità e della robustezza di **BitVF**. Infine, il Capitolo 6 presenta i risultati principali, discute le implicazioni per l’uso dei vuoti come sonde cosmologiche e propone possibili sviluppi futuri.

I risultati ottenuti mostrano che **BitVF** produce cataloghi di vuoti stabili e riproducibili, con frazioni di accoppiamento vuoto-a-vuoto superiori all’83% tra run indipendenti. Nel confronto con **VIDE**, **BitVF** si dimostra complessivamente più robusto nel tracciare i

vuoti tra spazio reale e dei redshift, con frazioni di accoppiamento più elevate e distanze normalizzate più piccole, sebbene con un comportamento differenziato nella mappatura dei volumi. La minore sensibilità agli effetti RSD di **BitVF** è evidente sia nell'analisi delle VSF che in quella dei profili di densità, e viene confermata dalle metriche di accoppiamento utilizzate.

Questo lavoro rappresenta un supporto alla validazione di **BitVF** come strumento per l'analisi dei vuoti cosmici, e apre la strada a future applicazioni su cataloghi osservativi reali. La crescente disponibilità di dati da grandi *survey*, combinata con la robustezza e l'affidabilità di algoritmi come **BitVF**, renderà possibile sfruttare appieno il potenziale cosmologico dei vuoti, contribuendo a vincolare i parametri del modello cosmologico e a testare teorie alternative della gravità.

# Capitolo 1

## Cosmologia

In questo primo capitolo verranno introdotti i fondamenti teorici e principi su cui si basa la cosmologia, le principali leggi che governano l'Universo su scale cosmologiche e verrà presentata una panoramica sulla sua composizione ed evoluzione.

### 1.1 Cenni di relatività generale

L'interazione cardine per la descrizione della formazione ed evoluzione delle strutture su scale dei Mpc è quella gravitazionale, modellizzata in cosmologia dalla teoria della relatività generale [Einstein, 1915]. L'obiettivo di questo capitolo è quello di introdurre i concetti basilari della teoria per facilitare la comprensione dei capitoli successivi.

Essa si basa sull'ipotesi che la presenza di massa, o energia, comporti una deformazione geometrica dello spaziotempo, i cui effetti siano empiricamente riscontrabili sotto forma di interazione gravitazionale. Lo spaziotempo viene matematicamente descritto come una varietà quadridimensionale e la sua geometria codificata nel *tensore metrico*,  $g_{\mu\nu}$ ; definito in modo che la distanza fra due eventi infinitesimamente distanti  $x^\mu = (ct; \vec{x})$  e  $x^\mu + dx^\mu = (c(t + dt); \vec{x} + d\vec{x})$  sia esprimibile come:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad (1.1)$$

dove viene utilizzata la convenzione di Einstein sulla contrazione degli indici: la ripetizione di un indice sottointende la somma su ognuno dei valori che può assumere, in questo caso 0, 1, 2, 3.

Le relazioni che legano la presenza di massa ed energia alla curvatura geometrica dello spaziotempo prendono il nome di equazioni di campo di Einstein e sono rappresentabili in forma compatta come:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}, \quad (1.2)$$

ove  $G$  è la costante gravitazionale universale;  $T_{\mu\nu}$  è il tensore energia-impulso, che porta i contributi della densità di energia, la pressione e i flussi di materia nell'universo; mentre  $R_{\mu\nu}$  e  $R$  sono rispettivamente il tensore di Ricci e lo scalare di Ricci, ottenibili da opportune contrazioni di un tensore di tipo (1,3) chiamato tensore di Riemann-Christoffel,  $R^\alpha_{\mu\nu\rho}$ .

In termini semplici, quest'ultimo racchiude le informazioni geometriche inerenti alla varietà spaziotemporale in esame.

## 1.2 Principio cosmologico e metrica di Friedmann-Lemaître-Robertson-Walker

Una delle prime evidenze osservative alla base della cosmologia moderna risale agli studi di Hubble, che dimostrarono la natura extragalattica della galassia di Andromeda (Fig. 1.1) e, successivamente, portarono alla scoperta dell'espansione dell'Universo [Hubble, 1929a,b].

Il quadro teorico attuale si fonda sull'ipotesi che l'Universo, su scale sufficientemente grandi ( $\gtrsim 100$  Mpc), sia omogeneo e isotropo. Queste proprietà costituiscono il Principio Cosmologico. L'omogeneità implica che le proprietà statistiche dell'Universo siano invarianti per traslazioni spaziali, mentre l'isotropia richiede l'assenza di direzioni privilegiate, ovvero invarianza per rotazioni spaziali. Evidenze osservative a supporto di tali ipotesi sono fornite, ad esempio, da misure fatte sulla radiazione cosmica di fondo (CMB) [Planck Collaboration et al., 2020a].

Nel contesto della relatività generale, queste simmetrie restringono fortemente la forma ammissibile della metrica spazio-temporale. In particolare, la metrica può essere scritta nella forma:

$$ds^2 = (cdt)^2 - g_{ij}dx^i dx^j, \quad (1.3)$$

dove  $i, j = 1, 2, 3$  sono indici spaziali.

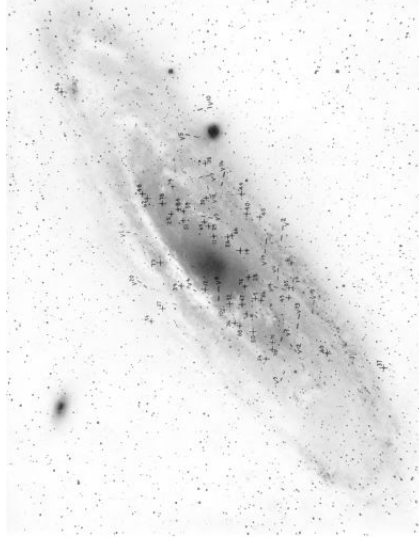


Figura 1.1: Tavola IV dell’articolo di Hubble [Hubble, 1929a]. L’immagine mostra la galassia di Andromeda (M31) con l’indicazione di singole stelle variabili Cefeidi, utilizzate da Hubble per determinare la distanza dell’oggetto. La scoperta che M31 si trovasse al di là dei confini della Via Lattea rappresentò il primo passo verso la cosmologia extragalattica moderna.

Imponendo omogeneità e isotropia, le componenti spaziali del tensore metrico assumono una forma specifica che, espressa in coordinate comoventi sferiche  $(r, \theta, \phi)$ , conduce alla metrica di Friedmann-Lemaître-Robertson-Walker (FLRW):

$$ds^2 = (cdt)^2 - a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2(\theta) d\phi^2) \right]. \quad (1.4)$$

In questa espressione:

- $a(t)$  è il fattore di scala, che descrive l’evoluzione temporale delle distanze cosmologiche;
- $k$  è la curvatura spaziale, che può assumere i valori  $k = -1, 0, +1$ , corrispondenti rispettivamente a geometria aperta (Universo iperbolico), piatta (Universo piatto) e chiusa (Universo sferico). Osservazioni cosmologiche, come ad esempio quelle della CMB, indicano la compatibilità dell’Universo con  $k = 0$ , entro gli errori sperimentali [Planck Collaboration et al., 2020b].

I parametri dinamici della metrica, in particolare  $a(t)$ , sono determinati dalle equazioni di campo di Einstein applicate a un fluido. Si assume pertanto che l'Universo possa essere descritto come un fluido ideale su scale cosmologiche, ipotesi sostenuta dal Principio Cosmologico.

### 1.2.1 Costante cosmologica ed espansione dell'Universo

Le equazioni di campo di Einstein possono essere generalizzate introducendo un termine additivo proporzionale alla metrica, noto come costante cosmologica ( $\Lambda$ ):

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \quad (1.5)$$

Originariamente, tale termine fu introdotto da Einstein per ottenere una soluzione statica delle equazioni di campo. Tuttavia, la scoperta dell'espansione cosmica rese questa interpretazione superata.

In un contesto moderno, la costante cosmologica è reinterpretata come una componente energetica con densità costante e pressione negativa, caratterizzata da equazione di stato  $w = -1$ ; si veda ad esempio il lavoro di Bianchi and Rovelli [2010] per una breve panoramica storica e una guida alla reinterpretazione di  $\Lambda$ .

Evidenze osservative indipendenti, come ad esempio misure di supernovae di tipo Ia, anisotropie della CMB e distribuzione delle strutture su larga scala, indicano che l'espansione dell'Universo è attualmente accelerata [Perlmutter et al., 1999][Riess et al., 1998][Planck Collaboration et al., 2020b].

In questo quadro, essa rappresenta il contributo dominante all'energia totale dell'Universo a tempi recenti ed è alla base del modello cosmologico standard  $\Lambda$ CDM.

## 1.3 Legge di Hubble-Lemaître

L'evidenza osservativa dell'espansione cosmica, inizialmente sistematizzata da Hubble [Hubble, 1929b], porta alla distinzione tra coordinate fisiche e coordinate comoventi. In un universo descritto dalla metrica FLRW, le coordinate comoventi  $(r, \theta, \phi)$  sono definite in modo tale che gli osservatori solidali con l'espansione mantengano coordinate spaziali costanti.

In tale quadro, la distanza propria tra due punti a tempo cosmologico  $t$  è definita come:

$$D_p(t) = a(t) \chi, \quad (1.6)$$

dove  $\chi$  è fornita da:

$$\chi = \int_0^r \frac{du}{\sqrt{1 - ku^2}}. \quad (1.7)$$

Essa è indipendente dal tempo, dunque la dipendenza temporale della distanza propria è interamente contenuta nel fattore di scala  $a(t)$ . In particolare, fissato il tempo presente  $t_0$ , si definisce la distanza comovente come:

$$D_c \equiv D_p(t_0) = a_0 \chi, \quad (1.8)$$

da cui segue la relazione generale:

$$D_p(t) = \frac{a(t)}{a_0} D_c. \quad (1.9)$$

Questa formulazione mostra che la struttura comovente dello spazio rimane invariata, mentre le distanze fisiche evolvono nel tempo attraverso il fattore di scala. Non vi è una contrazione o espansione delle coordinate comoventi, bensì una ridistribuzione delle distanze fisiche.

La velocità di recessione radiale tra due punti comoventi si definisce come:

$$v_{rad}(t) \equiv \frac{d}{dt} D_p(t). \quad (1.10)$$

Poiché  $\chi$  non dipende dal tempo, si ottiene:

$$v_{rad}(t) = \dot{a}(t) \chi = \frac{\dot{a}(t)}{a(t)} D_p(t), \quad (1.11)$$

ove  $\dot{a}(t) \equiv \frac{d}{dt} a(t)$ .

Introducendo quindi il parametro di Hubble:

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}, \quad (1.12)$$

possiamo giungere alla legge di Hubble-Lemaître, espressa nella forma:

$$v_{rad}(t) = H(t) D_p(t). \quad (1.13)$$

La Fig. 1.2 mostra la rappresentazione grafica originale di questa relazione, così come pubblicata da Hubble nel 1929.

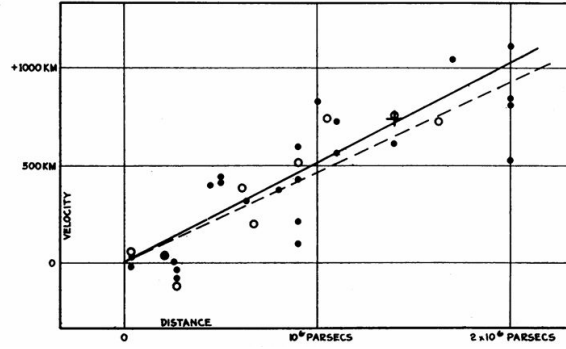


Figura 1.2: Diagramma velocità-distanza originale di Hubble [Hubble, 1929b]. L'asse delle ascisse riporta le distanze delle galassie (in Mpc), mentre l'asse delle ordinate le rispettive velocità di recessione (in km/s). La relazione lineare osservata costituì la prima prova empirica dell'espansione dell'Universo.

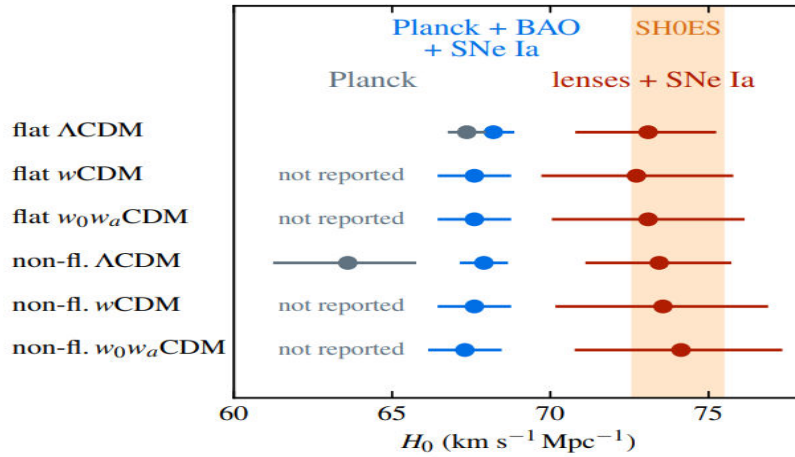


Figura 1.3: Confronto visivo tra diverse determinazioni della costante di Hubble  $H_0$ . Tra i modelli elencati sulla sinistra, quello di riferimento per questa trattazione è il *flat*  $\Lambda$ CDM. Adattata da Taubenberger et al. [2019].

Valutando la legge 1.13 al tempo presente si definisce il parametro  $H_0$ , detto costante di Hubble, che quantifica il tasso di espansione attuale dell'Universo.

Misure di questa costante mostrano una tensione statistica tra metodi indipendenti, come mostrato in Fig. 1.3. In particolare, quelle basate su supernovae di tipo Ia [Riess et al., 2019] forniscono valori tipicamente più elevati,  $H_0 = (74.03 \pm 1.42) \text{ km s}^{-1} \text{ Mpc}^{-1}$ , rispetto alle stime ottenute da anisotropie della radiazione cosmica di fondo [Planck Collaboration et al., 2020b],  $H_0 = (67.4 \pm 0.5) \text{ km s}^{-1} \text{ Mpc}^{-1}$ . La discrepanza potrebbe essere riconducibile a sistematiche osservative oppure a una deviazione dal modello  $\Lambda$ CDM.

### 1.3.1 Redshift cosmologico

In relatività generale, lo spostamento verso il rosso osservato per sorgenti cosmologiche in allontanamento non è interpretato come effetto Doppler in senso classico, ma come conseguenza dell'espansione dello spazio, interpretata da osservatori comoventi come dilatazione delle distanze fisiche nel loro tempo proprio.

La lunghezza d'onda osservata e quella emessa sono legate dalla relazione:

$$z \equiv \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}. \quad (1.14)$$

Utilizzando la metrica FLRW (Eq. 1.4) e considerando propagazione radiale ( $d\theta = d\phi = 0$ ), si ottiene la relazione generale:

$$1 + z = \frac{a_0}{a(t)}. \quad (1.15)$$

Per redshift piccoli ( $z \ll 1$ ) quindi, la legge di Hubble-Lemaître (Eq. 1.13) può essere linearizzata come:

$$v_{rad} \simeq cz \simeq H_0 D_p, \quad (1.16)$$

che rappresenta un'approssimazione locale valida a basse distanze cosmologiche.

Il redshift cosmologico è quindi una misura diretta dell'espansione dell'Universo e viene utilizzato come indicatore osservativo primario per la stima delle distanze su scale extragalattiche.

## 1.4 Equazioni di Friedmann

Come precedentemente accennato, l'assunzione di validità del principio cosmologico consente di descrivere l'Universo, su scale dell'ordine di centinaia di Mpc, come un fluido continuo, in prima approssimazione ideale, caratterizzato da densità  $\rho$ , pressione  $p$  e quadrivelocità  $u^\mu$ . Applicando la metrica FLRW alle equazioni di campo di Einstein (Eq. 1.2), con tensore energia-impulso  $T_{\mu\nu}$  di fluido perfetto, si ottiene un sistema di 16 equazioni, di cui solo due indipendenti, che governano l'evoluzione del fattore di scala  $a(t)$ . Tali equazioni sono:

$$\dot{a}^2 = -kc^2 + \frac{8\pi}{3}G\rho a^2, \quad (1.17)$$

$$\ddot{a} = -\frac{4\pi}{3}G\left(\rho + \frac{3P}{c^2}\right)a, \quad (1.18)$$

e prendono il nome di *equazioni di Friedmann*. Le quantità  $\rho$  e  $p$  rappresentano la somma dei contributi di tutte le componenti cosmologiche, dunque rispettivamente densità e pressione totali dell'Universo.

Nel limite di fluido ideale, la relazione tra esse è descritta dall'equazione di stato:

$$P = w\rho c^2, \quad (1.19)$$

dove  $w \in [-1, +1]$  è un parametro adimensionale che dipende dalla componente considerata.

Friedmann derivò queste equazioni assumendo un Universo statico, ovvero imponendo  $\dot{a} = 0$  e  $\ddot{a} = 0$ . Dall'Eq. 1.18 segue che tale condizione è soddisfatta solo se:

$$\rho = -\frac{3p}{c^2}, \quad (1.20)$$

che richiede pressione e/o densità totali negative.

Per ottenere soluzioni statiche coerenti fisicamente, vennero utilizzate le equazioni di campo nella forma generale (Eq. 1.5). Le equazioni di Friedmann assumono quindi la forma:

$$\dot{a}^2 = -kc^2 + \left(\frac{8\pi}{3}G\rho + \frac{\Lambda}{3}\right)a^2, \quad (1.21)$$

$$\ddot{a} = \left[ -\frac{4\pi}{3}G \left( \rho + \frac{3p}{c^2} \right) + \frac{\Lambda}{3} \right] a. \quad (1.22)$$

Queste equazioni descrivono sia universi statici sia in espansione accelerata. In particolare, il termine proporzionale a  $\Lambda$  consente  $\ddot{a} > 0$ , in accordo con osservazioni dell'espansione accelerata.

### 1.4.1 Modello $\Lambda$ CDM

Il modello  $\Lambda$ CDM rappresenta il quadro standard della cosmologia moderna. Esso si basa sulle seguenti assunzioni:

1. Vale il principio cosmologico.
2. La dinamica gravitazionale è descritta dalla relatività generale e la metrica è di tipo FLRW.
3. L'Universo è composto da materia ordinaria (detta barionica), radiazione (fotoni e neutrini relativistici), materia oscura fredda (CDM) ed energia oscura. La CDM è necessaria per la formazione delle strutture, mentre l'energia oscura, modellizzata tramite  $\Lambda$ , permette l'espansione accelerata in tempi cosmologici recenti.
4. Le perturbazioni primordiali sono gaussiane e quasi invarianti di scala, originate durante una fase inflazionaria.

Il parametro  $w$ , introdotto nell'equazione di stato (Eq. 1.19), assume valori caratteristici per le diverse componenti:

1.  $w = 0$  per materia non relativistica, dunque barionica e CDM;
2.  $w = \frac{1}{3}$  per la radiazione;
3.  $w = -1$  per l'energia oscura.

Assumendo  $w$  costante nel tempo, dall'equazione di conservazione dell'energia si ottiene:

$$\rho(t) = \rho_0 a(t)^{-3(1+w)} = \rho_0 (1+z)^{3(1+w)}. \quad (1.23)$$

Questa relazione permette di ricostruire l'evoluzione delle diverse componenti in funzione del redshift.

## 1.4.2 Storia termica dell'universo

L'Universo ha un'età finita, come evidenziato dal paradosso di Olbers: se l'Universo fosse infinito ed eterno, il cielo notturno sarebbe uniformemente luminoso. L'evoluzione cosmica è quindi caratterizzata da una successione di epoche distinte, in cui le diverse componenti energetiche hanno dominato la dinamica dell'Universo.

Una stima dell'ordine di grandezza dell'età cosmica si ottiene dalla legge di Hubble-Lemaître (1.13), assumendo un tasso di espansione costante:

$$t_H \sim \frac{1}{H_0}. \quad (1.24)$$

Il tempo di Hubble rappresenta una scala temporale caratteristica, ma non coincide con l'età reale, poiché il parametro di Hubble (1.12) evolve nel tempo. Misure osservative, come ad esempio quelle ottenute dalla missione Planck [Planck Collaboration et al., 2020b], indicano:

$$t_0 = (13.797 \pm 0.023) \text{ Gyr}. \quad (1.25)$$

L'evoluzione cosmica è descritta considerando le diverse componenti energetiche, ciascuna con equazione di stato di fluido perfetto (1.19).

Dalla conservazione dell'energia segue:

$$\rho_i(a) \propto a^{-3(1+w_i)}, \quad (1.26)$$

da cui:

- radiazione:  $\rho_r \propto a^{-4}$ ;
- materia:  $\rho_m \propto a^{-3}$ ;
- energia oscura:  $\rho_\Lambda = \text{costante}$ .

L'importanza relativa delle componenti varia nel tempo, definendo diverse epoche cosmologiche.

Per descrivere quantitativamente la composizione energetica dell'Universo si introducono i parametri di densità adimensionali:

$$\Omega_i(z) \equiv \frac{\rho_i(z)}{\rho_{\text{crit}}(z)}, \quad \rho_{\text{crit}}(z) = \frac{3H^2(z)}{8\pi G}, \quad (1.27)$$

dove  $\rho_i(z)$  rappresenta la densità di energia della componente  $i$ -esima (radiazione, materia o energia oscura), mentre  $\rho_{\text{crit}}(z)$  è la densità critica necessaria affinché l'Universo sia spazialmente piatto al redshift  $z$ .

La quantità  $H(z)$  è il parametro di Hubble in funzione del redshift, parametrizzazione equivalente a quella temporale, e contiene l'informazione dinamica sull'espansione. Ne segue che i parametri  $\Omega_i(z)$  dipendono dal tempo cosmologico, anche nel caso in cui alcune componenti abbiano densità fisica costante.

La somma dei parametri di densità per ognuna delle componenti definisce il parametro di densità totale:

$$\Omega_{\text{tot}}(z) = \Omega_r(z) + \Omega_m(z) + \Omega_\Lambda(z), \quad (1.28)$$

da cui si introduce il parametro di curvatura:

$$\Omega_K(z) = 1 - \Omega_{\text{tot}}(z). \quad (1.29)$$

In particolare,  $\Omega_K = 0$  corrisponde a geometria spaziale piatta, che come già detto è compatibile con le osservazioni più recenti.

L'evoluzione relativa delle diverse componenti determina le principali epoche cosmologiche, illustrate in Fig. 1.4 assieme all'evoluzione temporale delle densità di radiazione, materia ed energia oscura. Ricordando la relazione fra redshift e fattore di scala (Eq. 1.15): valori elevati di  $z$  corrispondono quindi a fattori di scala piccoli e, di conseguenza, a epoche cosmiche più remote.

1. Era della radiazione ( $z \gtrsim 10^4$ ), in cui  $\rho_r \gg \rho_m$ ;
2. era della materia, in cui  $\rho_m \gg \rho_r$  e domina la formazione delle strutture;
3. era dell'energia oscura, a bassi redshift, in cui  $\rho_\Lambda$  domina e l'espansione accelera.

Durante l'espansione cosmica, la temperatura della radiazione evolve come:

$$T(t) \propto \frac{1}{a(t)}, \quad (1.30)$$

in accordo con il redshift delle lunghezze d'onda dei fotoni.

Un evento chiave è la ricombinazione ( $z \sim 1100$ ), quando protoni ed elettroni si combinano per formare atomi neutri. Questo processo riduce drasticamente la sezione

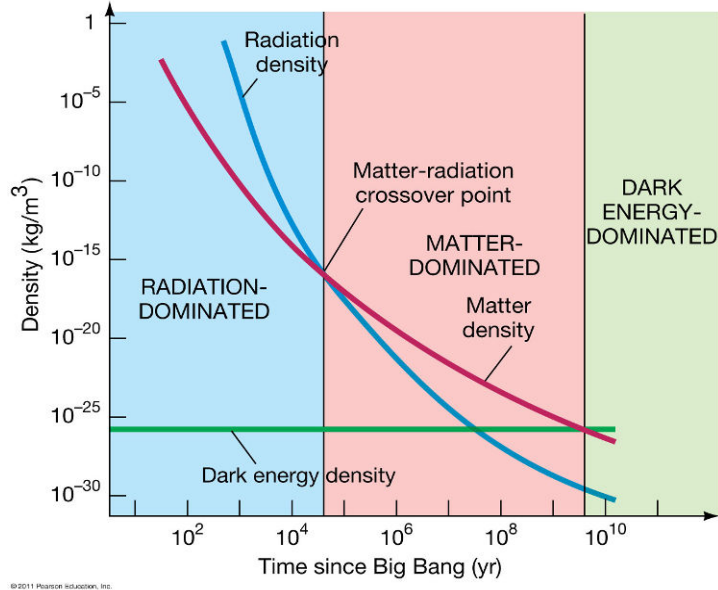


Figura 1.4: Andamento delle densità di energia  $\rho$  delle diverse componenti cosmologiche in funzione del tempo. Sono mostrate la densità di radiazione in blu, la densità di materia in rosso e la densità di energia oscura in verde. L'asse delle ordinate è in scala logaritmica ( $\text{kg}/\text{m}^3$ ), l'ascissa in anni dal Big Bang. Sono evidenziate tre epoche cosmologiche: era della radiazione in blu, era della materia in rosso ed era dell'energia oscura in verde, che arriva a tempi recenti. Tratta da Brau [2011].

d'urto per scattering dei fotoni, rendendo l'Universo trasparente. La radiazione emessa in questa fase è osservata oggi come CMB.

## 1.5 Materia ed energia oscura

Come già menzionato, le osservazioni cosmologiche indicano che la dinamica dell'Universo non è descrivibile considerando unicamente materia barionica e radiazione. Evidenze indipendenti — come ad esempio: dinamica degli ammassi [Zwicky, 1933], curve di rotazione galattiche [Rubin et al., 1980], supernovae di tipo Ia [Riess et al., 1998, Perlmutter et al., 1999] e *lensing* gravitazionale [Refregier, 2003, Massey et al., 2010] — richiedono l'introduzione di componenti aggiuntive non direttamente osservabili, note come materia oscura ed energia oscura.

## Abbondanze cosmologiche attuali

Misure della radiazione cosmica di fondo, come ad esempio quelle della missione Planck Collaboration et al. [2020b], permettono di determinare con precisione i parametri di densità al tempo presente:

$$\Omega_{\Lambda,0} \simeq 0.684, \quad \Omega_{\text{DM},0} \simeq 0.265, \quad \Omega_{\text{b},0} \simeq 0.049, \quad \Omega_{\text{r},0} \simeq 9 \times 10^{-5}. \quad (1.31)$$

Questi valori mostrano che l'Universo attuale è dominato dall'energia oscura, mentre la materia è prevalentemente CDM. La componente barionica costituisce solo una frazione minoritaria del contenuto totale, mentre la radiazione è oggi trascurabile dal punto di vista dinamico.

## Materia oscura

La materia oscura è una componente non barionica e non relativistica, che interagisce solamente tramite gravità.

Nel modello  $\Lambda$ CDM essa è descritta come un fluido a pressione nulla:

$$w_{\text{DM}} \simeq 0, \quad (1.32)$$

da cui segue una diminuzione puramente geometrica con l'espansione:

$$\rho_{\text{DM}}(a) \propto a^{-3}. \quad (1.33)$$

Questa dipendenza riflette la conservazione del numero di particelle in un volume comovente in espansione.

## Energia oscura

Nel modello  $\Lambda$ CDM, l'energia oscura è descritta dalla costante cosmologica, che può essere interpretata come una componente con pressione negativa:

$$w_{\Lambda} = -1, \quad (1.34)$$

$$\rho_{\Lambda}(a) = \text{cost.} \quad (1.35)$$

A differenza delle altre componenti, la sua densità non si diluisce con l'espansione. Di conseguenza, essa diventa dominante a bassi redshift ( $z \lesssim 1$ ), determinando una fase di espansione accelerata.

## Capitolo 2

# Formazione delle strutture cosmiche

Nel Capitolo precedente si è discusso come il modello  $\Lambda$ CDM descriva un Universo primordiale estremamente omogeneo e isotropo, caratterizzato da piccole fluttuazioni della densità e della temperatura, osservabili oggi nelle anisotropie della CMB, con ampiezza relativa dell'ordine di  $10^{-5}$ .

L'Universo osservato alle epoche attuali presenta invece una distribuzione fortemente strutturata della materia: galassie, gruppi e ammassi di galassie, filamenti e vaste regioni sottodense si organizzano in una rete complessa comunemente indicata come *cosmic web*. Comprendere come essa sia emersa a partire da condizioni iniziali quasi uniformi costituisce uno dei principali successi della cosmologia moderna.

Il meccanismo fisico alla base di questa evoluzione è l'instabilità gravitazionale. Piccole sovradensità iniziali esercitano infatti una maggiore attrazione gravitazionale sulle regioni circostanti, rallentandone l'espansione e richiamando progressivamente materia. Parallelamente, le regioni inizialmente meno dense tendono a svuotarsi ulteriormente, evolvendo verso le strutture oggi note come vuoti cosmici.

In questo capitolo verranno presentati i principali fondamenti teorici della formazione delle strutture cosmiche. In particolare si introdurrà la teoria lineare delle perturbazioni, si discuterà la crescita gravitazionale delle fluttuazioni, si presenteranno gli strumenti statistici utilizzati per descrivere il campo di densità e si accennerà al regime non lineare e al ruolo delle simulazioni numeriche.

## 2.1 Teoria lineare delle perturbazioni

### 2.1.1 Contrasto di densità

Per descrivere quantitativamente la non omogeneità della distribuzione di materia si introduce il contrasto di densità:

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad (2.1)$$

dove  $\rho(\vec{x}, t)$  rappresenta la densità locale di materia e  $\bar{\rho}(t)$  la densità media cosmica.

Per definizione, valori  $\delta > 0$  corrispondono a sovradensità, mentre  $\delta < 0$  descrivono sottodensità. Nelle prime fasi evolutive dell'Universo vale la condizione

$$|\delta(\vec{x})| \ll 1 \quad \forall \vec{x}, \quad (2.2)$$

che consente di trattare le fluttuazioni come piccole perturbazioni rispetto a un fondo omogeneo e isotropo.

### 2.1.2 Descrizione fluidodinamica

Su scale inferiori all'orizzonte cosmologico, l'evoluzione della materia può essere descritta in buona approssimazione mediante la fluidodinamica quasi-newtoniana. Chiamando  $\vec{x}$  le coordinate comoventi e  $\vec{v}(\vec{x}, t)$  la velocità peculiare, le equazioni fondamentali per un Universo in espansione sono:

- Equazione di continuità

$$\frac{\partial \rho(\vec{x}, t)}{\partial t} + 3H(t)\rho(\vec{x}, t) + \frac{1}{a(t)} \vec{\nabla}_{\vec{x}} \cdot (\rho(\vec{x}, t)\vec{v}(\vec{x}, t)) = 0, \quad (2.3)$$

che esprime la conservazione della massa;

- Equazione di Eulero

$$\frac{\partial \vec{v}(\vec{x}, t)}{\partial t} + H(t)\vec{v}(\vec{x}, t) + \frac{1}{a(t)}(\vec{v}(\vec{x}, t) \cdot \vec{\nabla}_{\vec{x}})\vec{v}(\vec{x}, t) = -\frac{1}{a(t)} \frac{\vec{\nabla}_{\vec{x}} P(\vec{x}, t)}{\rho(\vec{x}, t)} - \frac{1}{a(t)} \vec{\nabla}_{\vec{x}} \Phi(\vec{x}, t), \quad (2.4)$$

che descrive la conservazione della quantità di moto e in cui  $P(\vec{x}, t)$  è la pressione;

- Equazione di Poisson

$$\nabla^2 \Phi(\vec{x}, t) = 4\pi G a^2(t) \bar{\rho}(t) \delta(\vec{x}, t), \quad (2.5)$$

che lega il potenziale gravitazionale  $\Phi(\vec{x}, t)$  alla densità di massa.

Completano il sistema: un'equazione di stato termodinamica che legghi pressione e densità (Eq. 1.19) e le equazioni di Friedmann (Eq. 1.21, Eq. 1.22) che descrivono l'espansione del fondo cosmologico  $a(t)$ . D'ora in avanti, per semplicità, si ometteranno gli argomenti spaziali e temporali delle grandezze fisiche.

### 2.1.3 Instabilità di Jeans

Per comprendere il comportamento delle perturbazioni è utile considerare inizialmente un mezzo statico e uniforme. In tale caso la competizione tra gravità e pressione introduce una scala caratteristica detta lunghezza di Jeans:

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\bar{\rho}}}, \quad (2.6)$$

dove  $c_s = \sqrt{\partial P / \partial \rho}$  è la velocità del suono nel fluido.

Chiamando  $\lambda$  la lunghezza d'onda della perturbazione presa in considerazione, il confronto con la lunghezza di Jeans fornisce:

- per  $\lambda < \lambda_J$ , la pressione è sufficientemente efficace da contrastare la gravità e la perturbazione oscilla come un'onda acustica;
- per  $\lambda > \lambda_J$ , la gravità prevale e la perturbazione cresce nel tempo.

L'instabilità di Jeans costituisce il meccanismo elementare della formazione gravitazionale delle strutture.

Nel caso realistico di un Universo in espansione, l'evoluzione delle perturbazioni viene studiata utilizzando coordinate comoventi e linearizzando le equazioni precedenti, nel regime di validità della condizione (2.2).

Ogni modo di Fourier  $\delta_k$  del contrasto di densità soddisfa allora:

$$\ddot{\delta}_k + 2H_0 \dot{\delta}_k + \left( \frac{k^2 c_s^2}{a^2(t)} - 4\pi G \bar{\rho} \right) \delta_k = 0, \quad (2.7)$$

dove  $k$  è il modulo del vettore d'onda comovente.

Il termine di derivata prima rispetto al tempo  $2H\dot{\delta}_k$  ha natura dissipativa e rappresenta l'effetto frenante dell'espansione cosmica, spesso chiamato *attrito di Hubble*. L'equazione mostra che la crescita delle perturbazioni in un Universo dinamico risulta più lenta rispetto al caso statico.

## 2.2 Crescita delle perturbazioni

Per scale sufficientemente grandi, tali che il termine di pressione sia trascurabile, l'equazione (2.7) si riduce a:

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\bar{\rho}\delta = 0. \quad (2.8)$$

Essa ammette in generale due soluzioni indipendenti:

- un modo crescente  $\delta_+(t)$ , responsabile della formazione delle strutture;
- un modo decrescente  $\delta_-(t)$ , che si attenua nel tempo.

Nel modello di Universo dominato dalla materia ( $\Omega_m = 1$ ,  $\Omega_\Lambda = 0$ ), detto di Einstein-De Sitter, si ottiene:

$$\delta_+(t) \propto a(t) \propto t^{2/3}, \quad \delta_-(t) \propto t^{-1}. \quad (2.9)$$

Poiché il modo decrescente scompare rapidamente, l'evoluzione cosmologica è dominata dalla soluzione crescente.

### 2.2.1 Fattore di crescita

In modelli cosmologici con materia e costante cosmologica, la crescita non è semplicemente proporzionale al fattore di scala in tutte le epoche. Si introduce quindi il fattore di crescita lineare:

$$D(z) = \frac{\delta(z)}{\delta(0)}, \quad (2.10)$$

con  $\delta(0)$  contrasto di densità calcolato a tempi odierni, dunque a  $z = 0$ .

Esso quantifica l'evoluzione temporale delle perturbazioni lineari. Nel caso Einstein-de Sitter vale semplicemente:

$$D(z) = \frac{1}{1+z}. \quad (2.11)$$

Nel modello  $\Lambda$ CDM, la presenza dell'energia oscura rallenta la crescita alle epoche cosmiche recenti.

Si definisce inoltre il parametro

$$f = \frac{d \ln D(z)}{d \ln a(z)}, \quad (2.12)$$

che misura il tasso istantaneo di crescita delle perturbazioni. Una buona approssimazione nel modello  $\Lambda$ CDM è:

$$f \simeq \Omega_m^\gamma, \quad \gamma \simeq 0.55. \quad (2.13)$$

La misura osservativa di  $f$  costituisce un importante test della gravità su scale cosmologiche come indicato ad esempio da Linder [2005].

## 2.3 Descrizione statistica del campo di densità

A causa della natura quantistica delle condizioni primordiali, la distribuzione cosmica della materia non è predicibile in modo deterministico, quindi la sua descrizione richiede strumenti statistici.

### 2.3.1 Funzione di correlazione a due punti

La funzione di correlazione a due punti,  $\xi(r)$ , misura l'eccesso di probabilità di trovare due traccianti del campo densità, ad esempio galassie o ammassi di galassie, separati da distanza comovente  $r$  rispetto a una distribuzione casuale uniforme:

$$dP = n^2[1 + \xi(r)]dV_1dV_2, \quad (2.14)$$

dove  $n$  è la densità numerica media comovente e  $dV_1$ ,  $dV_2$  sono due volumi comoventi infinitesimi.

Se  $\xi(r) > 0$ , gli oggetti risultano correlati positivamente alla scala  $r$  e dunque con tendenza ad addensarsi; se  $\xi(r) < 0$  risultano anticorrelati alla medesima scala, quindi con tendenza ad allontanarsi; se infine  $\xi(r) = 0$ , la distribuzione è casuale.

### 2.3.2 Spettro di potenza e parametro $\sigma_8$

Lo spettro di potenza  $P(k)$  è la trasformata di Fourier della funzione di correlazione:

$$P(k) = \frac{1}{(2\pi)^3} \int \xi(r) e^{i\vec{k} \cdot \vec{r}} d\vec{r}. \quad (2.15)$$

Esso descrive la distribuzione della potenza delle fluttuazioni in funzione della scala spaziale: valori piccoli di  $k$  corrispondono a grandi scale, mentre grandi  $k$  corrispondono a piccole scale.

Secondo la teoria inflazionaria standard, le perturbazioni primordiali seguono una distribuzione casuale generata da fluttuazioni di un campo scalare, detto inflatone, e hanno ampiezze descritte da una gaussiana. Ne segue che lo spettro di potenza primordiale è parametrizzato solo da uno spettro di potenza, che nello scenario standard è descritto da una legge di potenza:

$$P_{\text{prim}}(k) \propto k^n, \quad (2.16)$$

dove  $n \simeq 1$  [Zeldovich, 1972]. Misure odierne della Planck Collaboration et al. [2020a] indicano  $n_{\text{oss}} \simeq 0.965$ , dunque l'invarianza di scala risulta ragionevole.

Una quantità legata allo spettro di potenza e largamente utilizzata è il parametro  $\sigma_8$ , definito come la deviazione standard del campo di densità filtrato su sfere di raggio  $8 h^{-1} \text{Mpc}$ :

$$\sigma_8^2 = \frac{1}{2\pi^2} \int_0^\infty P(k) k^2 W^2(kR) dk. \quad (2.17)$$

Esso è un parametro di scala e riferimento utile, poiché misura l'ampiezza delle fluttuazioni della materia su range tipici delle strutture locali.

## 2.4 Regime non lineare

### 2.4.1 Limiti della teoria lineare

La teoria lineare delle perturbazioni descrive correttamente l'evoluzione delle fluttuazioni fintanto che il contrasto di densità rimane molto piccolo:  $|\delta(\vec{x}, t)| \ll 1, \forall \vec{x}$ . In questo regime, le diverse scale evolvono indipendentemente e le equazioni possono essere linearizzate. Tuttavia, con il passare del tempo, le perturbazioni gravitazionali crescono fino a raggiungere valori dell'ordine dell'unità:  $|\delta| \sim 1$ . A questo punto i termini non lineari diventano importanti e la dinamica evolve verso la formazione di strutture collassate.

Nel regime non lineare non esistono in generale soluzioni analitiche per l'evoluzione delle perturbazioni (Eq. 2.8), ma alcuni modelli semplificati risultano estremamente utili per descrivere l'evoluzione gravitazionale delle strutture cosmiche.

### 2.4.2 Collasso sferico

Uno dei modelli analitici per descrivere l'evoluzione non lineare delle sovradensità è il modello del collasso sferico.

Si considera una regione sferica inizialmente più densa del fondo cosmologico, l'evoluzione della perturbazione attraversa quattro fasi principali:

1. Espansione rallentata: la sovradensità continua inizialmente a espandersi insieme all'Universo, ma con velocità inferiore rispetto al fondo cosmologico a causa della maggiore attrazione gravitazionale locale.
2. *Turn-around point*: la regione raggiunge un raggio massimo e la sua espansione si arresta.
3. Collasso gravitazionale: la sovradensità inizia a contrarsi sotto l'azione della gravità.
4. Virializzazione: l'energia potenziale gravitazionale viene convertita in energia cinetica dei moti peculiari interni e il sistema raggiunge uno stato quasi stazionario governato dal teorema del viriale.

Il modello del collasso sferico costituisce una modellizzazione della formazione degli aloni di materia oscura e permette di stimare quantità importanti come la densità critica di collasso.

### 2.4.3 Approssimazione di Zel'dovich

Un approccio più dettagliato all'evoluzione non lineare è fornito dall'approssimazione di Zel'dovich [Zel'dovich, 1970], largamente utilizzata nello studio della formazione delle strutture cosmiche.

A differenza della teoria lineare standard, formulata in coordinate euleriane, l'approssimazione di Zel'dovich adotta un approccio lagrangiano: in questo formalismo si segue direttamente il moto degli elementi di fluido a partire dalla loro posizione iniziale.

La posizione comovente di una particella può essere scritta come

$$\vec{x}(\vec{q}, t) = \vec{q} + D(t) \vec{\psi}(\vec{q}), \quad (2.18)$$

dove:

- $\vec{q}$  rappresenta la posizione lagrangiana iniziale;
- $\vec{x}$  la posizione euleriana finale, dunque quella osservata;
- $D(t)$  è il fattore di crescita lineare (2.10);
- $\vec{\psi}(\vec{q})$  è il campo di spostamento iniziale.

Si assume inoltre che il campo  $\vec{\psi}(\vec{q})$  sia descritto dal gradiente di un campo scalare  $\phi(\vec{q})$ , secondo:

$$\vec{\psi}(\vec{q}) = -\vec{\nabla} \phi(\vec{q}). \quad (2.19)$$

Considerando poi che l'equazione di Poisson (2.5), induce per il contrasto di densità la seguente forma:

$$\delta = -D \nabla^2 \phi, \quad (2.20)$$

allora è possibile giungere alla relazione:

$$\vec{\nabla} \cdot \vec{\psi} = -\delta. \quad (2.21)$$

Evidenze a supporto di essa possono essere trovate ad esempio nel lavoro di Falck et al. [2012].

Questo legame fra la divergenza del campo di spostamento e il contrasto di densità è lo strumento alla base del meccanismo di ricostruzione del campo di densità, operato dall'algoritmo su cui la tesi si incentra.

#### 2.4.4 Evoluzione delle sottodensità

Le regioni sottodense, comunemente chiamate vuoti cosmici, seguono un'evoluzione speculare rispetto alle sovradensità. In presenza di una sottodensità locale di materia, l'attrazione gravitazionale risulta minore rispetto al fondo cosmologico e la regione tende quindi a espandersi più rapidamente dell'Universo circostante.

Con l'evoluzione cosmica, tali regioni diventano progressivamente meno dense e assumono forme mediamente più sferiche. I moti gravitazionali trasportano la materia verso le regioni periferiche, contribuendo alla formazione della struttura a filamenti e pareti che delimita i vuoti cosmici.

La formazione dinamica dei vuoti è caratterizzata dal fenomeno di *shell crossing*, nel quale gusci interni di materia raggiungono velocità maggiori rispetto ai gusci esterni e finiscono per intersecarsi.

Poiché i vuoti si trovano in un regime moderatamente non lineare, essi risultano particolarmente interessanti dal punto di vista cosmologico; difatti le loro proprietà statistiche sono molto sensibili ai valori dei parametri cosmologici, tra cui quelli che descrivono l'energia oscura, e ad eventuali modifiche della teoria della relatività generale, discussa nella sezione 1.1.

#### 2.4.5 Bias di aloni oscuri e traccianti luminosi

Dal punto di vista osservativo, la distribuzione della materia (quasi totalmente oscura, dunque non interagente elettromagneticamente) non viene misurata in modo diretto, ma attraverso traccianti luminosi come galassie e ammassi.

Tali oggetti si formano preferenzialmente nelle regioni più dense del campo di materia oscura e risultano quindi distribuiti in modo diverso rispetto alla materia totale. Questo effetto prende il nome di *bias*.

Nel regime lineare, il contrasto di densità dei traccianti,  $\delta_{\text{tr}}$ , può essere messo in relazione con quello della materia,  $\delta_{\text{m}}$  come:

$$\delta_{\text{tr}} = b \delta_{\text{m}}, \quad (2.22)$$

dove  $b$  è detto parametro di bias, che può essere derivato tramite il rapporto fra le funzioni di correlazione a due punti (2.14) dei traccianti e della materia come mostrato da Kaiser [1984].

L'esistenza del bias osservativo ha conseguenze importanti nell'identificazione dei vuoti cosmici, poiché gli algoritmi di *void finding* operano generalmente su cataloghi osservativi di galassie piuttosto che sulla distribuzione sottostante della materia oscura.

## 2.5 Simulazioni cosmologiche N-corpi

Come già discusso, il regime non lineare non ammette soluzioni analitiche per l'equazione evolutiva delle perturbazioni, quindi lo studio della dinamica cosmologica si basa in larga misura su simulazioni numeriche a N-corpi.

In tali simulazioni, la materia oscura è rappresentata come un insieme discreto di particelle che campionano il campo di densità cosmico e che evolvono sotto l'azione della gravità in un Universo in espansione. L'evoluzione dinamica è descritta in coordinate comoventi dalle equazioni del moto:

$$\frac{d^2 \vec{x}_i}{dt^2} + 2H \frac{d\vec{x}_i}{dt} = -\frac{1}{a^2} \nabla \Phi, \quad (2.23)$$

in cui il termine  $2H \dot{\vec{x}}_i$  descrive l'effetto di attrito di Hubble, già discusso a riguardo dell'Eq. 2.8, mentre il potenziale gravitazionale  $\Phi$  è determinato dalla distribuzione di materia tramite l'equazione di Poisson (Eq. 2.5).

Dal punto di vista fisico, queste equazioni rappresentano la versione discreta del sistema fluido gravitazionale introdotto nelle sezioni precedenti, esteso al regime non lineare.

La soluzione numerica del problema a molti corpi richiede l'impiego di algoritmi approssimati per il calcolo delle forze di interazione gravitazionale fra particelle. Nei metodi Particle-Particle (PP) l'interazione viene calcolata direttamente tra tutte le coppie di particelle, con elevata accuratezza ma grande costo computazionale. Nei metodi Particle

Mesh (PM) le grandezze dinamiche vengono proiettate su una griglia e risolte in spazio di Fourier, in cui l'equazione di Poisson diviene algebrica. Questo metodo riduce il peso computazionale ma introduce un limite di risoluzione legato alla scala di *smoothing* utilizzata. Infine, i metodi Tree code combinano le due modalità precedenti: decomponendo lo spazio in regioni che contengono al più una particella e approssimando regioni periferiche come masse equivalenti, ottenendo una scalabilità tipica  $O(N \log N)$ .

Queste simulazioni rappresentano uno strumento fondamentale per il confronto tra teoria e osservazioni, in particolare per la calibrazione di funzioni statistiche quali la funzione di correlazione a due punti (Eq. 2.14) e lo spettro di potenza della materia (Eq. 2.15).

A titolo di esempio, la Fig. 2.1 illustra la distribuzione della materia oscura nella simulazione Uchuu [Ishiyama et al., 2021]. Già dal pannello superiore si riconosce la tipica struttura a ragnatela cosmica, con ammassi densi, filamenti e vuoti macroscopici. Scendendo verso i pannelli di scala via via più piccola, emergono progressivamente i singoli aloni di materia oscura e la loro struttura interna. Questa sequenza di ingrandimenti mette in evidenza la natura gerarchica del processo di formazione delle strutture, mostrando come le simulazioni N-corpi siano in grado di risolvere contemporaneamente scale cosmologiche e sub-galattiche, consentendo così un confronto diretto con le osservazioni su un ampio intervallo di scale.

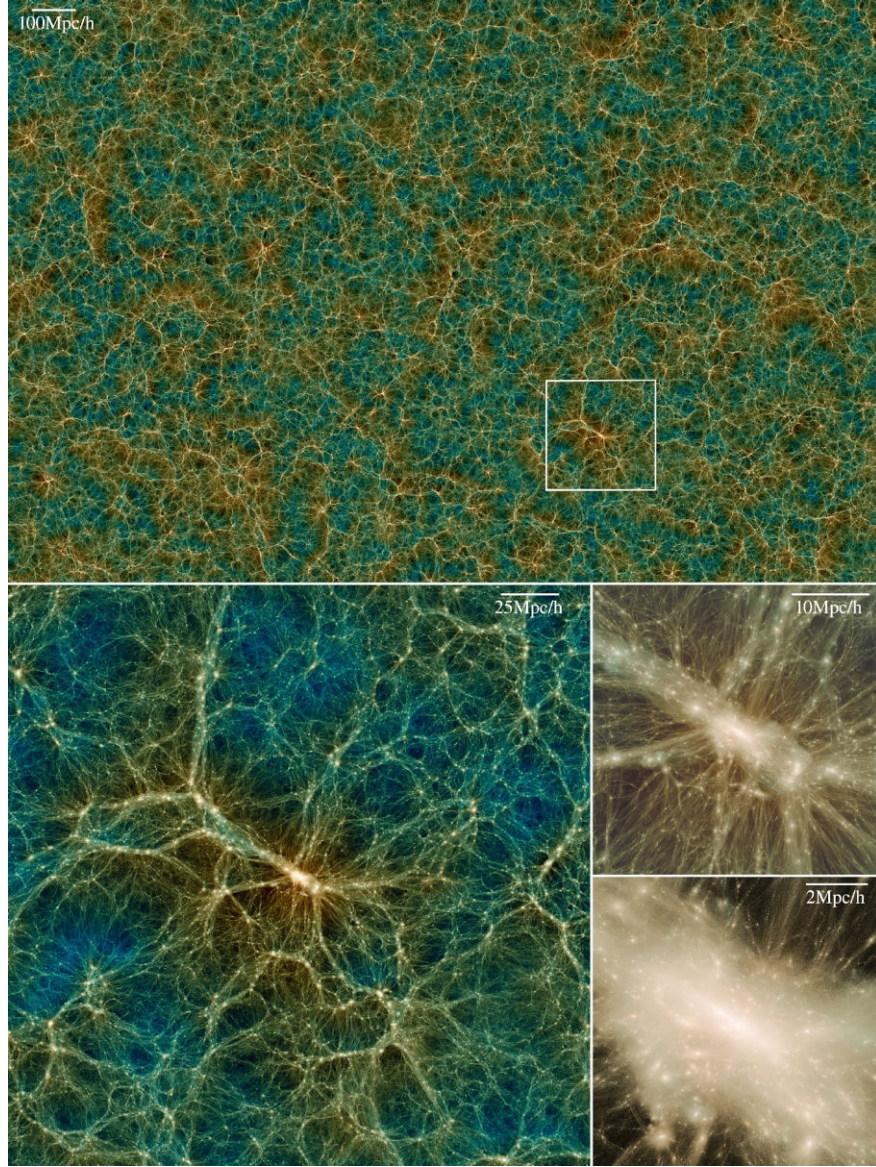


Figura 2.1: Distribuzione delle particelle di materia oscura nella simulazione Uchuu [Ishiyama et al., 2021] (volume complessivo  $2 h^{-1}\text{Gpc}$ ). Il pannello superiore mostra una regione di lato  $100 h^{-1}\text{Mpc}$ ; i pannelli inferiori mostrano ingrandimenti successivi alle scale di 25, 10 e  $2 h^{-1}\text{Mpc}$  (da sinistra a destra).

## Capitolo 3

# Back-in-time Void Finder

L'identificazione dei vuoti cosmici rappresenta il primo passo fondamentale per il loro utilizzo come sonde cosmologiche. Come discusso nel Capitolo 2, i vuoti sono strutture sottodense che occupano la maggior parte del volume dell'Universo e la loro evoluzione è determinata dall'instabilità gravitazionale a partire da piccole fluttuazioni primordiali. Tuttavia, la loro natura sotto-densa e la mancanza di una definizione univoca rendono la loro identificazione un problema complesso, che ha portato allo sviluppo di numerosi algoritmi basati su criteri diversi.

In questo capitolo verranno descritti i principali strumenti per l'identificazione di vuoti cosmici, classificando i metodi esistenti in tre categorie principali: metodi basati sulla densità, sulla geometria e sulla dinamica. Per ciascuna classe verranno illustrati i principi fondamentali, i vantaggi e le limitazioni, con particolare attenzione agli algoritmi a cui si farà riferimento nel seguito della tesi, come **VIDE** per i metodi geometrici e **BitVF** per quelli dinamici.

Successivamente, verranno introdotte le principali proprietà statistiche dei vuoti cosmici, ovvero la VSF e i profili di densità, soffermandosi su alcuni modelli teorici di riferimento e sul loro utilizzo per vincolare parametri cosmologici quali  $\Omega_m$  e  $\sigma_8$ .

Infine, si descriverà nel dettaglio l'algoritmo di ricerca dinamica **BitVF**. Verranno esposte le tre fasi principali dell'algoritmo, la procedura di *cleaning* necessaria per rendere i cataloghi confrontabili con i modelli teorici e l'implementazione all'interno delle librerie **CosmoBolognaLib** [Marulli et al., 2016–2026], disponibili pubblicamente su GitLab, utilizzate per le analisi presentate nei capitoli successivi.

### 3.1 Generalità sulla procedura di *void finding*

Per poter utilizzare i vuoti cosmici come sonde cosmologiche è necessario dapprima saperli individuare in modo univoco e robusto, attraverso l'utilizzo di traccianti. Tuttavia, non esiste al momento una definizione universalmente accettata di vuoto cosmico, problema messo in evidenza ad esempio da Colberg et al. [2008]; invece esiste una grande varietà di algoritmi per l'identificazione di queste strutture, basati su criteri e definizioni diverse. Seguendo la classificazione proposta da Lavaux and Wandelt [2009], i metodi si dividono in tre classi principali:

- Metodi basati sulla densità: identificano i vuoti come regioni sferiche (o di altra forma geometrica) in cui la densità dei traccianti scenda al di sotto di una certa soglia, tipicamente il 20% o 30% della densità media dell'Universo [Kauffmann and Fairall, 1991]. Questi metodi sono semplici da implementare ma soffrono del problema dello *shot noise*: il limitato numero di traccianti in zone sotto-dense comporta fluttuazioni statistiche difficili da controllare.
- Metodi basati sulla geometria: ricostruiscono il campo di densità attraverso tecniche di partizione dello spazio, come la tassellazione di Voronoi [Neyrinck, 2008]. I vuoti vengono poi identificati come insiemi di celle contigue a bassa densità. L'algoritmo più noto di questa classe è **VIDE** [Sutter et al., 2014], che sfrutta una procedura di *watershed* per delimitare i bacini di sottodensità. Questo metodo, implementato da Platen et al. [2007], può essere compreso con l'aiuto della Fig. 3.1: si immagina il campo di densità come una superficie, nella quale i minimi locali (prima immagine a sinistra) rappresentino i centri dei vuoti. Figurativamente, si inizia a riempire la superficie di "acqua" a partire dai punti di minimo e innalzando gradualmente il livello. Man mano che si riempiono d'acqua (immagine centrale), i bacini si espandono fino a incontrare i punti di massima altitudine (le creste), che separano nettamente un bacino dall'altro (immagine a destra). In questa analogia, i vuoti corrispondono ai bacini, mentre i filamenti e le pareti cosmiche corrispondono ai confini. Nel proseguo della tesi, verrà mostrato il risultato dell'analisi comparativa tra un'esecuzione di **VIDE** e una di **BitVF**. Questi metodi sono più sofisticati di quelli basati sulla sola densità, ma rimangono sensibili alla discretizzazione del campo di traccianti.

- Metodi basati sulla dinamica: non campionano direttamente il campo di densità, ma utilizzano i traccianti per ricostruire il campo di velocità [Forero-Romero et al., 2008, Lavaux and Wandelt, 2009]. I vuoti vengono poi identificati come regioni in cui i traccianti tendono a convergere quando si osserva il loro moto indietro nel tempo. Questa classe di metodi, a cui appartiene **BitVF**, risulta meno affetta da *shot noise* rispetto ai metodi precedentemente presentati e consente di ottenere campi di densità continui e regolari. La Fig. 3.2 illustra il principio alla base di questi metodi, mostrando i risultati ottenuti da Elyiv et al. [2015] con due diversi algoritmi dinamici. Le frecce nel pannello superiore rappresentano il campo di spostamento medio, ricostruito collegando le posizioni attuali dei traccianti (punti neri) con le loro posizioni lagrangiane (iniziali). Si osserva come le frecce puntino chiaramente verso le regioni di minima densità, indicando che i traccianti, se proiettati indietro nel tempo, tendono a convergere verso i centri dei vuoti. Nel pannello inferiore, il campo di divergenza (rappresentato dai colori) riproduce fedelmente il campo di densità: le regioni blu (divergenza negativa) corrispondono a sottodensità (vuoti), mentre le regioni rosse (divergenza positiva) corrispondono a sovradensità.

Un problema comune ai metodi basati sulla densità e sulla geometria è la loro sensibilità agli effetti delle RSD, ovvero effetti osservativi che alterano la posizione apparente dei traccianti lungo la linea di vista: quando si convertono i redshift osservati in distanze comoventi, il contributo delle velocità peculiari delle galassie si somma a quello dell'espansione cosmica, introducendo uno spostamento sistematico delle posizioni misurate. Questo effetto, difficilmente modellabile in modo accurato, introduce una componente di errore aggiuntiva nell'identificazione dei vuoti, poiché la distribuzione dei traccianti nello spazio dei redshift risulta distorto rispetto a quella reale.

Le RSD rappresentano per i vuoti un problema particolarmente complesso, poiché non si limitano a distorcere le statistiche, ma alterano il catalogo stesso dei vuoti, modificandone numero, posizioni e dimensioni. Per questo motivo, nel contesto di questa tesi le RSD sono considerate principalmente un effetto sistematico da mitigare, e l'obiettivo è valutare quanto i cataloghi di vuoti nello spazio dei redshift si discostino da quelli di riferimento nello spazio reale.

A questo proposito, i metodi dinamici, come **BitVF**, offrono un approccio alternativo per mitigare questi effetti. Invece di operare direttamente sul campo di densità osservato, essi ricostruiscono il campo di spostamento dei traccianti, ovvero la mappa che collega le

posizioni osservate nello spazio dei redshift alle corrispondenti posizioni che i traccianti avrebbero in assenza di velocità peculiari, dunque in spazio reale. Questa ricostruzione non fornisce le posizioni "vere" dei singoli traccianti, ma una stima statistica del campo di densità reale, meno affetta dallo *shot noise* e più robusta rispetto alle distorsioni indotte dalle RSD. Questo approccio permette di ottenere una stima del campo di densità reale meno affetta dallo *shot noise* e più robusta rispetto alle distorsioni indotte dalle RSD. Di conseguenza, si vuole investigare se BitVF possa produrre un campione di vuoti osservato più fedele a quello che si avrebbe in spazio reale, mitigando le sistematiche introdotte dagli effetti osservativi e migliorando l'affidabilità delle analisi cosmologiche basate sui vuoti. Questo aspetto verrà approfondito nel Capitolo 5, dove le metriche di accoppiamento e mappatura dei volumi mostreranno quantitativamente la maggiore stabilità di BitVF nel tracciare i vuoti tra spazio reale e spazio dei redshift rispetto a VIDE.

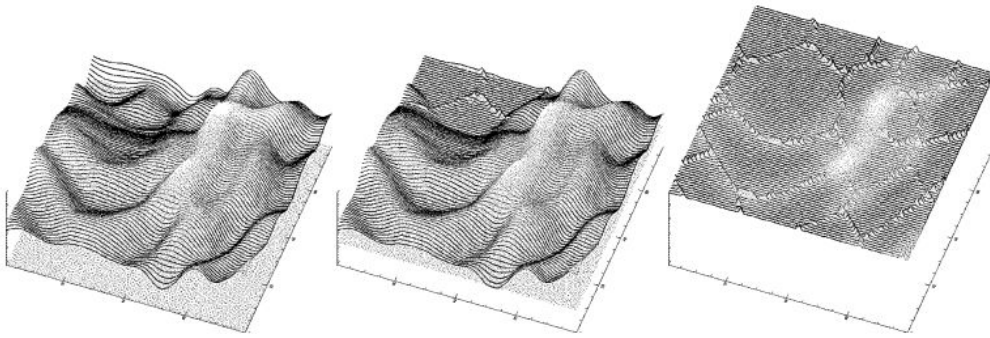


Figura 3.1: Illustrazione figurativa del principio del *watershed*. A sinistra: superficie da segmentare. Al centro: allagamento dai minimi locali; al confine tra bacini viene tracciata una demarcazione. A destra: rete di demarcazioni che definisce i volumi segmentati. Crediti: Platen et al. [2007].

## 3.2 Proprietà statistiche dei vuoti cosmici

Una volta identificato un catalogo di vuoti tramite un algoritmo, si possono estrarre informazioni cosmologiche attraverso due statistiche principali: la VSF e il profilo di densità. L'importanza dello studio di esse deriva dal fatto che sono ottimi strumenti per il vincolo di parametri cosmologici come  $\Omega_m$  e  $\sigma_8$  [Lehman et al., 2025].

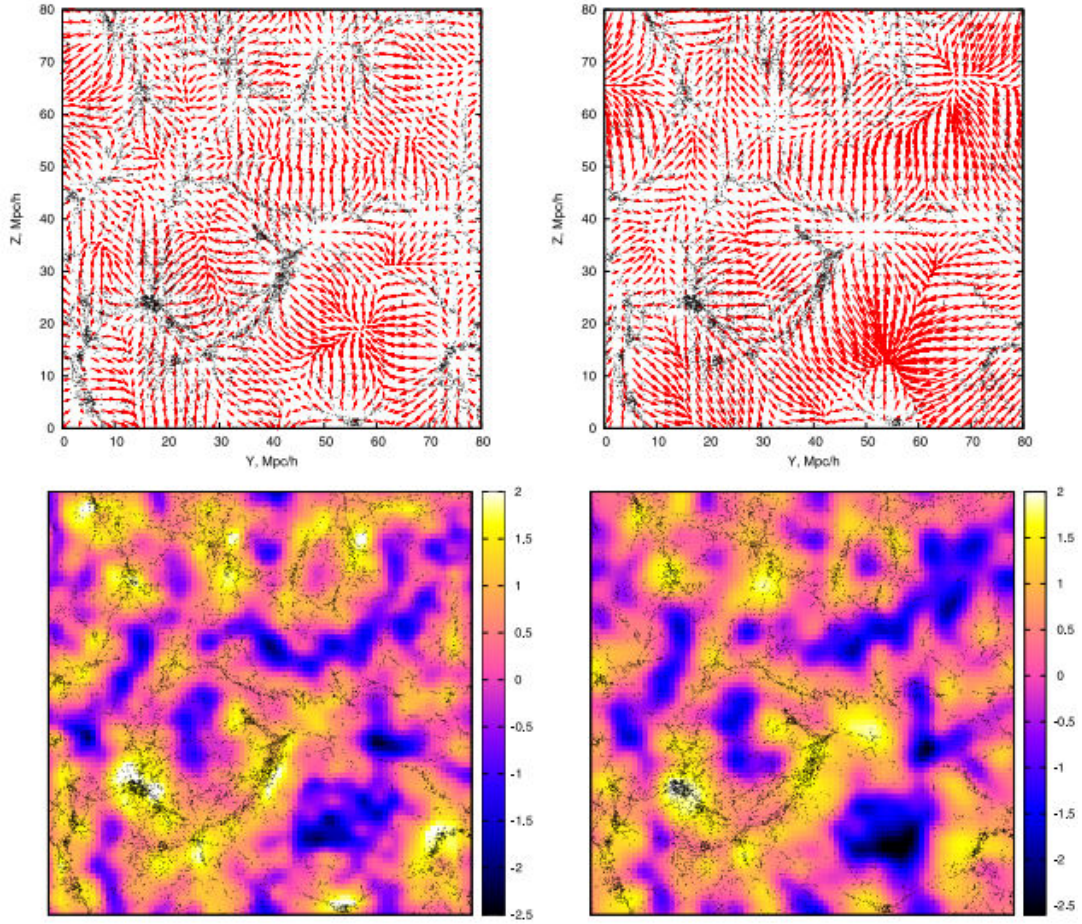


Figura 3.2: Ricostruzione del campo di spostamento (pannelli superiori) e della sua divergenza (pannelli inferiori) per due algoritmi dinamici: UVF (Uncorrelating Void Finder) a sinistra e LZVF (Lagrangian Zel'dovich Void Finder) a destra. La figura mostra una proiezione sull'asse Y-Z di una parte del volume simulato; gli assi orizzontale e verticale riportano rispettivamente le coordinate Y e Z espresse in  $h^{-1}\text{Mpc}$ . Crediti: Elyiv et al. [2015].

### 3.2.1 Void size function

Si definisce la VSF come la densità numerica comovente di vuoti,  $n$ , in funzione del loro raggio effettivo,  $R_v$ :  $dn/d \ln R_v(R_v, z)$ . Il primo e principale modello teorico per essa è stato proposto da Sheth and van de Weygaert [2004], da cui il nome modello SVdW, basato sul formalismo *excursion set* Bond et al. [1991]. Esso descrive la formazione di strutture cosmiche attraverso l'utilizzo delle proprietà statistiche del campo densità lineare, la quale viene filtrata a diverse scale cosmologiche e la sua varianza è espressa come:

$$\sigma^2(R) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k) |W(k, R)|^2 dk, \quad (3.1)$$

dove  $P(k)$  è lo spettro di potenza (Eq. 2.15) della materia e  $W(k, R)$  è una funzione filtro che smorza ogni fluttuazione di scala inferiore a  $R$ ; al variare della scala dunque, la variazione del contrasto di densità è considerabile come un *random walk*. Un vuoto o equivalentemente un alone si formano quando il *random walk* oltrepassa per la prima volta un valore critico:  $\delta_v \simeq -2.71$  (detta anche soglia di *shell crossing*) per le sotto-densità e  $\delta_c \simeq 1.686$  per le sovradensità.

I quattro processi principali descritti da questo formalismo sono mostrati in Fig. 3.3; in particolare si descrivono le formazioni di: sovradensità all'interno di una sovradensità di scala maggiore (*cloud-in-cloud*), una sovradensità in un vuoto di scala maggiore (*cloud-in-void*), un vuoto in un vuoto di scala maggiore (*void-in-void*), un vuoto all'interno di una sovradensità di scala maggiore (*void-in-cloud*). Quest'ultimo processo indica la possibilità che un vuoto presente all'interno di una sovradensità possa scomparire al seguito del collasso di quest'ultima.

Tuttavia, il modello presenta un'incongruenza, calcolando la frazione di volume totale occupata dai vuoti tramite:

$$\mathcal{F}(R) = \int_R^\infty V(r) \frac{dn}{d \ln r} \frac{dr}{r}, \quad (3.2)$$

si ottengono valori superiori all'unità per raggi piccoli. Ciò avviene poiché il modello richiede la conservazione in numero dei vuoti, i quali però hanno tendenza ad espandersi e sovrapporsi. Per risolvere il problema, Jennings et al. [2013] hanno proposto il modello Vdn, che assume la conservazione del volume totale occupato dai vuoti nel passaggio da regime lineare a non lineare. Questo è il modello ad oggi più utilizzato per il confronto

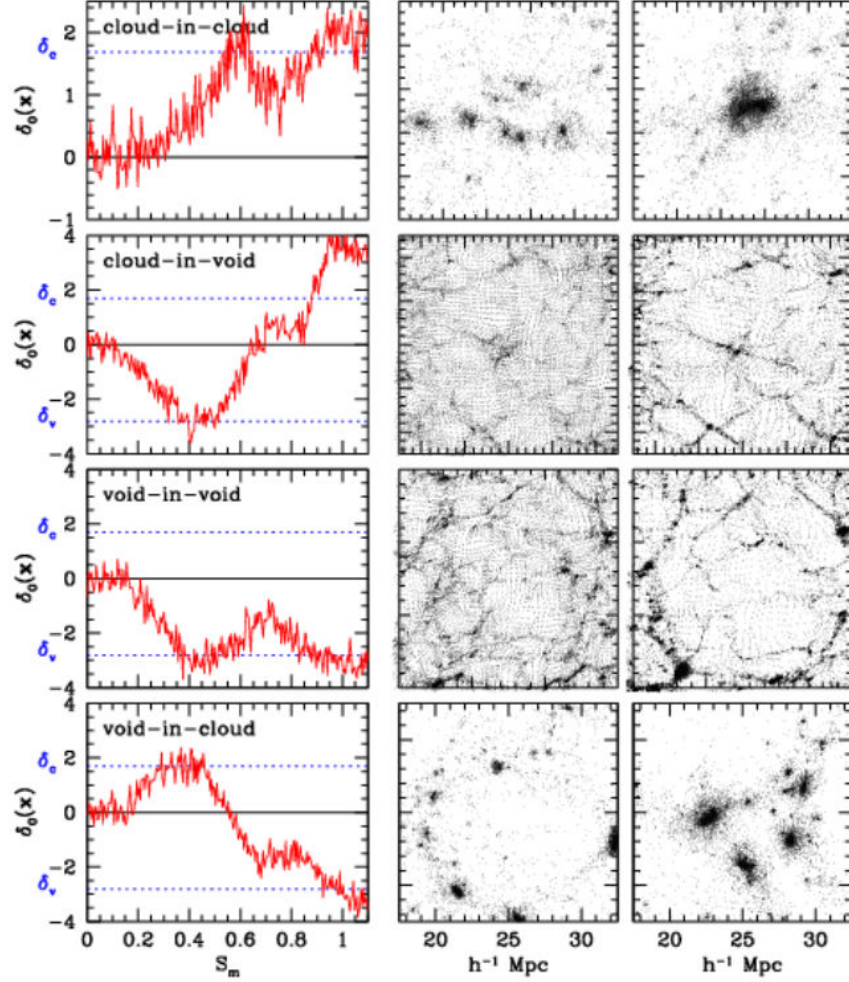


Figura 3.3: Rappresentazione schematica dei quattro processi del formalismo *excursion set*: *cloud-in-cloud*, *cloud-in-void*, *void-in-void* e *void-in-cloud* (dall'alto verso il basso). Per ciascun processo, il pannello di sinistra mostra il cammino casuale nel piano  $(\sigma^2, \delta)$ , dove l'asse orizzontale  $\sigma^2$  rappresenta la varianza del campo di densità filtrato, in unità arbitrarie. I pannelli centrale e destro mostrano l'evoluzione della distribuzione di particelle in due istanti successivi; in entrambi, l'asse orizzontale rappresenta una coordinata spaziale in  $h^{-1}\text{Mpc}$ . L'asse verticale di ogni pannello,  $\delta$ , è il contrasto di densità lineare. Le linee orizzontali tratteggiate nel pannello di sinistra indicano le soglie di collasso ( $\delta_c \simeq 1.686$ ) e di *shell-crossing* ( $\delta_v \simeq -2.71$ ). Crediti: Sheth and van de Weygaert [2004].

con simulazioni e osservazioni.

Per sfruttare il modello con traccianti "biasati" del campo di densità, come aloni e galassie, serve calibrare una relazione di bias [Contarini et al., 2019, 2022]:

$$\delta_{v,\text{DM}} = \frac{\delta_{v,\text{tr}}}{\mathcal{F}(b_{\text{eff}})}, \quad (3.3)$$

in cui  $\delta_{v,\text{DM}}$  è il contrasto di densità da usare per la materia oscura,  $\delta_{v,\text{tr}}$  è la soglia applicata ai traccianti (valore tipico  $-0.7$ ) e  $\mathcal{F}(b_{\text{eff}})$  è una funzione lineare del bias efficace.

### 3.2.2 Profili di densità dei vuoti

Il profilo di densità di un vuoto è definito come il valore medio del contrasto di densità  $\delta(r)$  in gusci sferici concentrici di raggio  $r$ , centrati sul centro del vuoto. Un modello empirico, proposto da Hamaus et al. [2014] molto utilizzato è:

$$\frac{\rho_v(r)}{\bar{\rho}} - 1 = \delta_c \frac{1 - (r/r_s)^\alpha}{1 + (r/r_v)^\beta}, \quad (3.4)$$

in cui  $\delta_c$  è il contrasto di densità centrale,  $r_v$  il raggio del vuoto,  $r_s$  un raggio di scala e  $\alpha, \beta$  le pendenze rispettivamente interna ed esterna del profilo.

Dal punto di vista statistico, il profilo di densità corrisponde al monopolo della funzione di cross-correlazione vuoto-tracciante, ottenuto mediando sfericamente la distribuzione dei traccianti attorno al centro dei vuoti. Tuttavia, i vuoti non sono perfettamente sferici nella loro forma intrinseca, e questo primo effetto è presente anche in spazio reale. A questo si aggiungono numerose distorsioni introdotte dal passaggio dallo spazio dei redshift a quello reale.

La principale fra queste è l'effetto delle distorsioni delle RSD, che si manifestano quando si convertono i redshift osservati in distanze comoventi trascurando il contributo delle velocità peculiari dei traccianti: tali velocità si sommano a quella di espansione cosmica, alterando la posizione apparente dei traccianti lungo la linea di vista e producendo un'elongazione caratteristica dei vuoti in quella direzione. Questo effetto è ben descritto dal quadrupolo del profilo di densità, che misura la deformazione lungo la linea di vista.

Oltre alle RSD, esistono però altre fonti di distorsione. La conversione redshift-distanza richiede l'assunzione di un modello cosmologico e di valori specifici per i parametri cosmologici; errori in queste assunzioni introducono distorsioni geometriche nella ricostruzione delle distanze comoventi. A queste si aggiungono gli errori sui redshift stessi, di natura sia statistica che sistematica, che alterano ulteriormente le posizioni ricostruite. Infine, effetti relativistici minori come il redshift gravitazionale, il redshift trasverso, il *lensing* gravitazionale e gli effetti di *wide-angle* possono contribuire alla distorsione delle mappe, sebbene con un impatto generalmente minore rispetto alle RSD, almeno per cataloghi spettroscopici con errori sui redshift piccoli.

In questa tesi ci si è concentrati in particolare sugli effetti delle RSD, che rappresentano la principale fonte di distorsione per i cataloghi spettroscopici utilizzati in questo lavoro. Per studiare questo fenomeno, si introduce il quadrupolo del profilo di densità, che misura la deformazione del profilo lungo la direzione della linea di vista. Parte del capitolo successivo verterà proprio sull'analisi congiunta di monopolo e quadrupolo per cataloghi di vuoti ottenuti con **BitVF** sia in spazio reale che in spazio dei redshift, confrontando poi i risultati con quelli derivanti da un catalogo trovato da **VIDE** partendo dalla medesima distribuzione di traccianti.

I profili di densità dei cataloghi di vuoti sono fortemente sensibili alla presenza di energia oscura e modifiche alla gravità, rendendoli un'ottima sonda cosmologica.

### 3.3 L'algoritmo BitVF

Il **BitVF** è un algoritmo di ricerca di vuoti cosmici basato su criteri dinamici, sviluppato da Sartori [2022] nella sua tesi magistrale e successivamente esteso a *survey* di galassie e validato da Sartori et al. [2026]. L'idea alla base è utilizzare i traccianti come particelle di prova, ricostruendo le loro traiettorie indietro nel tempo attraverso il principio di minima azione di Hamilton, fino a una configurazione iniziale omogenea e isotropa. Durante il processo, le regioni da cui i traccianti si allontanano sono identificate come vuoti.

L'algoritmo prevede tre fasi principali:

1. Ricostruzione del campo spostamento,  $\vec{\psi}(\vec{q})$ , che collega la posizione iniziale lagrangiana,  $\vec{q}$ , a quella attuale euleriana  $\vec{x}$ .

2. Calcolo del campo densità attraverso la relazione 2.21, che sfrutta l'approssimazione di Zel'dovich, Eq.(2.18).
3. Identificazione dei centri dei vuoti come minimi locali del campo divergenza e assegnazione del raggio effettivo.

Si descriverà ora ogni passo più dettagliatamente.

### 3.3.1 Ricostruzione del campo di spostamento

BitVF implementa due metodi per la ricostruzione del campo di spostamento:

- Ricostruzione esatta: disponendo di simulazioni N-corpi (vedi sezione 2.5) con particelle aventi identificativi univoci, si può ricostruire il campo in maniera esatta collegando la posizione attuale di ciascun tracciante a quella dello *snapshot* iniziale. Per quanto non sia applicabile a dati osservativi ma solo a simulazioni, resta un ottimo modo per validare l'approccio dinamico e statistico.
- Trasporto Ottimo (OT): questo è il metodo utilizzato per cataloghi reali. Sfruttando il principio di minima azione, traccianti e posizioni casuali vengono assegnati fra loro attraverso un procedimento di iterazione che minimizza la distanza totale degli spostamenti [Croft and Gaztañaga, 1997]. Questo approccio si basa sul formalismo del trasporto ottimo, le cui fondamentali matematiche sono state poste da Monge [1781] per il problema del trasporto di massa e successivamente sviluppate da Brenier [1991], che ha dimostrato l'esistenza e l'unicità della mappa di trasporto ottimo per il costo quadratico, e da Brenier [2003] per le generalizzazioni del problema.

Il metodo OT minimizza la seguente funzione costo:

$$S = \sum_{i=1}^N |\vec{x}_i - \vec{q}_{\pi(i)}|^2, \quad (3.5)$$

dove  $\vec{x}_i$  sono le posizioni euleriane (attuali) dei traccianti,  $\vec{q}_{\pi(i)}$  sono le posizioni lagrangiane (iniziali) assegnate tramite la permutazione  $\pi$  che minimizza la somma dei quadrati degli spostamenti. Questa formulazione deriva direttamente dal principio di minima

azione di Hamilton per un sistema auto-gravitante in espansione, nell'approssimazione di particelle di prova e nel limite dell'approssimazione di Zel'Dovich.

Andando più a fondo a riguardo del metodo OT, esso prevede dapprima la creazione di un catalogo di posizioni iniziali casuali, le quali vengono messe in corrispondenza con le posizioni reali dei traccianti. Il criterio affinché l'accoppiamento sia fisicamente corretto è che la distanza sia limitata a 4 volte la separazione media fra particelle (MPS).

Successivamente si selezionano quartetti di coppie (4 traccianti e le loro corrispettive posizioni iniziali) e si valutano le  $4! = 24$  permutazioni ottenibili; viene conservata la configurazione che minimizzi la somma delle distanze e il processo reiterato. La scelta di considerare dei quartetti è un compromesso fra costo computazionale e possibilità di considerare l'intero spazio delle configurazioni.

### 3.3.2 Calcolo del campo di densità

Ricostruito  $\vec{\psi}(\vec{q})$ , si calcola il suo campo di divergenza e si utilizza l'approssimazione di Zel'dovich [1970] per giungere alla relazione per il contrasto di densità, Eq. (2.21). `BitVF` calcola la divergenza su una griglia regolare usando il teorema di Gauss per la divergenza: per ogni cella, il flusso del campo di spostamento attraverso le sei facce viene sommato e diviso per il volume della cella. Se una cella non contiene valori misurati, il suo valore viene interpolato tramite un filtro gaussiano con deviazione standard pari alla dimensione della cella.

### 3.3.3 Identificazione dei vuoti e definizione dei raggi

Come già anticipato, le coordinate dei centri dei vuoti corrispondono ai minimi locali del campo di divergenza,  $\Theta(\vec{q})$ . Una cella è considerata centro di un vuoto se il valore di  $\Theta(\vec{q})$  è negativo e minore del valore calcolato nelle otto celle adiacenti nello spazio tridimensionale. Per aumentare la precisione sulla posizione del centro, viene fatta una media pesata sui valori di divergenza delle celle adiacenti:

$$\vec{x}_{\text{centro}} = \frac{\sum_{j=1}^{27} \Theta_j \vec{x}_j}{\sum_{j=1}^{27} \Theta_j}. \quad (3.6)$$

Una volta identificati i centri, si applica un algoritmo di *watershed* per delimitare i bacini di sottodensità associati a ciascun centro, ovvero le regioni dello spazio in cui la divergenza del campo di spostamento è negativa [Sartori, 2022]. Questo procedimento consente di assegnare a ciascun vuoto una regione di influenza ben definita, a partire dalla quale si definiscono due diverse misure di raggio.

**BitVF** associa infatti ad ogni vuoto due valori di raggio:

- $R_v$ : è il raggio equivalente della sfera avente volume pari al volume del bacino di sottodensità (o "vuoto naturale") identificato dal *watershed*. Questa definizione rappresenta il raggio del vuoto inteso come regione di divergenza negativa e viene utilizzata per caratterizzare la dimensione intrinseca dei vuoti così come identificati dall'algoritmo.
- $R_{\text{ext}}$ : è il raggio della sfera avente volume equivalente al volume del vuoto stesso, che può estendersi anche oltre il bacino naturale per coprire l'intero volume della simulazione o dell'osservazione. Questa definizione è analoga a quella adottata da **VIDE** e risulta più adatta per il confronto diretto tra cataloghi ottenuti con algoritmi diversi e per la misura della cross-correlazione vuoto-tracciante, i cui multipoli (monopolo e quadrupolo) forniscono rispettivamente il profilo di densità dei vuoti e la loro elongazione lungo la linea di vista.

Si noti che il raggio  $R_v$  basato sul contrasto di densità, tipicamente definito imponendo  $\delta_v = -0.7$  per i traccianti, non corrisponde alla definizione interna di **BitVF**, ma viene utilizzato solo in fase di *cleaning* per uniformare i cataloghi e renderli confrontabili con i modelli teorici come il Vdn [Jennings et al., 2013].

### 3.3.4 Parametri modificabili di BitVF

L'utilizzo di **BitVF** prevede l'assegnazione di un valore a diversi parametri liberi, che controllano le modalità con cui le procedure appena esposte avvengono. Di seguito vengono descritti i principali.

- **N\_rec**: controlla la densità del catalogo di posizioni casuali, che viene generato con una densità numerica pari a **N\_rec** volte quella del catalogo di traccianti. Un

valore più alto migliora l'accuratezza della ricostruzione a scapito del costo computazionale; per ottenere una ricostruzione accurata del campo di spostamento è sconsigliabile scendere sotto al valore 30.

- **threshold**: definisce la soglia di convergenza per la procedura di minimizzazione degli spostamenti. È un numero adimensionale compreso tra 0 e 1; valori prossimi a 0 garantiscono massima accuratezza, mentre valori prossimi a 1 eseguono solo la prima iterazione. Un buon compromesso tra accuratezza e tempo di calcolo è tipicamente compreso tra  $10^{-2}$  e  $10^{-3}$ .
- **cell\_size**: dimensione del lato delle celle della griglia per il calcolo del campo di divergenza, espressa in unità di MPS. Valori troppo piccoli causano un'occupazione delle celle statisticamente irrilevante; mentre valori troppo grandi riducono la risoluzione spaziale del catalogo di vuoti ottenuto.
- **sigma**: deviazione standard del filtro gaussiano applicato al campo di divergenza discusso in precedenza, espressa in unità di MPS. Un valore **sigma** = 0 disabilita il filtraggio, preservando la massima risoluzione a discapito di maggior rumore. Valori tipici per ottenere cataloghi puliti sono dell'ordine di 1 MPS, ma possono dipendere dalla fisica del campione analizzato.
- **seed**: seme per la generazione di numeri casuali, utilizzato nella creazione del catalogo di posizioni lagrangiane. Se impostato a 0, esso è generato automaticamente.

Nel prossimo capitolo, verranno specificati i valori effettivamente utilizzati per le analisi presentate in questa tesi.

### 3.4 Procedura di *void cleaning*

È possibile che un catalogo di vuoti ottenuto da **BitVF** contenga strutture non direttamente utilizzabili per analisi cosmologiche tramite VSF. Possono infatti essere presenti vuoti sovrapposti, sottodensità non sufficientemente profonde (cioè con contrasto di densità centrale non abbastanza vicino a  $-1$ ), vuoti mal centrati o vuoti che contengono al loro interno sovradensità significative (*cloud-in-void*). Per ovviare a questi problemi, si

applica una procedura di *cleaning*, sviluppata a tal proposito da Ronconi and Marulli [2017], che uniforma il catalogo e lo rende confrontabile con i modelli teorici.

La procedura si articola nei seguenti passi:

1. Filtro sulla densità centrale: si calcola la densità media entro un raggio di 2 MPS dal centro del vuoto e vengono rimossi i vuoti per cui il valore superi una soglia prestabilita.
2. Riscalatura del raggio: per ogni vuoto si ricostruisce il profilo di densità radiale e si determina il raggio  $R_v$  per cui il contrasto di densità medio nella sfera centrata sul vuoto è esattamente  $\delta_{v, \text{tr}} = -0.7$ . Questo passo uniforma i cataloghi rendendoli confrontabili con i modelli teorici. Poi si ripete la procedura con una soglia più alta e se non viene raggiunta il vuoto è rimosso, siccome ciò indica un errata assegnazione del centro.
3. Rimozione dei *cloud-in-void*: si calcola il contrasto di densità fra sfere centrate sul vuoto con raggi  $1.5R_v$  e  $R_v$ ; se risulta essere  $< 1$ , la struttura è considerata essere una *cloud-in-void* e quindi eliminata.
4. Filtro sui raggi minimi e massimi: si eliminano i vuoti con raggio inferiore a  $R_{\min}$  o superiore a  $R_{\max}$ ; i valori sono autonomamente impostabili, ricordando però che  $R_{\min}$  è legato alla risoluzione spaziale del catalogo di traccianti.
5. Rimozione delle sovrapposizioni: due vuoti sono considerati sovrapposti se la distanza tra i loro centri è minore della somma dei loro raggi. In caso di sovrapposizione, viene conservato il vuoto più profondo, dunque con densità centrale minore.

In Fig. 3.4 è proposto un minimale schema grafico che illustra tutti i principali procedimenti attuati dall'algoritmo: dall'inserimento di un catalogo di traccianti, scendendo alle procedure discusse nella sezione antecedente, fino al *cleaning* appena descritto, finendo con il catalogo di vuoti trovato.

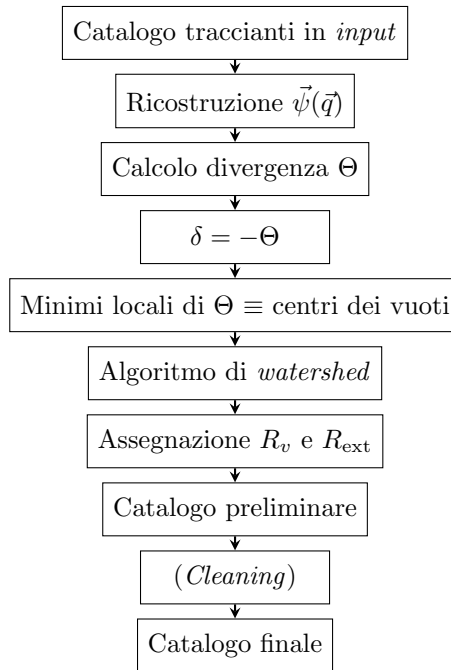


Figura 3.4: Diagramma di flusso dell’algoritmo BitVF. Le principali fasi procedono sequenzialmente dalla ricostruzione del campo spostamento al catalogo finale. Il *cleaning* è opzionale e produce il catalogo finale da utilizzare per le analisi cosmologiche.

## 3.5 CosmoBolognaLib

L'algoritmo `BitVF` è stato implementato all'interno delle `CosmoBolognaLib` (CBL) [Marulli et al., 2016–2026], un insieme di librerie C++ e Python *free software*, disponibili pubblicamente su GitLab, per calcoli cosmologici e analisi dati.

Esse offrono un'efficiente gestione di ampi cataloghi di oggetti astronomici; metodi per il calcolo di funzioni statistiche quali ad esempio: funzione di correlazione a due punti, spettro di potenza, VSF e molte altre; possibilità di generazione di cataloghi simulati in varie geometrie, in questo lavoro di tesi si utilizzerà geometria cubica per il catalogo di traccianti, ma è possibile generare cataloghi e utilizzare `BitVF` anche con geometria a cono luce e dati osservativi, dunque con maggior grado di realistica.

Le librerie sono inoltre organizzate in moduli tematici e, nel prossimo capitolo, verranno presentati risultati ottenuti tramite l'utilizzo di codici C++ e Python derivanti dalla modifica o il diretto utilizzo di codici di esempio, presenti nella sotto-sezione `cosmicVoids` della sezione `Examples` delle librerie.

## Capitolo 4

# Analisi preliminare di BitVF: proprietà statistiche e confronto con VIDE

In questo capitolo vengono presentati e discussi i risultati dell’analisi preliminare di una esecuzione isolata di BitVF e di un primo confronto delle proprietà statistiche tra i cataloghi di vuoti identificati con BitVF e quelli ottenuti con VIDE. L’obiettivo principale è duplice: da un lato, si intende caratterizzare le proprietà dei cataloghi di vuoti identificati con BitVF e valutarne la robustezza; dall’altro, confrontare le sue prestazioni con quelle di VIDE al fine di validare il nuovo metodo e metterne in luce i punti di forza e le eventuali limitazioni.

Tutti i vuoti analizzati sono stati identificati a partire da due cataloghi di aloni di materia oscura, uno in spazio reale e uno in spazio dei redshift, che fungono da traccianti della struttura su scala cosmologica. La simulazione è caratterizzata da una geometria cubica con lato di  $938 h^{-1} \text{Mpc}$ , un redshift di riferimento  $z = 1$  e una massa minima per gli aloni di  $2.16 \times 10^{12} M_{\odot}$ . La scelta di questa simulazione è motivata dalla sua risoluzione e dal volume sufficiente a garantire un campione statisticamente significativo di vuoti. La linea di vista è stata impostata lungo l’asse Y, in modo da poter studiare in modo coerente gli effetti RSD nei confronti tra i cataloghi in spazio reale, in cui non sono presenti, e in spazio dei redshift, nel quale se ne tiene conto.

L’analisi si articola in diverse fasi: inizialmente viene presentata una esecuzione pre-

liminare di **BitVF** per verificarne il funzionamento e la consistenza dei risultati; successivamente si procede a un confronto sistematico tra i cataloghi prodotti da **BitVF** e quelli di **VIDE**, sia in spazio reale che in spazio dei redshift; si analizza inoltre l'effetto del parametro `merging_threshold` sui cataloghi **VIDE**; infine, viene eseguita una procedura di *cleaning* sui cataloghi **BitVF** per valutarne l'impatto sulle proprietà statistiche dei vuoti.

Le sezioni del capitolo seguono le fasi dell'analisi presentate sopra.

## 4.1 Esecuzione preliminare di BitVF

Come analisi preliminare, si è innanzitutto testato il funzionamento di **BitVF**, eseguendo l'algoritmo, implementato all'interno delle CBL, con i seguenti parametri: `N_rec = 30`, `threshold = 10-2`, `cell_size = 1`, `sigma = 1` e `seed = 0`. La scelta di `N_rec = 30` e `threshold = 10-2` è stata dettata dalla necessità di mantenere un costo computazionale contenuto, garantendo al contempo una risoluzione statistica sufficiente per un'analisi dei cataloghi prodotti.

Per visualizzare il catalogo di vuoti individuato, si è reso necessario selezionare una porzione del catalogo di traccianti lungo la linea di vista Y. Lo spessore della "fetta" è stato scelto in modo da essere significativamente inferiore al lato della *box* ( $938 h^{-1}\text{Mpc}$ ), ma al contempo abbastanza grande da contenere un numero statisticamente rilevante di vuoti. Utilizzando come scala di riferimento la separazione media tra particelle (MPS) del catalogo di traccianti, stimata essere di  $8.51 h^{-1}\text{Mpc}$ , si è scelto uno spessore di 3 MPS, pari a circa  $25.5 h^{-1}\text{Mpc}$ , centrato a metà della lunghezza totale del *box* lungo la direzione Y.

Sovrapponendo i vuoti ai traccianti, i primi sono stati rappresentati utilizzando la definizione di raggio equivalente fornita da **BitVF** (Sezione 3.3.3) tramite cerchi: neri per i vuoti individuati in spazio reale e verdi per quelli in spazio dei redshift. I traccianti dei rispettivi spazi sono visualizzati rispettivamente in blu e rosso. La Fig. 4.1 mostra il risultato di questa sovrapposizione per la fetta selezionata.

Come indicato nella legenda della figura, il numero, la posizione e la dimensione dei vuoti individuati nei due spazi non sono necessariamente identici. Questo risultato è dovuto al fatto che lo spazio dei redshift è contaminato dal contributo delle velocità peculiari dei traccianti lungo la linea di vista, che alterano la conversione redshift-distanza e quindi la ricostruzione delle posizioni comoventi. Di conseguenza, le strutture appaiono

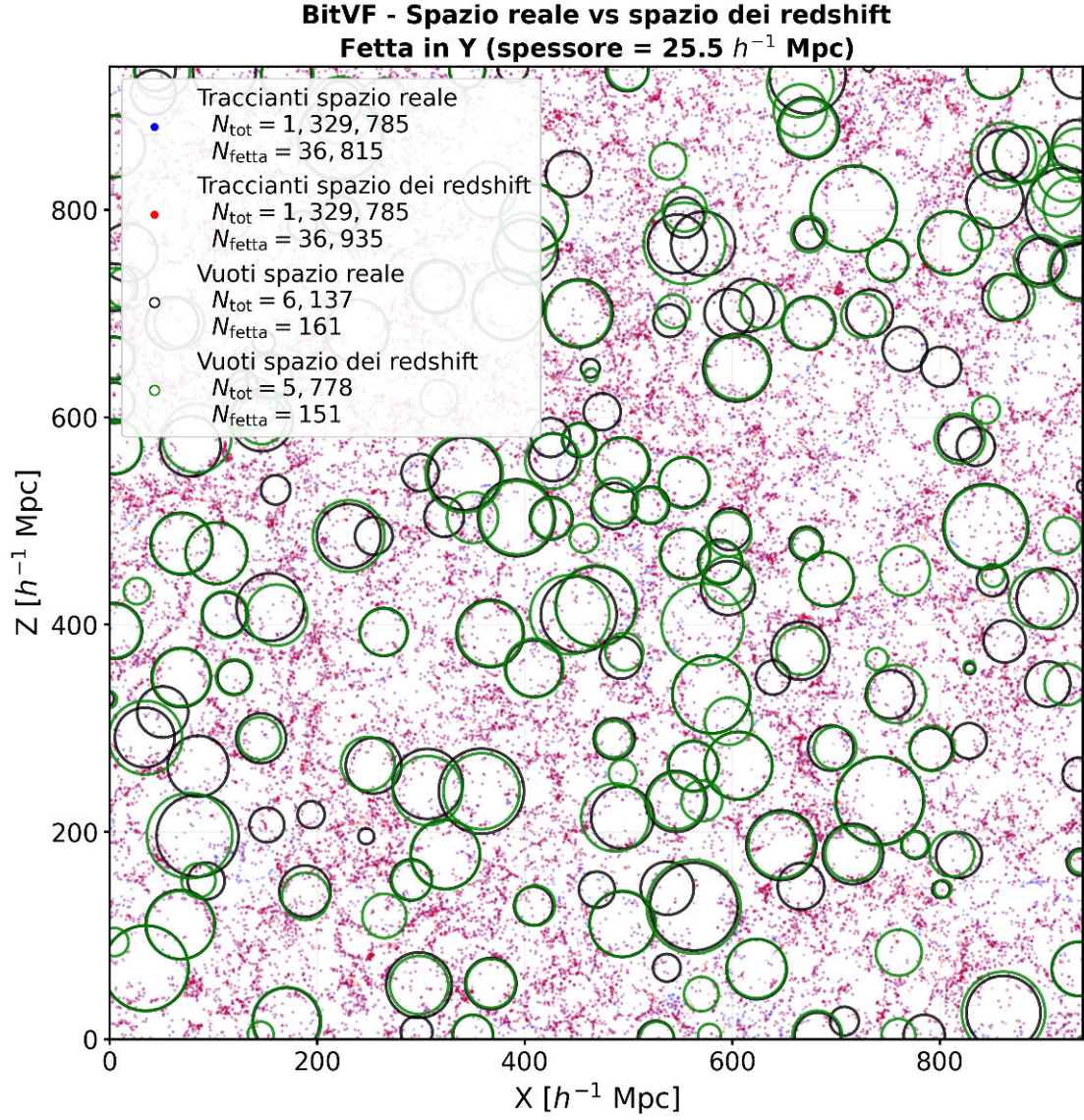


Figura 4.1: Distribuzione dei vuoti individuati da BitVF in una fetta di 3 MPS ( $\approx 25.5 h^{-1} \text{ Mpc}$ ) lungo la direzione Y. I cerchi neri rappresentano i vuoti in spazio reale, quelli verdi i vuoti in spazio dei redshift; i traccianti sono mostrati rispettivamente in blu e rosso. Gli assi X e Z sono espressi in  $h^{-1} \text{ Mpc}$ . La legenda riporta il numero di traccianti e vuoti sia nel volume totale che nella fetta selezionata, per entrambi gli spazi.

distorte e la corrispondenza tra i cataloghi di vuoti ottenuti nei due spazi non è perfetta. Questo aspetto sarà approfondito nel Capitolo 5 attraverso metriche di accoppiamento dedicate.

Nelle analisi che seguono, i risultati verranno presentati applicando un taglio inferiore ai raggi di  $15 h^{-1}\text{Mpc}$ . Questo taglio è motivato dal fatto che, per scale inferiori, l'informazione statistica risulta instabile e caratterizzata da fluttuazioni significative, come emerso dall'analisi delle VSF dei cataloghi di vuoti (si veda la Sezione 4.2).

## 4.2 Confronto tra BitVF e VIDE

Prima di procedere con l'analisi di confronto fra le proprietà statistiche dei due algoritmi, è utile definire la notazione che verrà utilizzata nel seguito per le quantità misurate nei due spazi. Per la void size function (VSF) si utilizzeranno i simboli  $\text{VSF}_R$  e  $\text{VSF}_Z$  per indicare rispettivamente il valore della funzione in spazio reale e in spazio dei redshift. Analogamente, si adotteranno le notazioni  $\text{MP}_R$  e  $\text{MP}_Z$  per il monopolo, e  $\text{QP}_R$  e  $\text{QP}_Z$  per il quadrupolo.

Le incertezze statistiche associate alle diverse quantità sono state stimate con metodologie differenti, scelte in base alla natura della grandezza considerata. Per la VSF, che è essenzialmente un conteggio di vuoti in bin di raggio, si è utilizzato l'errore Poissoniano, che rappresenta una ragionevole approssimazione delle fluttuazioni statistiche per questo tipo di misura. Per monopolo e quadrupolo, invece, trattandosi di quantità che coinvolgono correlazioni spaziali tra vuoti e traccianti, si è reso necessario l'uso del metodo Jackknife, implementato suddividendo il volume in  $N = 4 \times 4 \times 4 = 64$  regioni e ricalcolando i multipoli escludendo una regione alla volta; la varianza è stata poi stimata tramite la formula standard:

$$\sigma_{\text{JK}}^2 = \frac{N-1}{N} \sum_i (\theta_i - \bar{\theta})^2. \quad (4.1)$$

Questo approccio permette di tener conto della *sample variance* e delle covarianze tra diversi bin radiali.

Come primo passo dell'analisi vera e propria, si è proceduto a un confronto sistematico tra i cataloghi di vuoti prodotti da BitVF e quelli ottenuti con VIDE, utilizzando lo stesso catalogo di traccianti simulati. In particolare, sono state confrontate la VSF, i profili di densità (monopoli) e le elongazioni lungo la linea di vista (quadrupoli), introdotti rispettivamente nelle Sezioni 3.2.1 e 3.2.2.

Per garantire un catalogo accurato, **BitVF** è stato eseguito su *cluster* di calcolo con i seguenti parametri: `N_rec` = 50, `threshold` =  $10^{-3}$ , `cell_size` = 0.7937, `sigma` = 1 e `seed` = 0. Il valore di `cell_size`, espresso in unità di MPS, assicura che ogni cella contenga almeno 25 posizioni casuali, garantendo così una buona risoluzione statistica [Sartori, 2022].

Per quanto riguarda **VIDE**, il catalogo utilizzato in questa sezione è stato ottenuto con parametro `merging_threshold` =  $10^{-9}\bar{n}$ . L'effetto di questo parametro e un confronto con una configurazione alternativa verranno discussi nella Sezione 4.3.

L'analisi è stata condotta utilizzando la seconda definizione di raggio fornita da **BitVF** (Sezione 3.3.3), ovvero il raggio esteso,  $R_{ext}$ . Questa scelta è motivata dal fatto che essa si allinea maggiormente alla definizione di raggio adottata da **VIDE**, rendendo il confronto tra i due metodi più significativo.

Le Figg. 4.2a e 4.2b mostrano rispettivamente i vuoti individuati da **BitVF** e **VIDE** in spazio reale e in spazio dei redshift, sovrapposti ai rispettivi traccianti e prendendo in esame la medesima porzione lungo la linea di vista Y discussa in precedenza. In entrambi gli spazi, **VIDE** identifica un numero di vuoti significativamente maggiore rispetto a **BitVF**; la differenza più contenuta si osserva nello spazio dei redshift, dove **VIDE** conta 1359 vuoti in più rispetto a **BitVF**.

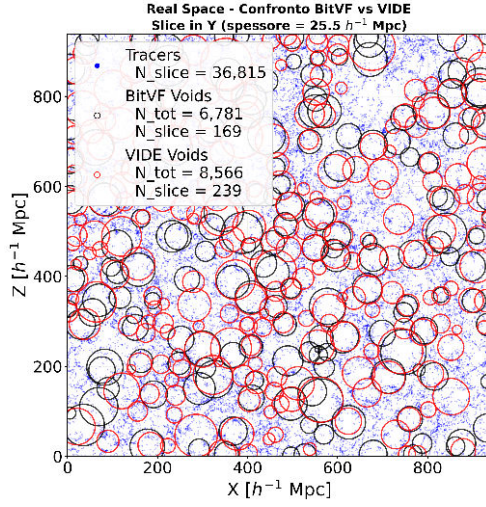
La Fig. 4.2c presenta un confronto diretto tra i cataloghi dei due algoritmi, affiancando le identificazioni in spazio reale e spazio dei redshift per ciascun metodo. La differenza nelle popolazioni di vuoti tra i due spazi risulta qui evidente per entrambi gli algoritmi, sebbene con intensità diverse: **VIDE** mostra una variazione più pronunciata nel passaggio da spazio reale a spazio dei redshift, mentre **BitVF** sembra preservare una maggiore coerenza tra i due cataloghi.

La Fig. 4.3 mostra le VSF in spazio reale ( $VSF_R$ ) e spazio dei redshift ( $VSF_Z$ ) per entrambi i metodi. Si osservano curve confrontabili entro le incertezze su scale maggiori di  $\approx 25 h^{-1}\text{Mpc}$  fino a  $\approx 60 h^{-1}\text{Mpc}$ , mentre su scale inferiori a  $25 h^{-1}\text{Mpc}$  la VSF di **VIDE** risulta sistematicamente più elevata. Questo comportamento indica che **VIDE** identifica un numero maggiore di vuoti di piccole dimensioni a parità di raggio, in accordo con quanto osservato nella visualizzazione dei cataloghi (Fig. 4.2c). Le bande colorate attorno alle curve in spazio reale rappresentano le incertezze statistiche Poissoniane. Nei pannelli inferiori sono riportati i residui  $(VSF_Z - VSF_R)/\sigma$ , calcolati in quadratura sugli errori dei due cataloghi; le bande orizzontali sfumate indicano le regioni a  $\pm 1\sigma$  e  $\pm 2\sigma$ . Per

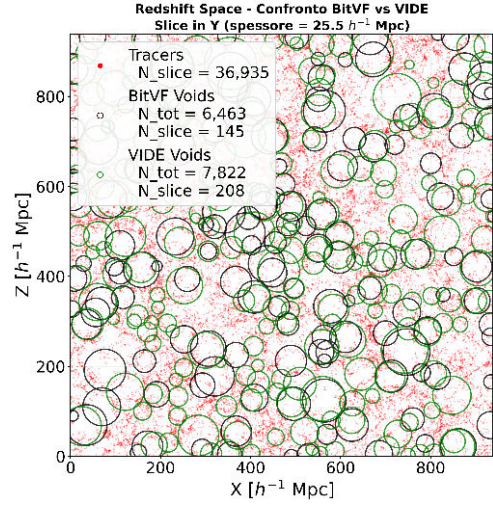
**BitVF**, la frazione di punti entro  $2\sigma$  è del  $(92 \pm 7)\%$ , mentre per **VIDE** scende al  $(57 \pm 13)\%$ , indicando che **BitVF** è più stabile nel passaggio da spazio reale a spazio dei redshift. La correlazione tra i valori nei due spazi per **BitVF** è stata stimata a  $r = (0.998 \pm 0.002)$ , leggermente superiore a quella di **VIDE**:  $r = (0.996 \pm 0.002)$ , ma confrontabile entro le incertezze, mentre le differenze medie percentuali sono del  $-5.0\%$  per **BitVF** e del  $-8.9\%$  per **VIDE**, suggerendo che l'effetto RSD riduce il numero di vuoti in modo più marcato per **VIDE**.

La Fig. 4.4 mostra i monopoli in entrambi gli spazi con i relativi residui sottostanti. I monopoli mostrano che **BitVF** e **VIDE** producono profili di densità con ampiezze diverse: **BitVF** raggiunge un valore minimo di  $\xi_0 = -(0.978 \pm 0.002)$ , mentre **VIDE** si ferma a  $\xi_0 = -(0.843 \pm 0.008)$ ; inoltre il monopolio di **BitVF**, in entrambi gli spazi, rimane al di sotto di quello di **VIDE** fino a circa  $s/R_v \approx 1.5$ , indicando che i profili di densità dei vuoti identificati da **BitVF** sono più profondi su queste scale. I residui  $(MP_Z - MP_R)/\sigma$  evidenziano che **BitVF** è complessivamente più stabile nel passaggio da uno spazio all'altro: la deviazione standard dei residui di **BitVF** ( $\sigma_\delta = 0.049$ ) è infatti circa due volte più piccola di quella di **VIDE** ( $\sigma_\delta = 0.121$ ). La correlazione tra i monopoli di **BitVF** e **VIDE** è molto elevata in entrambi gli spazi: entrambe stimate a  $r = (0.997 \pm 0.002)$ ; indicando che, nonostante le differenze nella sensibilità all'effetto RSD, i profili di densità prodotti dai due metodi sono sostanzialmente compatibili.

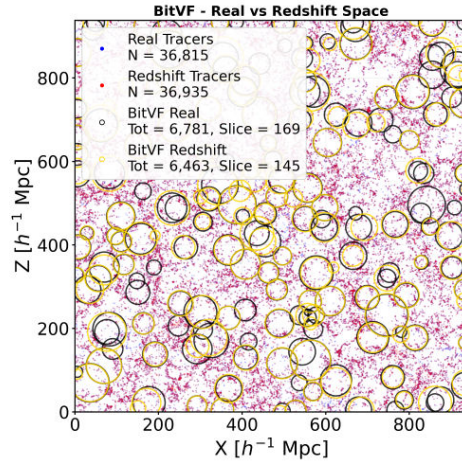
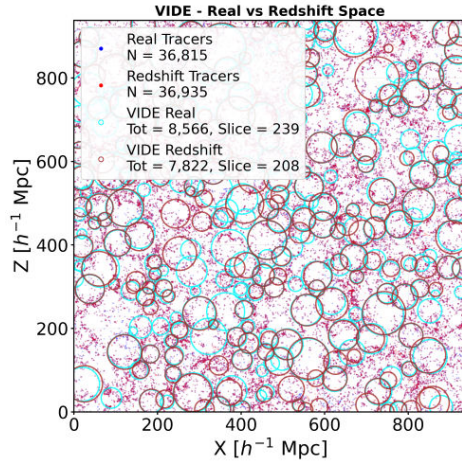
La Fig. 4.5 mostra i quadrupoli per entrambi i metodi nei due spazi. Il quadrupolo, che misura l'elongazione dei vuoti lungo la linea di vista, rivela alcune differenze tra i due metodi. In spazio reale, i quadrupoli di **BitVF** e **VIDE** sono coerentemente compatibili con zero a tutte le scale, come atteso in assenza di distorsioni lungo la linea di vista. In spazio dei redshift, invece, la correlazione tra i quadrupoli dei due metodi è calcolata a  $r = (0.90 \pm 0.05)$ , suggerendo che gli effetti RSD inducono un'elongazione fisicamente coerente in entrambi i cataloghi, sebbene con ampiezze differenti nelle varie scale.



(a) Vuoti individuati da BitVF (neri) e VIDE (rossi) in spazio reale, sovrapposti ai traccianti (blu). La legenda riporta il numero di traccianti e vuoti per entrambi i cataloghi, sia sul volume totale che nella fetta selezionata.



(b) Vuoti individuati da BitVF (neri) e VIDE (verdi) in spazio dei redshift, sovrapposti ai traccianti (rossi). La legenda riporta il numero di traccianti e vuoti per entrambi i cataloghi, sia sul volume totale che nella fetta selezionata.



(c) Confronto tra i vuoti identificati nei due spazi. A sinistra: VIDE in azzurro (spazio reale) e marrone (spazio dei redshift). A destra: BitVF in nero (spazio reale) e giallo (spazio dei redshift). Le legende riportano i numeri di traccianti e vuoti per ciascun catalogo.

Figura 4.2: Confronto tra i cataloghi di vuoti individuati da BitVF e VIDE a partire dagli stessi traccianti. Gli assi X e Z sono espressi in  $h^{-1}\text{Mpc}$ ; la fetta lungo Y ha spessore 3 MPS ( $\approx 25.5 h^{-1}\text{Mpc}$ ).

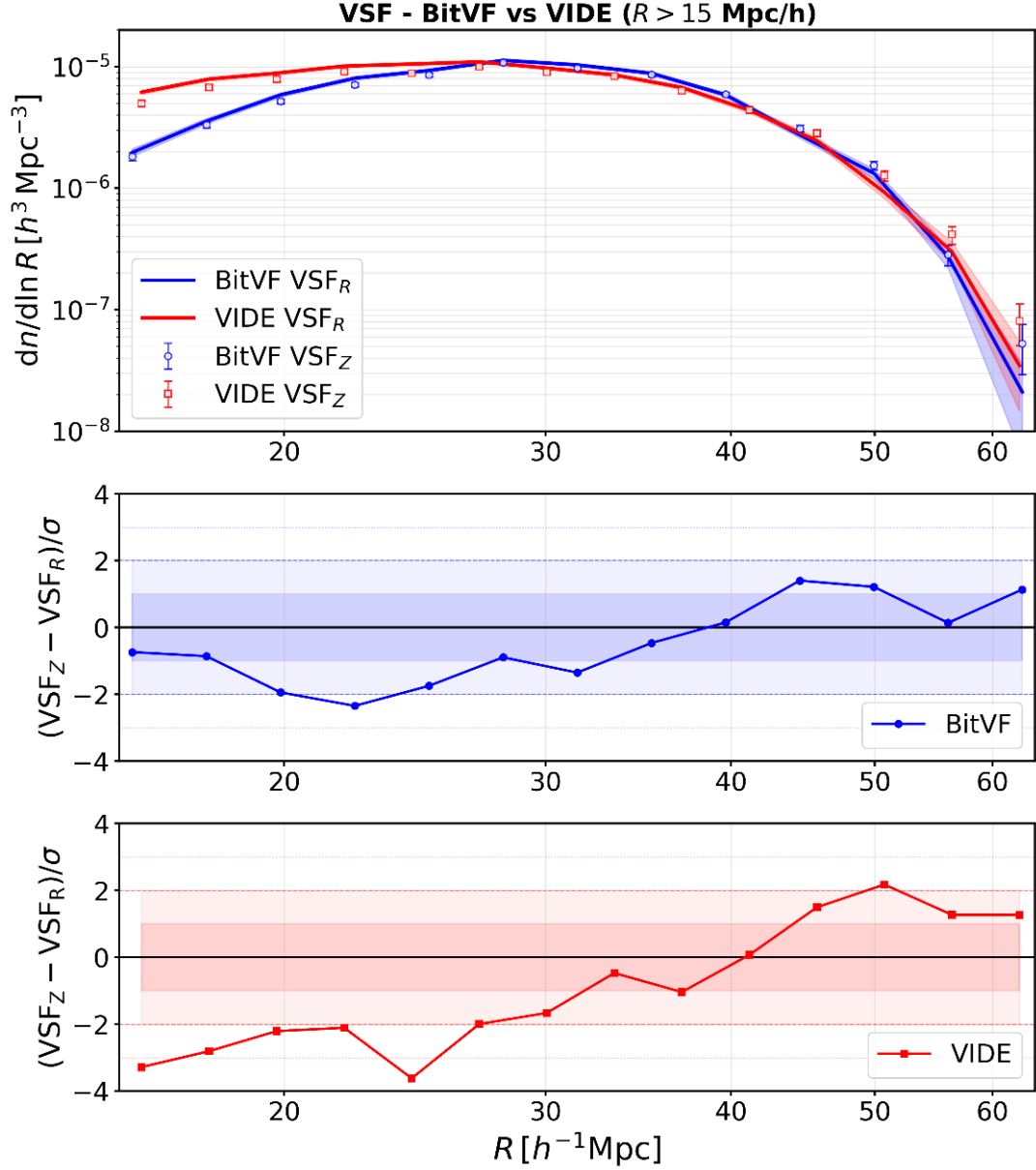


Figura 4.3: Confronto tra le VSF ottenute da BitVF (blu) e VIDE (rosso), per raggi  $R \geq 15 h^{-1} \text{Mpc}$ . Le curve continue rappresentano i cataloghi in spazio reale ( $\text{VSF}_R$ ), mentre i simboli con barre di errore indicano quelli in spazio dei redshift ( $\text{VSF}_Z$ ). Le bande colorate attorno alle curve continue rappresentano le incertezze statistiche (Poissoniane) sui valori in spazio reale. Nei pannelli inferiori sono riportati i residui  $(\text{VSF}_Z - \text{VSF}_R)/\sigma$ , calcolati in quadratura sugli errori dei due cataloghi; le bande orizzontali sfumate evidenziano le regioni a  $\pm 1\sigma$  e  $\pm 2\sigma$ .

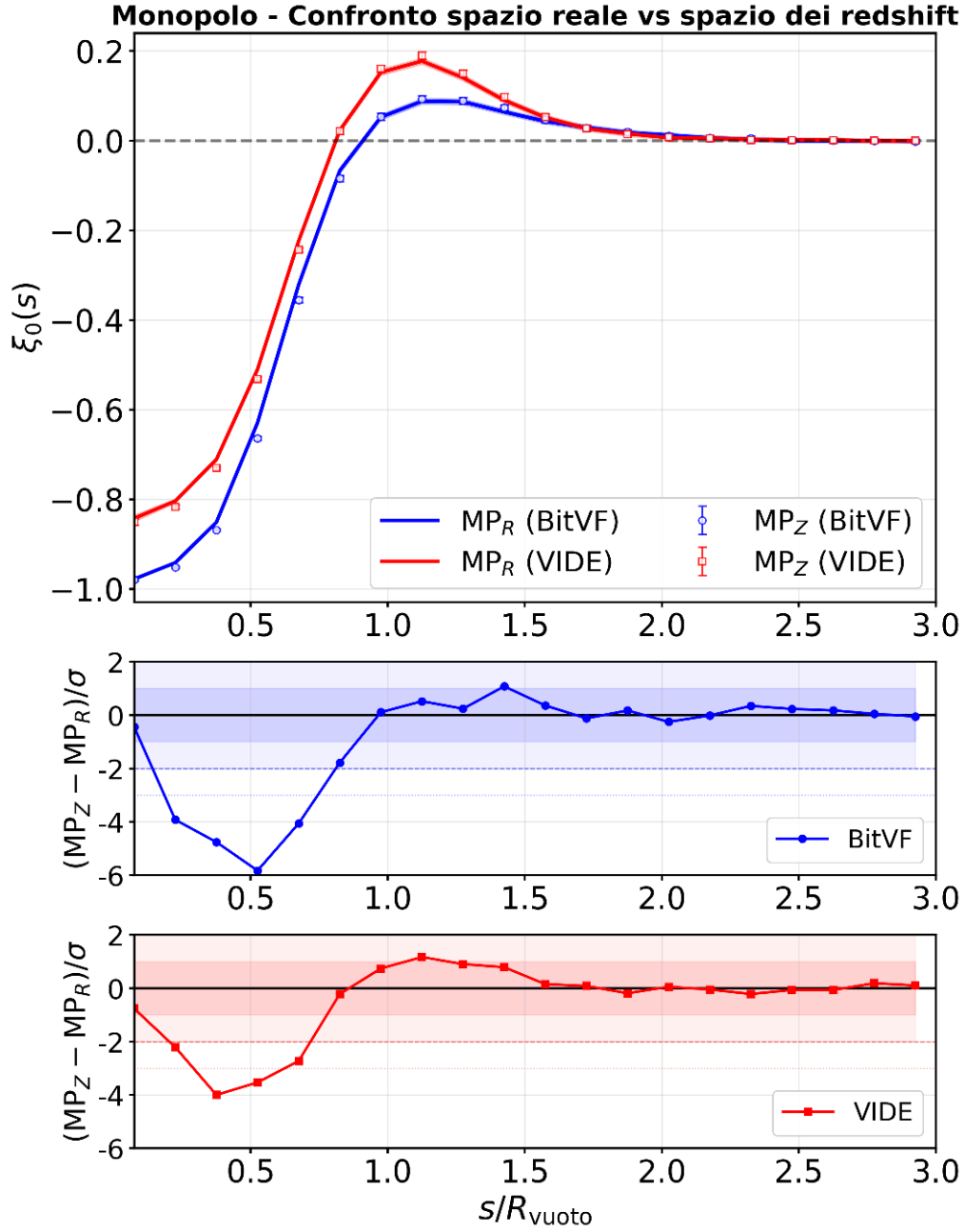


Figura 4.4: Confronto dei monopoli (profili di densità) di BitVF (blu) e VIDE (rosso) in spazio reale ( $MP_R$ , linee continue) e spazio dei redshift ( $MP_Z$ , simboli con barre di errore). L'asse delle ascisse è espresso in unità di  $s/R_v$ . Nei pannelli inferiori sono riportati i residui  $(MP_Z - MP_R)/\sigma$  per ciascun metodo, calcolati in quadratura sugli errori; le bande orizzontali grigie indicano le regioni a  $\pm 1\sigma$  e  $\pm 2\sigma$ .

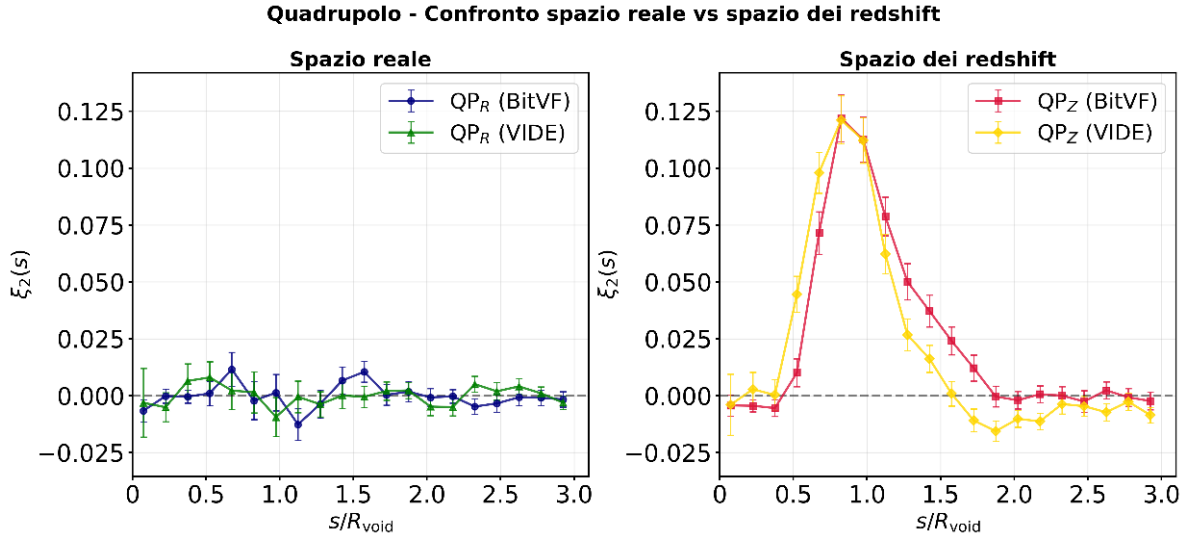


Figura 4.5: Confronto dei quadrupoli di BitVF e VIDE. Il pannello di sinistra mostra il quadrupolo in spazio reale ( $QP_R$ ), quello di destra in spazio dei redshift ( $QP_Z$ ). BitVF è rappresentato in blu (spazio reale) e rosso (spazio dei redshift); VIDE in verde (spazio reale) e giallo (spazio dei redshift). L'asse delle ascisse è espresso in unità di  $s/R_v$ , dove  $s$  è la separazione comovente e  $R_v$  il raggio del vuoto. Le barre di errore rappresentano le incertezze statistiche stimate tramite Jackknife.

### 4.3 Analisi del parametro `merging_threshold` per cataloghi VIDE

Come accennato nella Sezione 4.2, il parametro `merging_threshold` controlla la fusione dei vuoti adiacenti identificati da VIDE. Per comprendere il suo funzionamento, è necessario chiarire innanzitutto la natura fisica di questo parametro. In VIDE, la soglia di fusione è un valore di densità, misurato in unità della densità media dei traccianti,  $\bar{n}$  [Hamaus et al., 2014]. Due vuoti adiacenti vengono fusi se la densità minima lungo il confine che li separa è inferiore al valore di soglia impostato. Questa logica è alla base dell’algoritmo *watershed* su cui VIDE si basa: la soglia rappresenta il massimo livello di densità del confine che può essere superato per fondere due bacini adiacenti in un unico vuoto.

In questa analisi sono stati considerati due casi estremi:

- `merging_threshold = 0`: in base alla convenzione adottata nel codice, questo valore corrisponde a una soglia infinita. La condizione per la fusione (densità del confine  $<$  soglia) è quindi sempre verificata, e tutti i vuoti adiacenti vengono fusi. Questo produce un catalogo con un numero ridotto di vuoti, ma di dimensioni molto maggiori, che arrivano a coprire regioni estese della distribuzione di materia.
- `merging_threshold = 10-9 $\bar{n}$` : questo è il valore di default di VIDE, scelto per essere così piccolo da impedire qualsiasi fusione. La densità minima sul confine tra due vuoti non scende mai a valori così estremamente bassi, quindi la condizione per fondersi non viene mai soddisfatta. Questo catalogo preserva la struttura fine del campo di densità, restituendo un numero maggiore di vuoti di piccole dimensioni.

Sebbene in questa sede ci si sia concentrati sui due estremi, è importante sottolineare che il parametro ammette valori intermedi. Ad esempio, in letteratura viene spesso utilizzato un valore di  $0.2\bar{n}$  per la soglia di fusione [Sutter et al., 2014], che ha un significato fisico ben preciso: deriva dal modello di espansione sferica di una perturbazione in un universo Einstein-de Sitter e segna la densità caratteristica all’interno della perturbazione quando avviene il *shell-crossing* ai suoi confini. Valori intermedi come questo permettono la formazione di una gerarchia di vuoti con strutture ”genitore-figlio”, in cui i vuoti più piccoli sono contenuti all’interno di vuoti più grandi.

Per comprendere l’impatto di questo parametro sulle proprietà statistiche dei cataloghi, si sono confrontati i risultati ottenuti con le due configurazioni estreme. Nelle figure riportate si utilizzerà l’acronimo ”tsh” per indicare il parametro. La Fig. 4.6 mostra le VSF per i cataloghi VIDE con `merging_threshold` =  $10^{-9}\bar{n}$  (utilizzato nel confronto con BitVF) e con `merging_threshold` = 0 (soglia infinita). La differenza più evidente è il range di raggi accessibile: mentre la configurazione con soglia  $10^{-9}\bar{n}$  produce vuoti con raggi fino a circa  $60 h^{-1}\text{Mpc}$ , quella con soglia 0 estende il campionamento fino a quasi  $240 h^{-1}\text{Mpc}$ , rivelando la presenza di vuoti su scale molto più grandi e una forte riduzione del numero di vuoti su scale inferiori ai  $40 h^{-1}\text{Mpc}$ ; questo comportamento è una diretta conseguenza della fusione di vuoti adiacenti.

L’analisi dei residui, riportata nei pannelli inferiori della figura e sintetizzata nella Tabella 4.1, mostra che entrambe le configurazioni sono caratterizzate da una correlazione molto elevata tra i due spazi ( $r > 0.99$ ), indicando una buona stabilità complessiva. Tuttavia, la configurazione con soglia  $10^{-9}\bar{n}$  presenta una differenza media percentuale tra spazio reale e dei redshift del  $-8.9\%$  e una frazione di punti entro  $2\sigma$  del  $57.1\%$ , valori che suggeriscono una maggiore sensibilità agli effetti RSD rispetto alla configurazione con soglia 0 ( $-8.2\%$  e  $80.0\%$  rispettivamente). Questo risultato è probabilmente dovuto al fatto che la fusione dei vuoti adiacenti tende a mediare le anisotropie locali indotte dagli effetti RSD, rendendo il catalogo con soglia 0 statisticamente più stabile nonostante la presenza di vuoti di dimensioni maggiori.

Tabella 4.1: Statistiche degli effetti RSD sulla VSF per le due configurazioni di VIDE.

|                       | <code>merging_threshold</code> = 0 | <code>merging_threshold</code> = $10^{-9}\bar{n}$ |
|-----------------------|------------------------------------|---------------------------------------------------|
| Correlazione $r$      | $0.997 \pm 0.002$                  | $0.996 \pm 0.002$                                 |
| Differenza % media    | $-8.2 \pm 1.6\%$                   | $-8.9 \pm 1.6\%$                                  |
| Punti entro $2\sigma$ | $80 \pm 10\%$                      | $57 \pm 13\%$                                     |

La Fig. 4.7 mostra i monopoli per le due configurazioni di VIDE. I monopoli rivelano che la configurazione con soglia  $10^{-9}\bar{n}$  produce profili di densità significativamente più profondi in entrambi gli spazi, con un valore minimo di  $\xi_0 \simeq -0.85$  in spazio dei redshift contro il  $\xi_0 \simeq -0.45$  della configurazione con soglia 0. Questa differenza è attribuibile alla presenza di vuoti più piccoli e meglio definiti, che generano contrasti di densità più marcati. I residui  $(MP_Z - MP_R)/\sigma$  mostrano che la configurazione con soglia 0 è

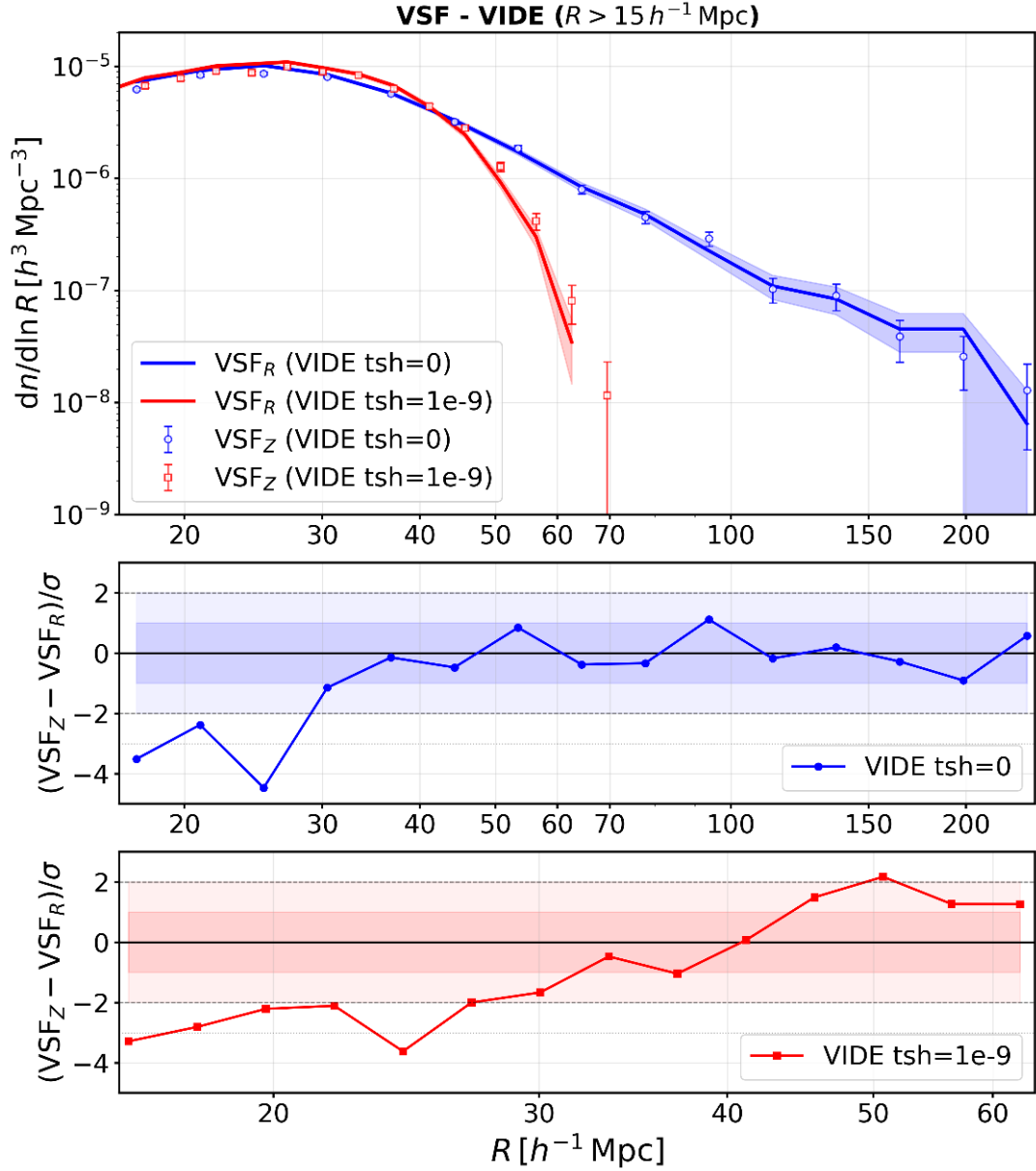


Figura 4.6: Confronto tra le VSF di VIDE con `merging_threshold` =  $10^{-9}\bar{n}$  (rosso) e `merging_threshold` = 0 (blu), per raggi  $R \geq 15 h^{-1} \text{Mpc}$ . Le curve continue rappresentano i cataloghi in spazio reale ( $\text{VSF}_R$ ), i simboli con barre di errore quelli in spazio dei redshift ( $\text{VSF}_Z$ ). L'asse delle ascisse riporta il raggio dei vuoti in  $h^{-1} \text{Mpc}$ ; l'asse delle ordinate la densità numerica  $dn/d\ln R$  in  $h^3 \text{Mpc}^{-3}$ . Nei pannelli inferiori sono riportati i residui  $(\text{VSF}_Z - \text{VSF}_R)/\sigma$ ; le bande orizzontali grigie indicano le regioni a  $\pm 1\sigma$  e  $\pm 2\sigma$ .

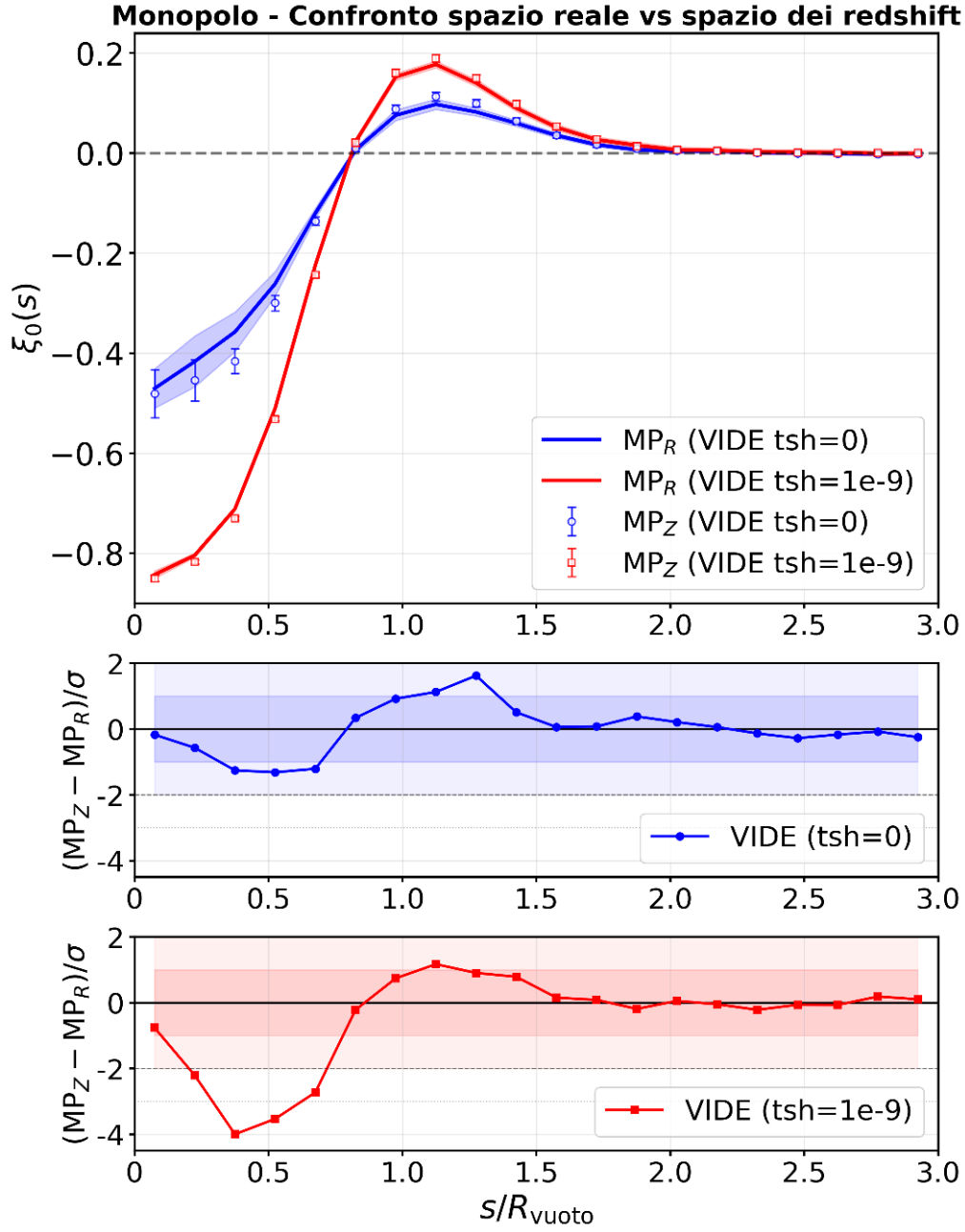


Figura 4.7: Confronto dei monopoli (profili di densità) di VIDE per le due configurazioni di `merging.threshold`. La configurazione con soglia  $10^{-9}\bar{n}$  è rappresentata in blu, quella con soglia 0 in rosso. Le curve continue rappresentano i cataloghi in spazio reale ( $MP_R$ ), i simboli con barre di errore quelli in spazio dei redshift ( $MP_Z$ ). L'asse delle ascisse è espresso in unità di  $s/R_v$ . Nei pannelli inferiori sono riportati i residui  $(MP_Z - MP_R)/\sigma$  per ciascuna configurazione; le bande orizzontali grigie indicano le regioni a  $\pm 1\sigma$  e  $\pm 2\sigma$ .

complessivamente più stabile nel passaggio da uno spazio all'altro, con una deviazione standard dei residui  $\sigma_\delta = 0.08$  contro  $\sigma_\delta = 0.12$  della configurazione con soglia  $10^{-9}\bar{n}$ , in accordo con quanto osservato per la VSF.

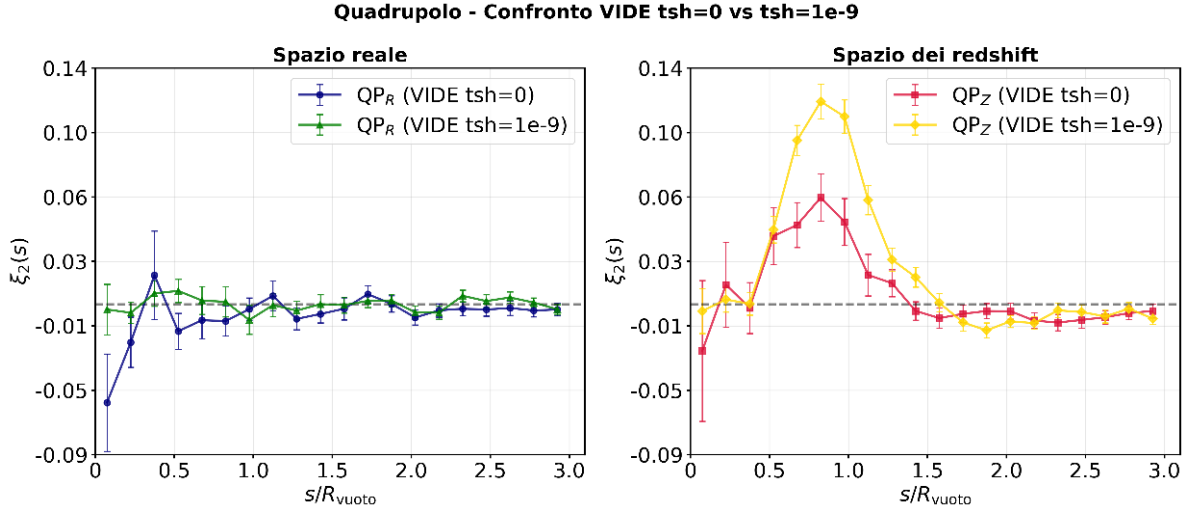


Figura 4.8: Confronto dei quadrupoli di VIDE per le due configurazioni di `merging_threshold`. Il pannello di sinistra mostra il quadrupolo in spazio reale ( $QP_R$ ), quello di destra in spazio dei redshift ( $QP_Z$ ). La configurazione con soglia  $10^{-9}\bar{n}$  è rappresentata in blu, quella con soglia 0 in rosso. L'asse delle ascisse è espresso in unità di  $s/R_v$ . Le barre di errore rappresentano le incertezze statistiche stimate tramite Jackknife.

I quadrupoli, mostrati in Fig. 4.8, mostrano un comportamento generalmente simile tra le due configurazioni: in entrambi i casi, il quadrupolo in spazio reale ( $QP_R$ ) è compatibile con zero entro le barre di errore quasi ovunque, come atteso in assenza di distorsioni lungo la linea di vista, mentre in spazio dei redshift ( $QP_Z$ ) si osserva un'elongazione coerente, con correlazione tra i quadrupoli delle due configurazioni elevata, stimata  $r = (0.92 \pm 0.04)$ , suggerendo che l'effetto RSD induce un'elongazione fisicamente simile, indipendentemente dalla scala considerata. Tuttavia, l'ampiezza di quest'ultima varia significativamente tra le due configurazioni: quella con soglia 0 raggiunge un valore massimo di  $\xi_2 \simeq 0.05$ , mentre quella con soglia  $10^{-9}\bar{n}$  arriva a  $\xi_2 \simeq 0.125$ , indicando come la configurazione con soglia  $10^{-9}\bar{n}$  fornisca una maggiore informazione sulle distorsioni lungo la linea di vista.

In conclusione, la scelta del parametro `merging_threshold` ha un impatto signifi-

cativo sulle proprietà statistiche dei cataloghi di vuoti di **VIDE**. La configurazione con soglia infinita (dunque con valore 0 assegnato al parametro) rivela la presenza di vuoti su scale molto grandi, fino a  $240 h^{-1}\text{Mpc}$ , a scapito di una minore risoluzione statistica e di profili di densità meno profondi. La configurazione con soglia  $10^{-9}\bar{n}$ , invece, produce cataloghi di vuoti più piccoli e meglio definiti, ma risulta complessivamente più sensibile agli effetti RSD e copre un intervallo di scale più limitato. Per il confronto con **BitVF** presentato nella Sezione 4.2, si è scelta la configurazione con soglia  $10^{-9}\bar{n}$ , in quanto produce cataloghi di vuoti di dimensioni confrontabili con quelli di **BitVF** e più adatti a un’analisi dettagliata delle proprietà statistiche su scale intermedie.

### 4.3.1 *Cleaning* dei cataloghi **BitVF**

Come approfondimento aggiuntivo, si è voluto verificare l’effetto della procedura di *cleaning* (descritta nella Sezione 3.4) sui cataloghi di **BitVF** utilizzati nell’analisi di riproducibilità. Questa analisi, che costituisce un elemento di corredo rispetto agli obiettivi principali della tesi, consente di valutare come la procedura di *cleaning* modifichi le proprietà statistiche dei vuoti e come si rifletta sui risultati di accoppiamento tra esecuzioni indipendenti.

Ottenute due esecuzioni indipendenti di **BitVF** con medesimi parametri, il *cleaning* è stato applicato con i seguenti parametri principali. Il parametro `delta_r` = [18.0, 40.0]  $h^{-1}\text{Mpc}$  definisce l’intervallo di raggi entro cui viene ricostruito il profilo di densità di ciascun vuoto: il limite inferiore è stato scelto per escludere scale troppo piccole dove lo *shot-noise* potrebbe compromettere la ricostruzione, mentre il limite superiore rappresenta la scala massima entro cui il profilo è ben descritto dal modello Vdn [Jennings et al., 2013]. Il parametro `initial_radius` specifica il raggio iniziale da cui partire per la procedura di riscalatura del raggio: se non viene fornito, viene utilizzato il raggio  $R_{\text{ext}}$  calcolato da **BitVF**. Il parametro booleano `rescaling`, se impostato a `True`, abilita la procedura di riscalatura del raggio che ridimensiona ciascun vuoto in modo che il suo contrasto di densità medio raggiunga il valore  $\delta_{v,\text{tr}} = -0.7$  (specificato da `deltav_NL`); questa operazione uniforma i cataloghi rendendoli confrontabili con i modelli teorici come il Vdn [Jennings et al., 2013]. La soglia di contrasto di densità per la riscalatura del raggio è stata impostata a `deltav_NL` =  $-0.7$ , valore standard per cataloghi di vuoti che consente il confronto con modelli teorici di evoluzione non lineare [Ronconi and Marulli,

2017]. La procedura di rimozione delle sovrapposizioni è stata attivata (`overlap = True`), con il criterio di conservazione del vuoto più profondo (ovvero con densità centrale minore). Infine, il parametro `ratio = 0.5` controlla la soglia di sovrapposizione tra vuoti: due vuoti vengono considerati sovrapposti se la distanza tra i loro centri è minore di `ratio` volte la somma dei loro raggi; con `ratio = 0.5`, la sovrapposizione è consentita fino a metà del diametro del vuoto più piccolo, e se il centro di un vuoto cade all'interno di un vuoto più grande, il vuoto più piccolo viene eliminato.

L'applicazione della procedura ha comportato una significativa riduzione del numero di vuoti in entrambi i cataloghi, come evincibile dalla legenda della Figura 4.9, che presenta le proiezioni in porzioni di spessore  $\Delta Y = 25.5 h^{-1} \text{Mpc}$  dei cataloghi delle due esecuzioni, sia in spazio reale che in spazio dei redshift, mostrando la distribuzione dei vuoti dopo il *cleaning*. I vuoti evidenziati con cerchi tratteggiati sono quelli accoppiati tramite la metrica del primo vicino, discussa nel Capitolo 5.

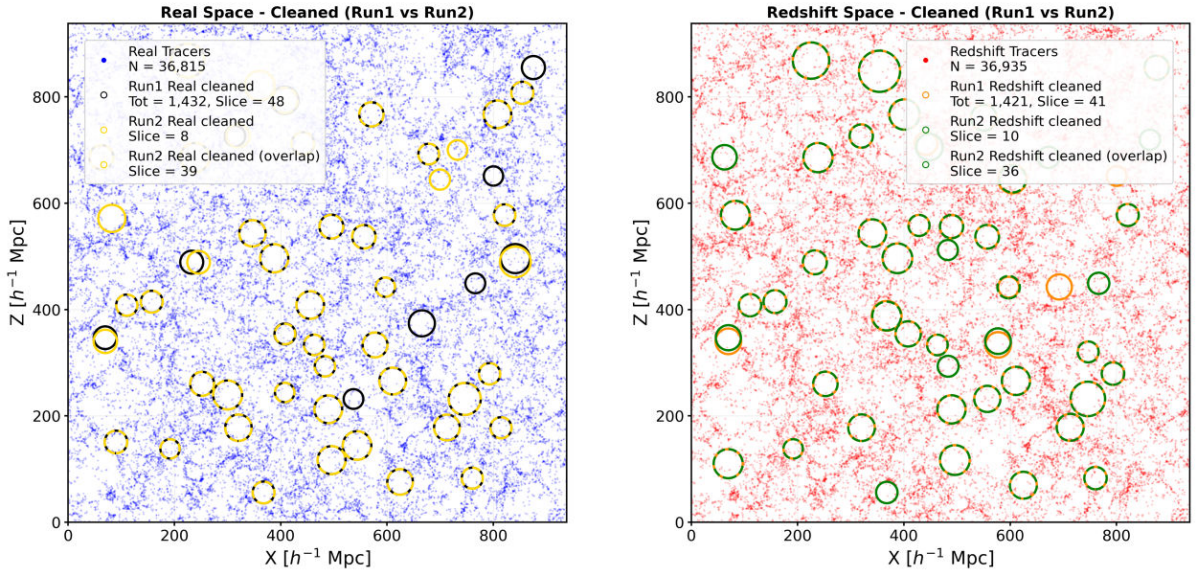


Figura 4.9: Confronto dei cataloghi delle due esecuzioni di **BitVF** dopo l'applicazione della procedura di *cleaning*. I pannelli di sinistra e destra mostrano rispettivamente lo spazio reale e lo spazio dei redshift. Nei rispettivi spazi, i vuoti della Esecuzione 1 sono rappresentati in nero e verde, quelli della Esecuzione 2 in giallo e arancione, mentre i vuoti accoppiati sono evidenziati con tratteggio. La proiezione è effettuata sulla medesima porzione di spessore  $\Delta Y = 25.5 h^{-1} \text{Mpc}$  considerata in precedenza.

Questa analisi, sebbene marginale rispetto al fine principale della tesi, suggerisce che

l'applicazione del *cleaning* selezioni vuoti più profondi e "puri", i quali risultano meno sensibili alle sistematiche indotte dalle distorsioni in fase di identificazione. L'incremento della frazione di accoppiamenti osservato dopo il *cleaning* supporta questa interpretazione, indicando che la procedura non introduce distorsioni sistematiche tra le due esecuzioni, ma agisce in modo consistente su entrambi i cataloghi, migliorando la corrispondenza tra i vuoti identificati. Una conferma più accurata richiederebbe tuttavia la ripetizione sistematica delle analisi sui cataloghi puliti; per questo motivo, quanto osservato ha qui valore prevalentemente indicativo.

## Capitolo 5

# Validazione di BitVF: analisi di riproducibilità e confronto tra spazio reale e spazio dei redshift

In questo capitolo vengono presentate e discusse le analisi di validazione di **BitVF**, con particolare attenzione alla riproducibilità dell’algoritmo e alla sua robustezza nell’identificazione dei vuoti nello spazio dei redshift. L’analisi si articola in tre sezioni principali, corrispondenti a tre diverse metriche di confronto tra cataloghi di vuoti.

Prima di procedere, è opportuno chiarire la scelta della definizione di raggio adottata per **BitVF** nel confronto con **VIDE**. Come discusso nella Sezione 3.3.3, **BitVF** fornisce due definizioni di raggio: il raggio equivalente del bacino di sottodensità ( $R_v$ ) e il raggio esteso ( $R_{\text{ext}}$ ). Per il confronto tra spazio reale e dei redshift per **BitVF**, si è scelto di utilizzare esclusivamente il raggio esteso. Questa scelta è motivata dalla maggiore compatibilità di  $R_{\text{ext}}$  con la definizione di raggio adottata da **VIDE**. Tuttavia, come mostrato nella Sezione 5.1.1, le due definizioni di raggio fornite da **BitVF** producono risultati sostanzialmente equivalenti per quanto riguarda le proprietà statistiche dei cataloghi. L’uso di  $R_{\text{ext}}$  garantisce quindi una maggiore omogeneità nel confronto con **VIDE** senza perdita di informazione statistica.

Per tutte le analisi, i cataloghi di vuoti utilizzati sono fondamentalmente gli stessi descritti nel Capitolo 4: due esecuzioni indipendenti di **BitVF** (`N_rec` = 50, `threshold` =  $10^{-3}$ , `cell.size` = 0.7937, `sigma` = 1) e il catalogo **VIDE** con `merging_threshold` =

$10^{-9}$ .

Si noti che le due esecuzioni di **BitVF** nello spazio reale e spazio dei redshift sono state condotte con parametro `seed` = 0, che corrisponde a una generazione automatica del seme per il catalogo di posizioni lagrangiane iniziali. Di conseguenza, i due cataloghi sono stati ottenuti con semi diversi, introducendo una componente stocastica aggiuntiva che si somma agli effetti delle RSD. In linea di principio, per isolare esclusivamente gli effetti RSD, sarebbe opportuno utilizzare lo stesso seme per entrambe le esecuzioni, oppure normalizzare lo scatter osservato rispetto a quello che si otterrebbe confrontando due esecuzioni indipendenti nello stesso spazio. La riproducibilità di **BitVF** tra esecuzioni indipendenti con semi diversi è robusta (frazioni di identificazione sempre  $> 83\%$ ,  $r \simeq 0.98$ ) ma nel confronto tra lo spazio reale e dei redshift la frazione di identificazioni scende, come si vedrà nel corso del Capitolo, a valori inferiori a quelli tra esecuzioni indipendenti nello stesso spazio. Ciò indica che, sebbene la maggior parte della riduzione di corrispondenza sia attribuibile agli effetti RSD, l'effetto del seme non sia trascurabile.

Va inoltre chiarito che i cataloghi di vuoti nello spazio reale rappresentano il punto di riferimento per questa analisi: essi costituiscono il catalogo che **BitVF** e **VIDE** individuano a partire dalla distribuzione dei traccianti non distorto dalle velocità peculiari. Il catalogo nello spazio dei redshift, invece, si discosta dal catalogo di riferimento a causa degli effetti RSD e tanto più il catalogo nello spazio dei redshift si discosta da quello nello spazio reale, tanto maggiore è l'impatto delle distorsioni sull'identificazione dei vuoti. L'obiettivo delle analisi presentate in questo capitolo è quindi duplice: da un lato, valutare quale algoritmo produca cataloghi di vuoti nello spazio dei redshift più simili a quelli nello spazio reale; dall'altro, comprendere come le RSD modifichino i cataloghi di vuoti e quali caratteristiche di questi effetti siano catturate dai diversi algoritmi.

Prima di procedere con la discussione dei risultati, è opportuno sottolineare alcuni aspetti metodologici che caratterizzano tutte le analisi presentate in questo capitolo. In primo luogo, in tutti gli istogrammi viene riportato il numero di *outliers* ( $N_{\text{outliers}}$ ): ovvero il numero di vuoti che, per ciascuna distribuzione, cadono al di fuori del range visualizzato. In particolare, per gli istogrammi delle distanze normalizzate il range è fissato a  $[0, 0.1]$ , valore dettato dalla soglia  $s = 0.1$  discussa nella Sezione 5.1, mentre per gli istogrammi delle differenze relative il range è determinato dinamicamente come il percentile 1% e 99% della distribuzione, per escludere code estreme non rappresentative. I vuoti al di fuori di questi range vengono esclusi dalla rappresentazione grafica in quanto

statisticamente non rilevanti per la comprensione del comportamento complessivo, ma il loro numero è sempre riportato in legenda. Si noti che le scale degli assi delle ordinate non sono uniformate tra i diversi pannelli, poiché i conteggi assoluti variano da un istogramma all'altro; tale scelta è stata adottata per garantire la massima chiarezza nella visualizzazione di ciascuna distribuzione. In secondo luogo, per tutto il capitolo si utilizzerà il termine *match* come sinonimo di identificazione o corrispondenza tra vuoti di cataloghi diversi, intendendo una relazione biunivoca: una volta che un vuoto di un catalogo è stato identificato con uno dell'altro catalogo, nessuno dei due potrà essere identificato con un altro vuoto. In terzo luogo, le Analisi 1 e 2 non richiedono alcuna assunzione sulla forma dei vuoti, basandosi unicamente sulle distanze tra i centri e sui raggi effettivi. L'Analisi 3, invece, assume la sfericità dei vuoti per il calcolo dei volumi come  $\frac{4}{3}\pi R^3$  e per identificare i centri che cadono all'interno di tali volumi. Questa è una semplificazione che, sebbene garantisca una sufficiente validità dei risultati, costituisce comunque un'approssimazione della loro morfologia reale. In un lavoro futuro, o in un'analisi più raffinata, si dovrebbe considerare il bacino reale di sottodensità, con la propria forma effettiva, per ottenere una descrizione più accurata delle proprietà dei vuoti e delle loro distorsioni.

Un ulteriore aspetto metodologico, che sarà discusso più in dettaglio nella Sezione 6.4, riguarda la futura possibilità di combinare le tre metriche in un unico criterio di identificazione. Idealmente, si potrebbe assegnare a ogni coppia di vuoti (uno da ciascun catalogo) i valori delle tre metriche e selezionare le migliori combinazioni in base a un criterio globale, ad esempio richiedendo che una coppia soddisfi almeno due dei tre criteri oppure definendo una metrica composita che bilanci distanza spaziale e somiglianza in raggio. Questo approccio permetterebbe di identificare quelle coppie di vuoti che persistono indipendentemente dal metodo utilizzato, fornendo un campione di vuoti "robusti" alle distorsioni indotte dalle RSD. L'analisi di queste coppie persistenti con ogni metrica identificativa, così come lo studio dei vuoti che vengono identificati da un criterio ma non da un altro (e delle loro eventuali caratteristiche distintive in termini di posizione, dimensione o profilo di densità), rappresenta un interessante sviluppo futuro.

## 5.1 Analisi 1: metrica del primo vicino

La metrica del primo vicino definisce una corrispondenza biunivoca tra i due cataloghi: un vuoto del catalogo di riferimento viene identificato con il vuoto del secondo catalogo più vicino, a condizione che sia soddisfatta la disuguaglianza:

$$\frac{d}{R_1 + R_2} < s, \quad (5.1)$$

dove  $d$  è la distanza tra i centri,  $R_1$  e  $R_2$  sono i raggi dei due vuoti e  $s = 0.1$  è un parametro di soglia adimensionale. La distanza  $d$  tra i centri viene normalizzata rispetto alla somma dei raggi dei due vuoti,  $R_1 + R_2$ . Questa normalizzazione rende la metrica indipendente dalla scala assoluta dei vuoti e consente un confronto omogeneo tra cataloghi con diverse distribuzioni di raggi: la stessa distanza assoluta può essere significativa per vuoti piccoli ma irrilevante per vuoti grandi. Da questo punto in poi, quando si parlerà di distanza normalizzata, si intenderà sempre la quantità  $d/(R_1 + R_2)$  definita nell'(5.1).

La procedura di identificazione garantisce una corrispondenza biunivoca tra i due cataloghi: vengono calcolate tutte le possibili coppie di vuoti (uno da ciascun catalogo) che soddisfano il criterio della soglia. Le coppie vengono poi ordinate per distanza normalizzata crescente, dunque dalla migliore alla peggiore, e selezionate in quest'ordine, accettando una coppia solo se entrambi i vuoti coinvolti non sono già stati identificati in precedenza. Questa procedura non dipende dall'ordine di processamento dei vuoti e produce l'identificazione globalmente ottimale in termini di minimizzazione delle distanze normalizzate tra le coppie identificate, garantendo al contempo che ogni vuoto venga identificato al massimo una volta. La scelta del valore di  $s$  rappresenta un compromesso tra la selettività dell'identificazione e la robustezza statistica dell'analisi.

Con questa analisi abbiamo prodotto nove grafici, organizzati in tre blocchi da tre ciascuno: il primo blocco riguarda il confronto tra le due esecuzioni indipendenti di **BitVF** (Esecuzione1 ed Esecuzione2), il secondo il confronto tra spazio reale e spazio dei redshift per **BitVF**, e il terzo il confronto tra spazio reale e spazio dei redshift per **VIDE**. Per ciascun blocco sono stati generati un istogramma delle distanze normalizzate delle identificazioni, un istogramma delle differenze relative dei raggi e uno *scatter plot* dei raggi associati.

### 5.1.1 Esecuzione1 vs Esecuzione2: riproducibilità di BitVF

Come discusso nell'introduzione al Capitolo e nella Sezione 3.3.4, BitVF utilizza il parametro `seed` per la generazione del catalogo di posizioni lagrangiane iniziali e, quando impostato a 0, il seme viene generato automaticamente, introducendo una componente stocastica nell'identificazione dei vuoti. È quindi essenziale verificare che, a parità di parametri di input, l'algoritmo produca cataloghi consistenti e riproducibili.

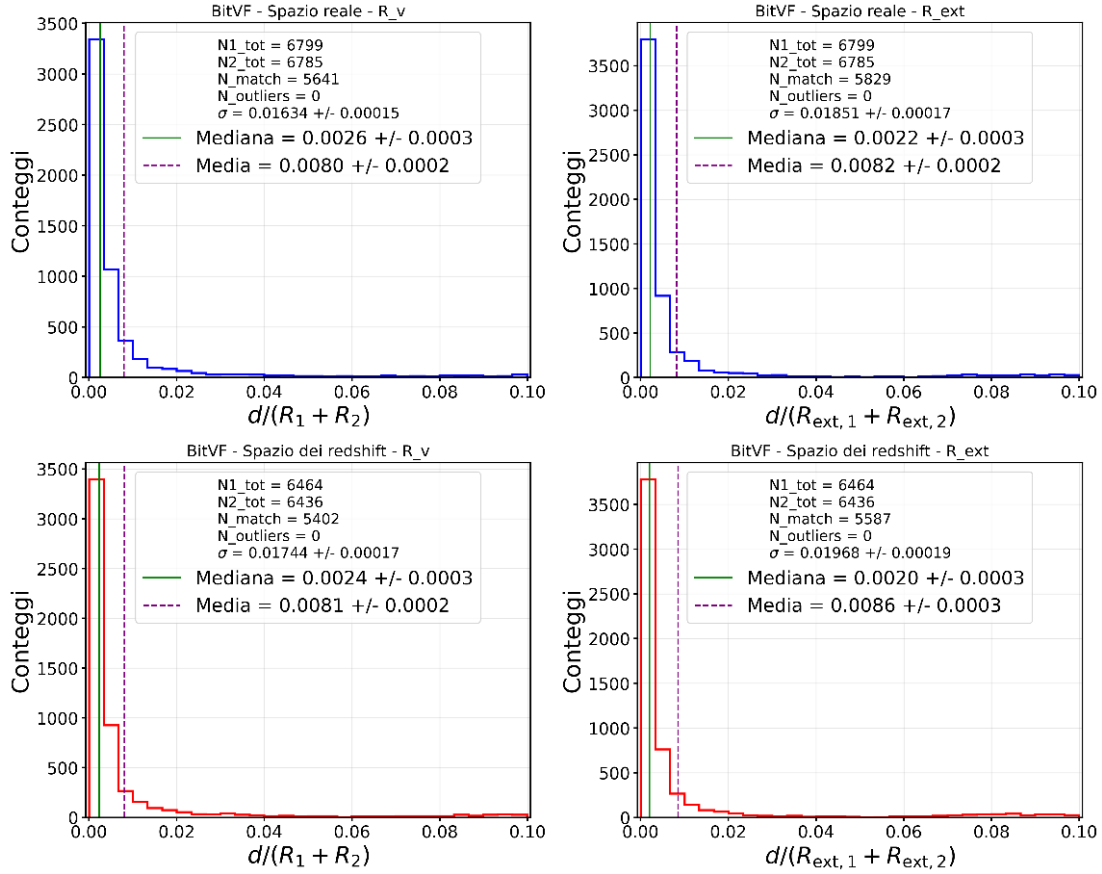


Figura 5.1: Istogramma delle distanze normalizzate  $d/(R_1 + R_2)$  per l'identificazione tra due esecuzioni di BitVF. I pannelli di sinistra e destra si riferiscono rispettivamente alle definizioni di raggio  $R_v$  e  $R_{ext}$ ; i pannelli superiori e inferiori corrispondono rispettivamente a spazio reale e dei redshift. Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,tot}$ ,  $N_{2,tot}$ ), il numero di identificazioni ( $N_{match}$ ), il numero di *outliers* ( $N_{outliers}$ ) e la deviazione standard  $\sigma$  della distribuzione. Le linee verticali indicano i valori mediano (verde, continua) e medio (viola, tratteggiata) della distribuzione.

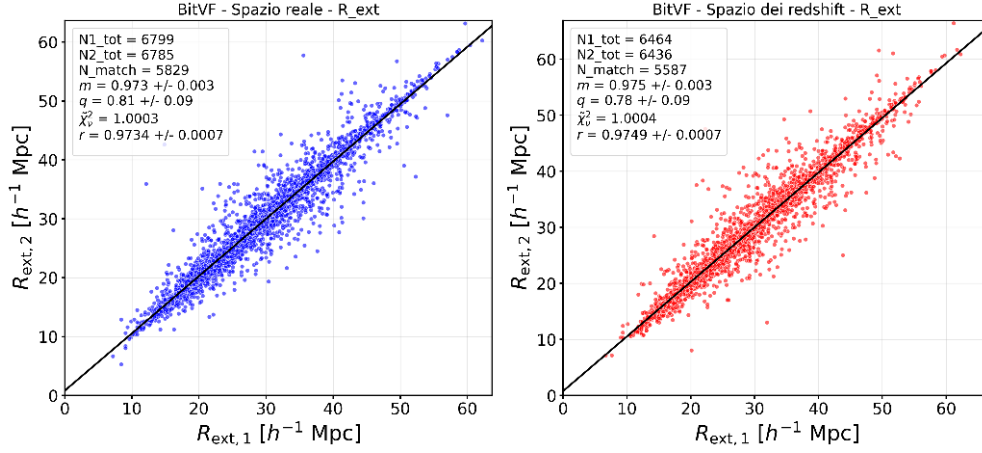


Figura 5.2: *Scatter plot* dei raggi dei vuoti identificati per l'identificazione tra due esecuzioni di BitVF. I pannelli di sinistra e destra mostrano rispettivamente spazio reale e dei redshift; per entrambi è utilizzata la definizione di raggio  $R_{\text{ext}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,\text{tot}}$ ,  $N_{2,\text{tot}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), la pendenza  $m$ , l'intercetta  $q$ , il  $\chi^2_v$  e il coefficiente di correlazione  $r$  del fit lineare. La linea continua rappresenta il miglior fit lineare.

Ottenute due esecuzioni indipendenti di BitVF con i parametri descritti nell'introduzione al Capitolo, si è passati al confronto fra di esse. La Fig. 5.1 mostra gli istogrammi delle distanze normalizzate rispetto alla somma dei raggi. Gli istogrammi presentano un picco pronunciato a valori prossimi allo zero, con valori mediani  $\text{Med}[d/(R_1 + R_2)] \sim 0.002$  e valori medi  $\langle d/(R_1 + R_2) \rangle \sim 0.008$  per entrambe le definizioni di raggio e per entrambi gli spazi. La mediana, in questo contesto, risulta più rappresentativa della tendenza centrale della distribuzione rispetto alla media, poiché la distribuzione delle distanze normalizzate è fortemente asimmetrica e presenta una coda verso valori più elevati: la presenza di poche identificazioni con distanze relativamente grandi (ma comunque al di sotto della soglia  $s = 0.1$ ) sposta la media verso valori più alti, mentre la mediana, essendo più robusta agli outlier, fornisce una stima più affidabile della tipica distanza di identificazione. Le frazioni di identificazione risultano superiori all'83% per entrambe le definizioni di raggio.

Lo *scatter plot* dei raggi, in Fig. 5.2, mostra un allineamento ottimo lungo la bisettrice, come testimoniato dai parametri del fit lineare: coefficienti di correlazione  $r \simeq 0.97$ ,  $\chi^2_v \simeq 1$  e pendenza  $m \simeq 0.97$  per entrambi gli spazi. Questi risultati confermano che le due esecuzioni identificano sostanzialmente gli stessi vuoti, con differenze trascurabili

attribuibili esclusivamente alle fluttuazioni statistiche introdotte dal seme casuale.

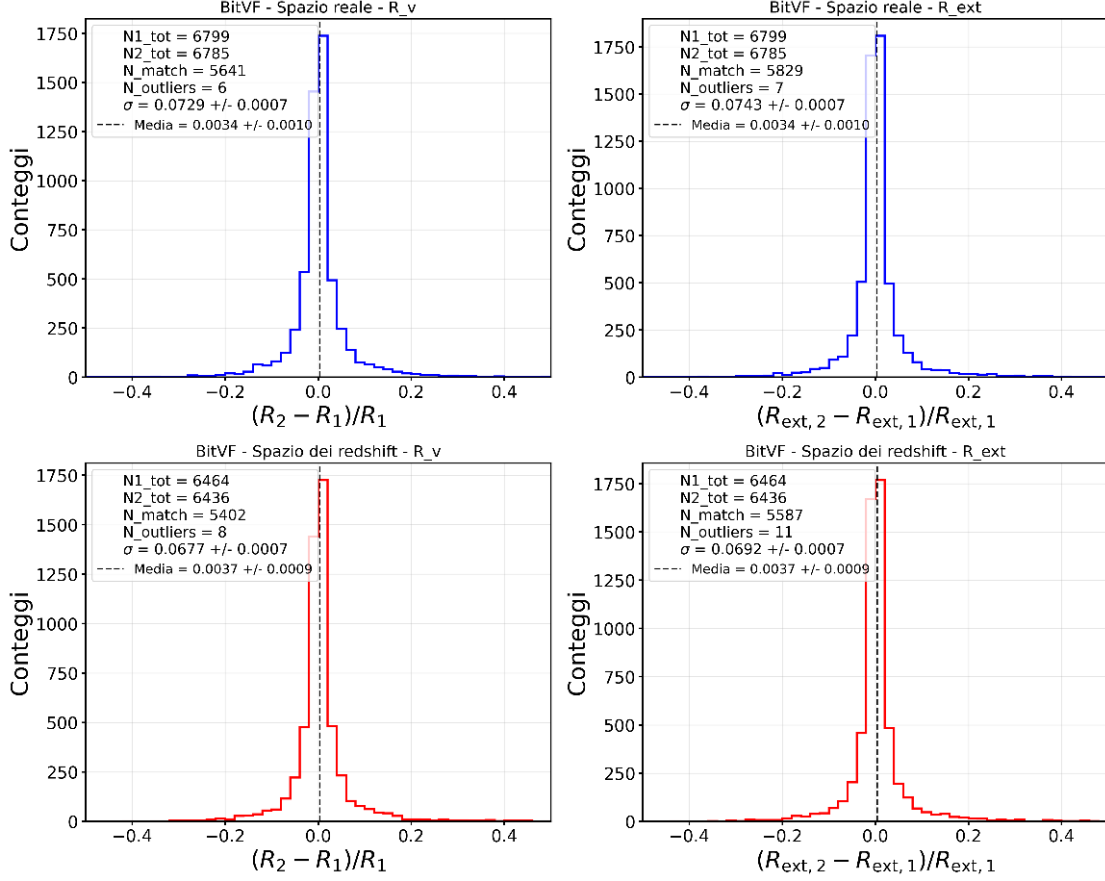


Figura 5.3: Distribuzione delle differenze relative dei raggi  $(R_2 - R_1)/R_1$  per i vuoti identificati tra le due esecuzioni. I pannelli di sinistra e destra si riferiscono rispettivamente alle definizioni di raggio  $R_v$  e  $R_{ext}$ ; i pannelli superiori e inferiori corrispondono rispettivamente a spazio reale e dei redshift. Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,tot}$ ,  $N_{2,tot}$ ), il numero di identificazioni ( $N_{match}$ ), il numero di *outliers* ( $N_{outliers}$ ) e la deviazione standard  $\sigma$  della distribuzione. La linea tratteggiata nera indica il valore medio della distribuzione.

La Fig. 5.3 mostra la distribuzione delle differenze relative dei raggi, che risulta centrata su zero con dispersione ridotta, confermando ulteriormente l'assenza di bias sistematici tra le due esecuzioni.

La Tabella 5.1 riporta le statistiche complete delle identificazioni.

Tabella 5.1: Statistiche delle identificazioni tra le due esecuzioni di **BitVF** con metrica del primo vicino.

| Spazio   | Def.             | $N_1$ | $N_2$ | Match | Fraz.              | $\text{Med}[d/(R_1 + R_2)]$ | $\sigma_{\text{Med}}$ | $\langle (R_2 - R_1)/R_1 \rangle$ | $\sigma_{\langle (R_2 - R_1)/R_1 \rangle}$ |
|----------|------------------|-------|-------|-------|--------------------|-----------------------------|-----------------------|-----------------------------------|--------------------------------------------|
| Reale    | $R_v$            | 6799  | 6785  | 5641  | $(83.1 \pm 0.5)\%$ | $0.0026 \pm 0.0003$         | $0.0163 \pm 0.0002$   | $0.0034 \pm 0.0010$               | $0.0729 \pm 0.0007$                        |
| Reale    | $R_{\text{ext}}$ | 6799  | 6785  | 5829  | $(85.8 \pm 0.4)\%$ | $0.0022 \pm 0.0003$         | $0.0185 \pm 0.0002$   | $0.0034 \pm 0.0010$               | $0.0743 \pm 0.0007$                        |
| Redshift | $R_v$            | 6464  | 6436  | 5402  | $(83.8 \pm 0.5)\%$ | $0.0024 \pm 0.0003$         | $0.0174 \pm 0.0002$   | $0.0037 \pm 0.0009$               | $0.0677 \pm 0.0007$                        |
| Redshift | $R_{\text{ext}}$ | 6464  | 6436  | 5587  | $(86.6 \pm 0.4)\%$ | $0.0020 \pm 0.0003$         | $0.0197 \pm 0.0002$   | $0.0037 \pm 0.0009$               | $0.0692 \pm 0.0007$                        |

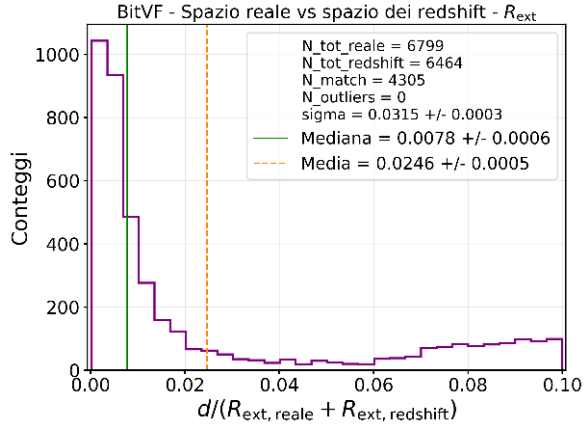
### 5.1.2 **BitVF** e **VIDE**: compatibilità tra spazio reale e dei redshift

Ricordiamo come l'obiettivo principale di questa analisi sia valutare, tra **BitVF** e **VIDE**, quale algoritmo produca cataloghi di vuoti nello spazio dei redshift più simili a quelli ottenuti nello spazio reale, in termini di: numero di vuoti, raggi, posizioni e proprietà statistiche. In particolare, si è voluto testare l'ipotesi, avanzata da Sartori et al. [2026], secondo cui **BitVF** mantiene maggiore coerenza fra i cataloghi individuati nei due spazi, rispetto a quanto fatto da **VIDE**.

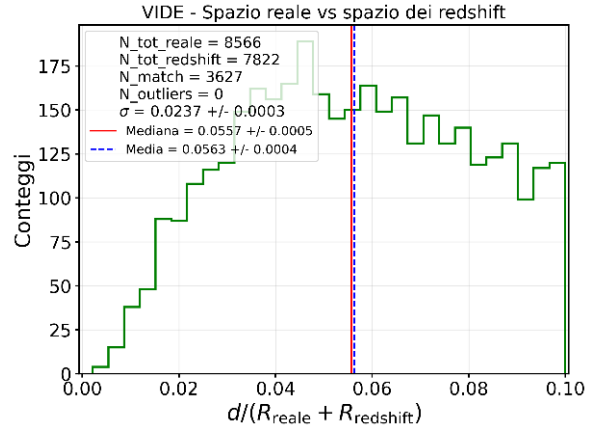
Come si evince dalla Fig. 5.4, la distribuzione delle distanze normalizzate per **BitVF** è estremamente piccata a zero, mentre per **VIDE** è notevolmente più ampia e spostata verso valori più elevati. Questa differenza è il risultato più rilevante emerso dall'analisi di validazione: **BitVF** identifica i centri dei vuoti nello spazio dei redshift con una precisione straordinariamente maggiore rispetto a **VIDE**, come discusso in dettaglio nel seguito.

La Fig. 5.4 mostra gli istogrammi delle distanze normalizzate. Per **BitVF**, la distribuzione è estremamente piccata a valori prossimi allo zero, con valori mediani  $\text{Med}[d/(R_{\text{reale}} + R_{\text{redshift}})] = (0.0078 \pm 0.0006)$  e valori medi  $\langle d/(R_{\text{reale}} + R_{\text{redshift}}) \rangle = (0.0246 \pm 0.0005)$ . Per **VIDE**, invece, la distribuzione è notevolmente più ampia e spostata verso valori più elevati, con mediana  $\text{Med}[d/(R_{\text{reale}} + R_{\text{redshift}})] = (0.0557 \pm 0.0005)$  e media  $\langle d/(R_{\text{reale}} + R_{\text{redshift}}) \rangle = (0.0563 \pm 0.0004)$ . Questo indica che i centri dei vuoti di **VIDE** subiscono spostamenti spaziali significativamente maggiori tra spazio reale e dei redshift rispetto a quelli di **BitVF**. In particolare, la differenza tra le mediane dei due algoritmi è di un fattore  $\sim 7$ , a testimonianza della straordinaria precisione con cui **BitVF** identifica i centri dei vuoti nello spazio dei redshift. Questo risultato, come si vedrà nel Capitolo 6, costituisce il risultato più interessante e rilevante emerso dall'intera analisi di validazione.

La Fig. 5.5 mostra le distribuzioni delle differenze relative dei raggi  $(R_{\text{redshift}} - R_{\text{reale}})/R_{\text{reale}}$  per i vuoti identificati tra spazio reale e spazio dei redshift. Per **BitVF**,

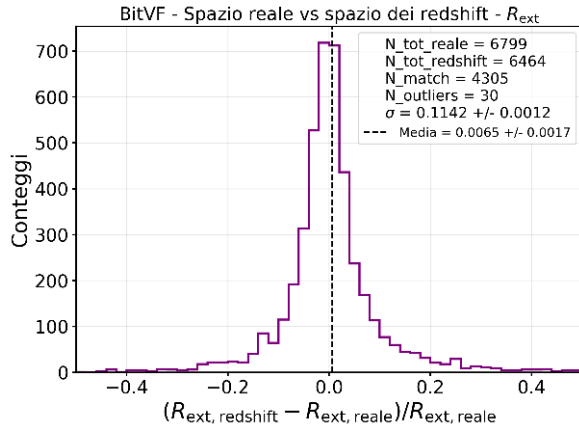


(a) BitVF - Istogramma delle distanze normalizzate per  $R_{\text{ext}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. Le linee verticali indicano i valori mediano (verde, continua), medio (arancione, tratteggiata) e soglia (nera, tratteggiata).

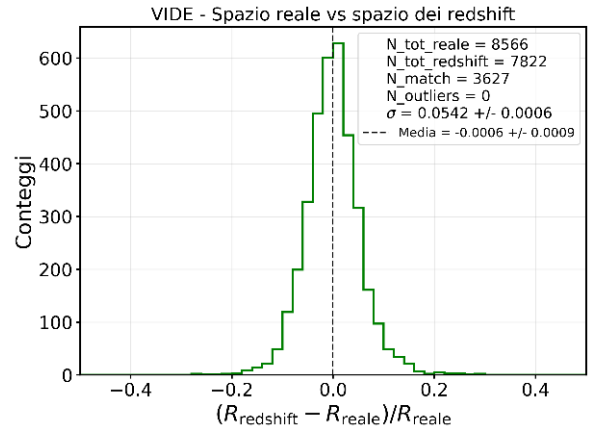


(b) VIDE - Istogramma delle distanze normalizzate. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. Le linee verticali indicano i valori mediano (rosso, continua), medio (blu, tratteggiata) e soglia (nera, tratteggiata).

Figura 5.4: Istogrammi delle distanze normalizzate  $d/(R_{\text{reale}} + R_{\text{redshift}})$  per i vuoti identificati tra spazio reale e dei redshift con la metrica del primo vicino.



(a) **BitVF** - Distribuzione delle differenze relative dei raggi per  $R_{\text{ext}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. La linea tratteggiata nera indica il valore medio della distribuzione.



(b) **VIDE** - Distribuzione delle differenze relative dei raggi. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. La linea tratteggiata nera indica il valore medio della distribuzione.

Figura 5.5: Distribuzioni delle differenze relative dei raggi  $(R_{\text{redshift}} - R_{\text{reale}})/R_{\text{reale}}$  tra spazio reale e dei redshift.

la distribuzione è centrata su un valore prossimo a zero ma presenta una dispersione più ampia ( $\sigma \simeq 0.11$ ), consistente con la maggiore popolazione di match inclusi: l'algoritmo dinamico tende a mantenere un numero più elevato di corrispondenze, includendo anche coppie di vuoti i cui raggi subiscono variazioni lievi ma non trascurabili nel passaggio da uno spazio all'altro. Questa maggiore inclusività suggerisce che **BitVF** è in grado di identificare i vuoti nonostante gli effetti RSD con una copertura più estesa del campione, sebbene a discapito di una precisione puntuale nella corrispondenza dei raggi.

Per **VIDE**, invece, la distribuzione è molto più stretta ( $\sigma \simeq 0.05$ ) e centrata su zero, suggerendo una maggiore selettività dell'identificazione. L'algoritmo geometrico tende a identificare preferenzialmente vuoti con raggi molto simili tra i due spazi, a discapito di una frazione di identificazioni inferiore (come evidenziato in Tabella 5.2) e di una minore precisione nelle corrispondenze centro-centro. Questo comportamento riflette la diversa natura dei due algoritmi: **VIDE**, basato sulla tassellazione di Voronoi e sul *watershed*, identifica vuoti la cui forma e dimensione sono strettamente legate alla struttura locale del campo di densità, rendendo le corrispondenze più stringenti in termini di raggio ma meno precise dal punto di vista posizionale. **BitVF**, basato sulla dinamica delle particelle e sulla ricostruzione del campo di spostamento, risulta meno sensibile alle distorsioni indotte dalle RSD, consentendo di mantenere un campione più ampio di vuoti tra i due spazi e una maggiore precisione nelle corrispondenze centro-centro.

La Fig. 5.6 mostra gli *scatter plot* dei raggi. Per **BitVF**, utilizzando la definizione di raggio  $R_{\text{ext}}$ , la pendenza del fit lineare è  $m = (0.971 \pm 0.006)$ ; per **VIDE**, invece, si ottiene  $m = (1.025 \pm 0.003)$ , indicando che, in questo caso, i raggi dei vuoti nello spazio dei redshift tendono ad essere sistematicamente più ampi rispetto a quelli nello spazio reale.

La Tabella 5.2 riporta le statistiche di identificazione per entrambi i cataloghi.

Tabella 5.2: Statistiche di identificazione tra spazio reale e dei redshift per **BitVF** e **VIDE** con metrica del primo vicino. Per **BitVF** è riportato il risultato per la definizione di raggio esteso  $R_{\text{ext}}$ .

| Catalogo     | Def.             | $N_{\text{real}}$ | $N_{\text{redshift}}$ | Match | Fraz.              | Med[ $d/(R_1 + R_2)$ ] | $\sigma_{\text{Med}}$ | $\langle \Delta R/R \rangle$ | $\sigma_{\Delta R/R}$ |
|--------------|------------------|-------------------|-----------------------|-------|--------------------|------------------------|-----------------------|------------------------------|-----------------------|
| <b>BitVF</b> | $R_{\text{ext}}$ | 6799              | 6464                  | 4305  | $(64.9 \pm 0.6)\%$ | $0.0078 \pm 0.0006$    | $0.0315 \pm 0.0003$   | $0.0065 \pm 0.0017$          | $0.1142 \pm 0.0012$   |
| <b>VIDE</b>  | $R$              | 8566              | 7822                  | 3627  | $(44.3 \pm 0.5)\%$ | $0.0557 \pm 0.0005$    | $0.0237 \pm 0.0003$   | $-0.0006 \pm 0.0009$         | $0.0542 \pm 0.0006$   |

I risultati rivelano delle differenze nel comportamento dei due cataloghi. **BitVF** mostra una frazione di identificazioni più elevata (64.9%) rispetto a **VIDE** (44.3%), indican-

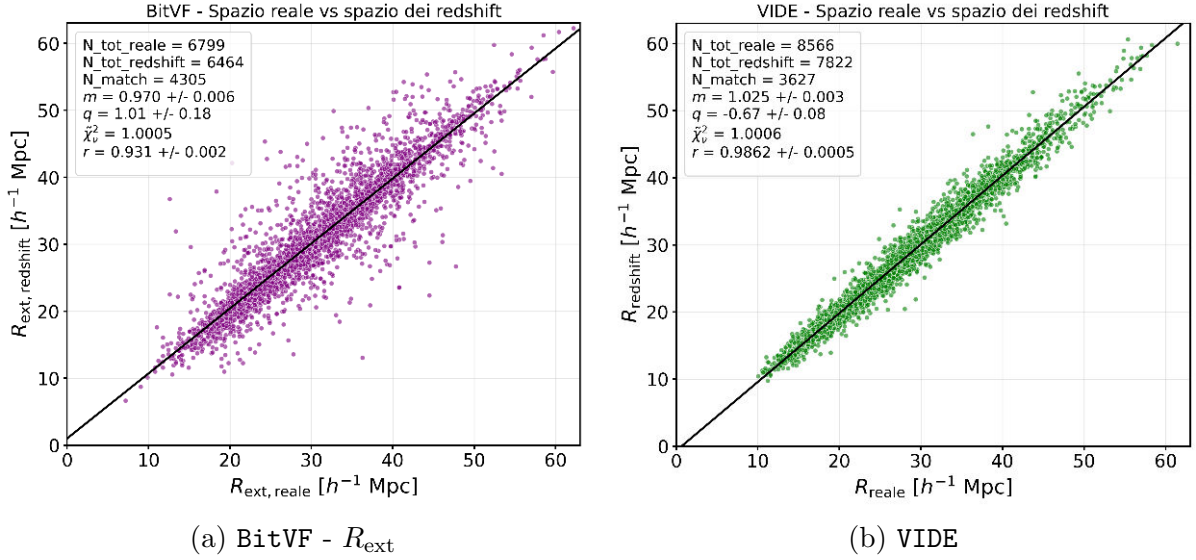


Figura 5.6: *Scatter plot* dei raggi dei vuoti identificati tra i due spazi. La linea continua rappresenta il miglior fit lineare; nelle rispettive legende sono riportate pendenza, intercetta,  $\tilde{\chi}^2_\nu$  e coefficiente di correlazione.

do una maggiore corrispondenza spaziale tra i vuoti identificati nei due spazi. Inoltre, le distanze normalizzate per BitVF sono significativamente più piccole ( $\text{Med}[d/(R_{\text{real}} + R_{\text{redshift}})] \simeq 0.008$ ) rispetto a VIDE ( $\simeq 0.056$ ), confermando che BitVF traccia i vuoti tra i due spazi con una precisione spaziale superiore. La distribuzione delle differenze relative per VIDE risulta più stretta ( $\sigma_{\Delta R/R} \simeq 0.05$  contro  $\simeq 0.11$  per BitVF), riflettendo una maggiore selettività dell'identificazione, come già discusso.

## 5.2 Analisi 2: metrica della somiglianza di raggio con filtro spaziale

L'Analisi 2 introduce una metrica di identificazione alternativa, basata sulla somiglianza di raggio piuttosto che sulla distanza tra i centri. Per ogni vuoto del catalogo di riferimento, con raggio  $R_{\text{ref}}$ , si cercano tutti i centri nel catalogo target, con raggi  $R_{\text{target}}$ , che cadono all'interno del raggio del vuoto ( $d \leq R_{\text{ref}}$ ). Tra questi, si seleziona quello che minimizza la differenza dei raggi:

$$\Delta R = |R_{\text{target}} - R_{\text{ref}}|. \quad (5.2)$$

La presenza del filtro spaziale  $d \leq R_{\text{ref}}$  garantisce che la somiglianza di raggio venga valutata solo tra vuoti che sono spazialmente vicini, evitando accoppiamenti tra vuoti lontani che potrebbero avere raggi simili ma che non corrispondono fisicamente. Questa metrica risponde quindi all'esigenza di valutare la corrispondenza tra vuoti basandosi sulla combinazione di posizione spaziale e dimensione.

Anche in questo caso la procedura garantisce una corrispondenza biunivoca tra i due cataloghi: vengono calcolate tutte le possibili coppie che soddisfano il filtro spaziale, ordinate per differenza di raggio crescente, e selezionate in quest'ordine assicurando che ogni vuoto venga identificato al massimo una volta. Questa procedura è ottimale e non dipende dall'ordine di processamento dei vuoti.

Questa analisi produce esattamente gli stessi grafici della precedente, ma con criterio di identificazione differente. Come prima, i nove grafici sono organizzati in tre blocchi da tre ciascuno: il primo blocco riguarda il confronto tra le due esecuzioni indipendenti di **BitVF** (Esecuzione1 vs Esecuzione2), il secondo il confronto tra spazio reale e dei redshift per **BitVF**, e il terzo il confronto tra spazio reale e dei redshift per **VIDE**. Per ciascun blocco vengono generati: un istogramma delle distanze normalizzate delle identificazioni, uno *scatter plot* dei raggi associati e un istogramma delle differenze relative dei raggi. Per quanto riguarda **BitVF**, si riportano esclusivamente i risultati ottenuti con la definizione di raggio esteso  $R_{\text{ext}}$ ; la validità di questa scelta è già stata discussa all'inizio del capitolo.

### 5.2.1 Esecuzione1 vs Esecuzione2: riproducibilità di **BitVF**

Applicando il criterio di somiglianza di raggio con filtro spaziale alle due esecuzioni indipendenti di **BitVF**, si ottengono i risultati mostrati nelle figure seguenti.

La Fig. 5.7 mostra gli istogrammi delle distanze normalizzate per le quattro combinazioni di spazio e raggio di **BitVF**. Le distribuzioni sono estremamente piccate a valori prossimi allo zero, con valori mediani  $\text{Med}[d/(R_1 + R_2)] \simeq 0.003$  e valori medi  $\langle d/(R_1 + R_2) \rangle \simeq 0.030$ , dunque più alti di quelli mediani di un fattore 10, per tutte le combinazioni, in buon accordo con i valori riportati in Tabella 5.3. Come già discusso per l'Analisi 1, la mediana risulta più rappresentativa della tipica distanza di identificazione rispetto alla media in questo contesto, poiché quest'ultima è sensibile alla coda della

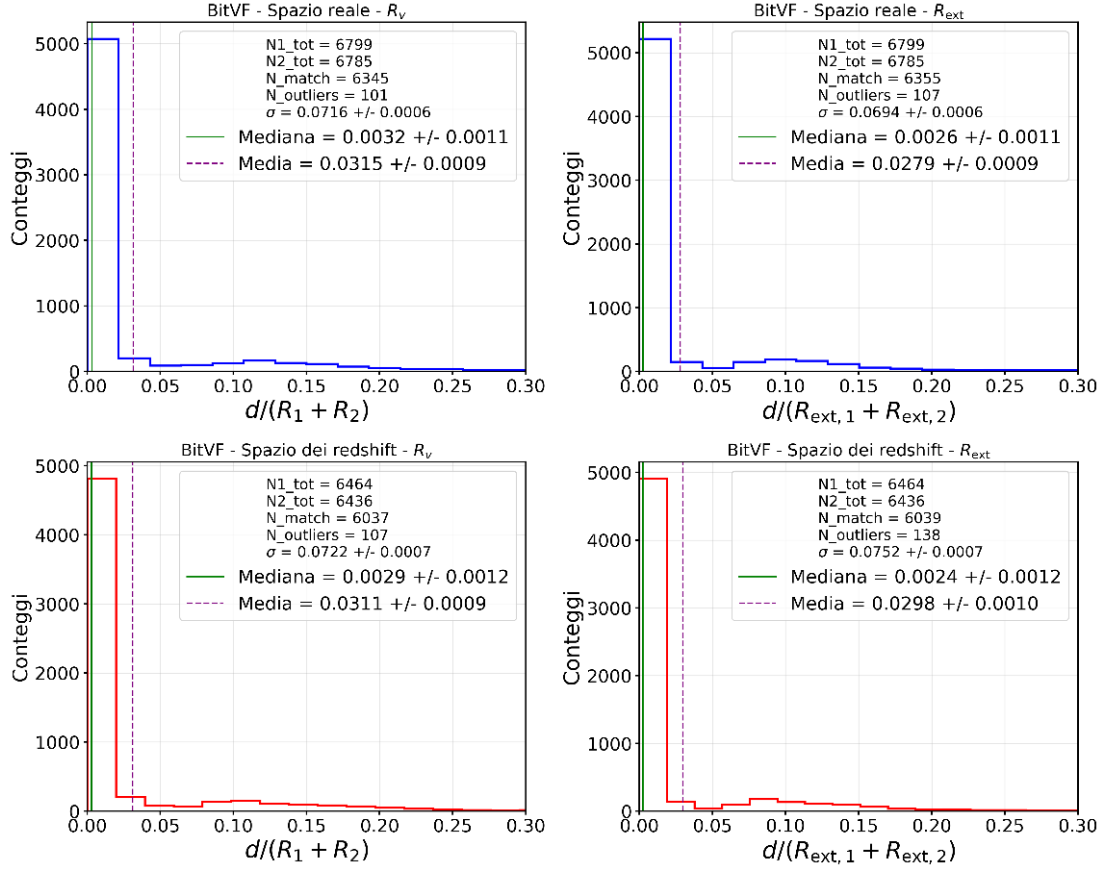


Figura 5.7: Istogramma delle distanze normalizzate  $d/(R_1 + R_2)$  per l'identificazione per somiglianza di raggio con filtro spaziale tra due esecuzioni di BitVF. I quattro pannelli mostrano le distribuzioni per le combinazioni di spazio (reale e dei redshift) e definizione di raggio ( $R_v$  e  $R_{ext}$ ). Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,tot}$ ,  $N_{2,tot}$ ), il numero di identificazioni ( $N_{match}$ ), il numero di *outliers* ( $N_{outliers}$ ) e la deviazione standard  $\sigma$  della distribuzione. Le linee verticali indicano i valori mediano (verde, continua) e medio (viola, tratteggiata) della distribuzione.

distribuzione verso valori più elevati. Questo sostiene l'ipotesi che la riproducibilità di BitVF sia robusta rispetto alla metrica di identificazione utilizzata.

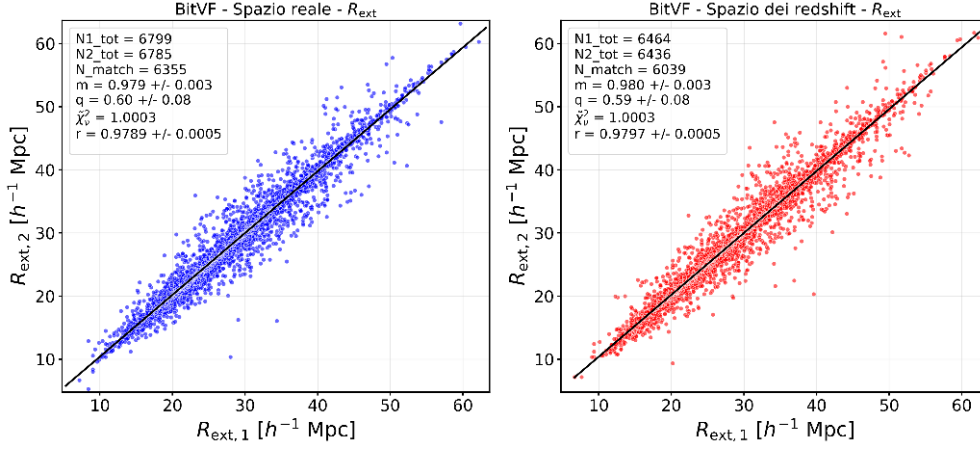


Figura 5.8: *Scatter plot* dei raggi estesi  $R_{\text{ext}}$  per l'identificazione per somiglianza di raggio con filtro spaziale tra due esecuzioni di BitVF. I pannelli di sinistra e destra mostrano rispettivamente spazio reale e dei redshift. Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,\text{tot}}$ ,  $N_{2,\text{tot}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), la pendenza  $m$ , l'intercetta  $q$ , il  $\tilde{\chi}^2_{\nu}$  e il coefficiente di correlazione  $r$  del fit lineare. La linea continua rappresenta il miglior fit lineare.

La Fig. 5.8 mostra gli *scatter plot* dei raggi estesi per le identificazioni in spazio reale e dei redshift. In entrambi i casi, i punti sono estremamente allineati lungo la bisettrice, con coefficienti di correlazione  $r \simeq 0.98$ , pendenze  $m \simeq 0.98$  e  $\tilde{\chi}^2_{\nu} \simeq 1$ , indicando un ottimo accordo con il modello lineare.

La Fig. 5.9 mostra la distribuzione delle differenze relative dei raggi. In tutti i casi, le distribuzioni risultano centrate su zero, con valori medi  $\langle (R_2 - R_1)/R_1 \rangle \simeq 0.002$  e deviazioni standard  $\sigma \simeq 0.06$ , confermando l'assenza di bias sistematici tra le due esecuzioni.

La Tabella 5.3 riporta le statistiche principali dell'analisi.

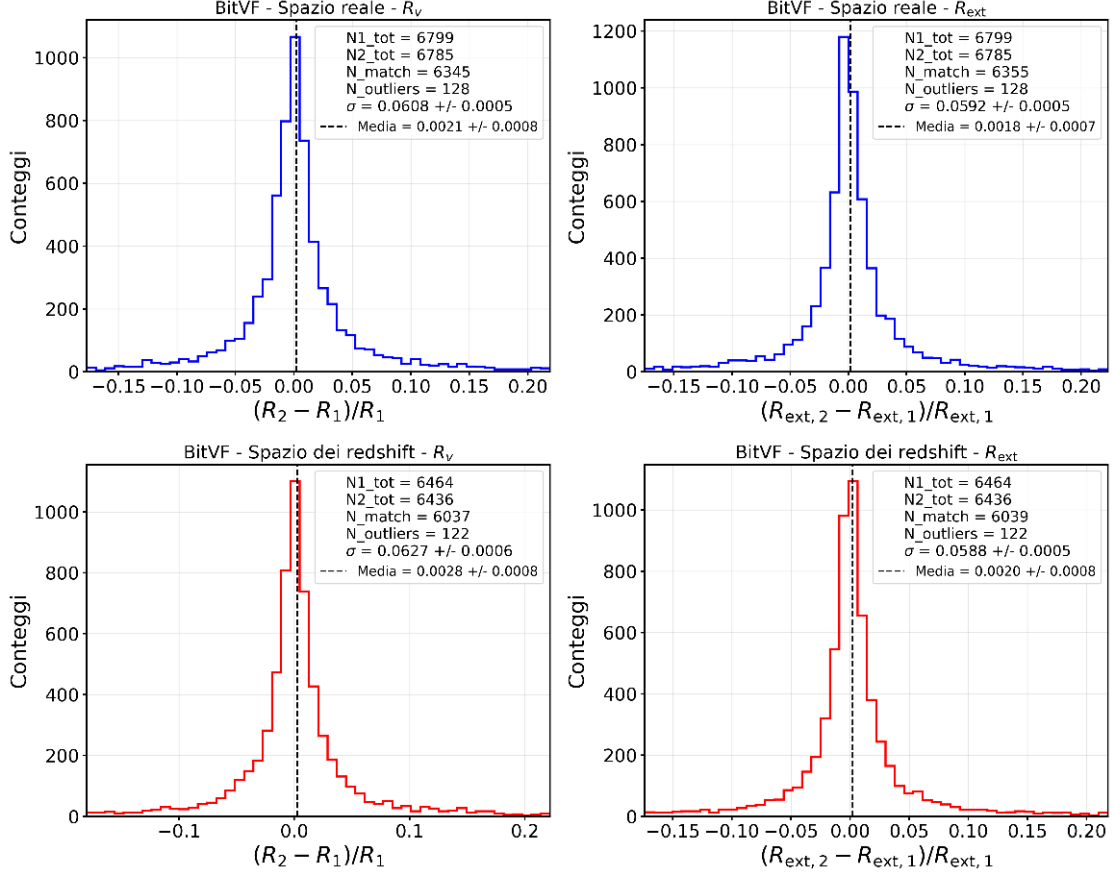


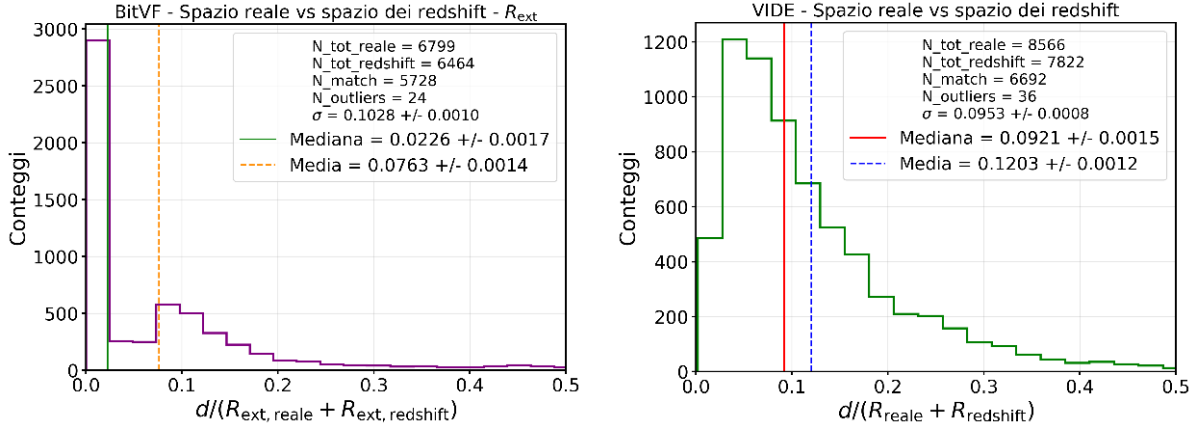
Figura 5.9: Distribuzione delle differenze relative dei raggi  $(R_2 - R_1)/R_1$  per l'identificazione per somiglianza di raggio con filtro spaziale tra le due esecuzioni. I quattro pannelli mostrano le distribuzioni per le combinazioni di spazio (reale e dei redshift) e definizione di raggio ( $R_v$  e  $R_{ext}$ ). Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,tot}$ ,  $N_{2,tot}$ ), il numero di identificazioni ( $N_{match}$ ), il numero di *outliers* ( $N_{outliers}$ ) e la deviazione standard  $\sigma$  della distribuzione. La linea tratteggiata nera indica il valore medio della distribuzione.

Tabella 5.3: Statistiche delle identificazioni tra le due esecuzioni di BitVF con metrica della somiglianza di raggio con filtro spaziale.

| Spazio   | Def.      | $N_1$ | $N_2$ | Match | Fraz.              | $Med[d/(R_1 + R_2)]$ | $\sigma_{Med}$      | $\langle (R_2 - R_1)/R_1 \rangle$ | $\sigma_{\langle (R_2 - R_1)/R_1 \rangle}$ |
|----------|-----------|-------|-------|-------|--------------------|----------------------|---------------------|-----------------------------------|--------------------------------------------|
| Reale    | $R_v$     | 6799  | 6785  | 6345  | $(93.4 \pm 0.3)\%$ | $0.0032 \pm 0.0011$  | $0.0716 \pm 0.0006$ | $0.0021 \pm 0.0008$               | $0.0608 \pm 0.0005$                        |
| Reale    | $R_{ext}$ | 6799  | 6785  | 6355  | $(93.6 \pm 0.3)\%$ | $0.0026 \pm 0.0011$  | $0.0694 \pm 0.0006$ | $0.0018 \pm 0.0007$               | $0.0592 \pm 0.0005$                        |
| Redshift | $R_v$     | 6464  | 6436  | 6037  | $(93.6 \pm 0.3)\%$ | $0.0029 \pm 0.0012$  | $0.0722 \pm 0.0007$ | $0.0028 \pm 0.0008$               | $0.0627 \pm 0.0006$                        |
| Redshift | $R_{ext}$ | 6464  | 6436  | 6039  | $(93.6 \pm 0.3)\%$ | $0.0024 \pm 0.0012$  | $0.0752 \pm 0.0007$ | $0.0020 \pm 0.0008$               | $0.0588 \pm 0.0005$                        |

### 5.2.2 BitVF e VIDE: compatibilità tra spazio reale e dei redshift

Applicando il criterio di somiglianza di raggio con filtro spaziale al confronto tra spazio reale e dei redshift per BitVF e VIDE, si ottengono i risultati mostrati nelle figure seguenti. Per BitVF si utilizza esclusivamente la definizione di raggio esteso  $R_{\text{ext}}$ .



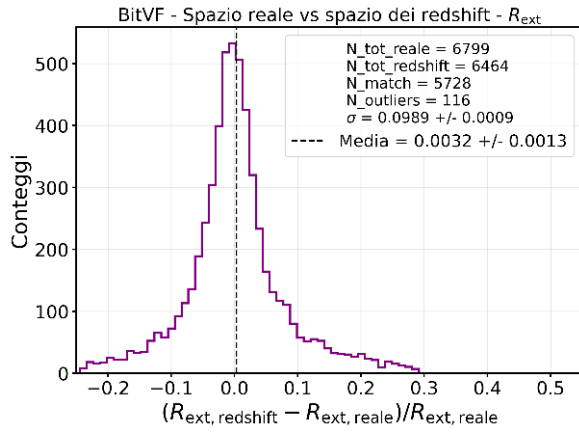
(a) BitVF - Istogramma delle distanze normalizzate per  $R_{\text{ext}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. Le linee verticali indicano i valori mediano (verde, continua) e medio (arancione, tratteggiata) della distribuzione.

(b) VIDE - Istogramma delle distanze normalizzate. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. Le linee verticali indicano i valori mediano (rosso, continua) e medio (blu, tratteggiata) della distribuzione.

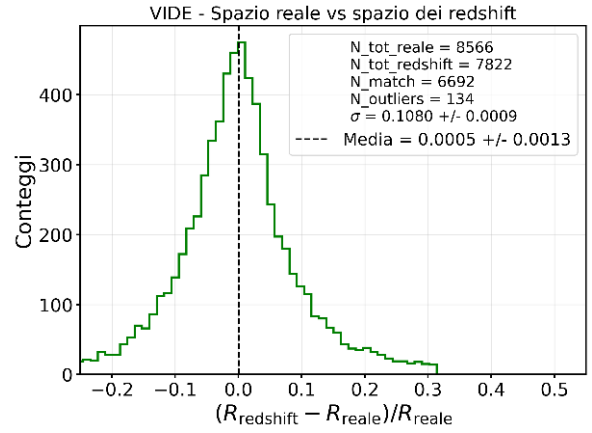
Figura 5.10: Istogrammi delle distanze normalizzate  $d/(R_{\text{reale}} + R_{\text{redshift}})$  per l'identificazione per somiglianza di raggio con filtro spaziale tra spazio reale e dei redshift.

La Fig. 5.10 mostra gli istogrammi delle distanze normalizzate. Per BitVF, la distribuzione è piccata a zero, con valori mediani  $\text{Med}[d/(R_{\text{reale}} + R_{\text{redshift}})] = (0.0226 \pm 0.0017)$  e valori medi  $\langle d/(R_{\text{reale}} + R_{\text{redshift}}) \rangle = (0.0763 \pm 0.0014)$ . Per VIDE, la distribuzione è più ampia e spostata verso valori più elevati, con mediana  $\text{Med}[d/(R_{\text{reale}} + R_{\text{redshift}})] = (0.0921 \pm 0.0015)$  e media  $\langle d/(R_{\text{reale}} + R_{\text{redshift}}) \rangle = (0.1203 \pm 0.0012)$ , indicando che anche con il criterio di somiglianza di raggio con filtro spaziale, BitVF mantiene una maggiore precisione spaziale.

La Fig. 5.11 mostra le distribuzioni delle differenze relative dei raggi, complessivamente analoghe a quelle dell'analisi precedente.



(a) **BitVF** - Differenze relative dei raggi per  $R_{\text{ext}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. La linea tratteggiata nera indica il valore medio della distribuzione.



(b) **VIDE** - Differenze relative dei raggi. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), il numero di *outliers* ( $N_{\text{outliers}}$ ) e la deviazione standard  $\sigma$  della distribuzione. La linea tratteggiata nera indica il valore medio della distribuzione.

Figura 5.11: Distribuzione delle differenze relative dei raggi  $(R_{\text{redshift}} - R_{\text{reale}})/R_{\text{reale}}$  per l'identificazione per somiglianza di raggio con filtro spaziale.

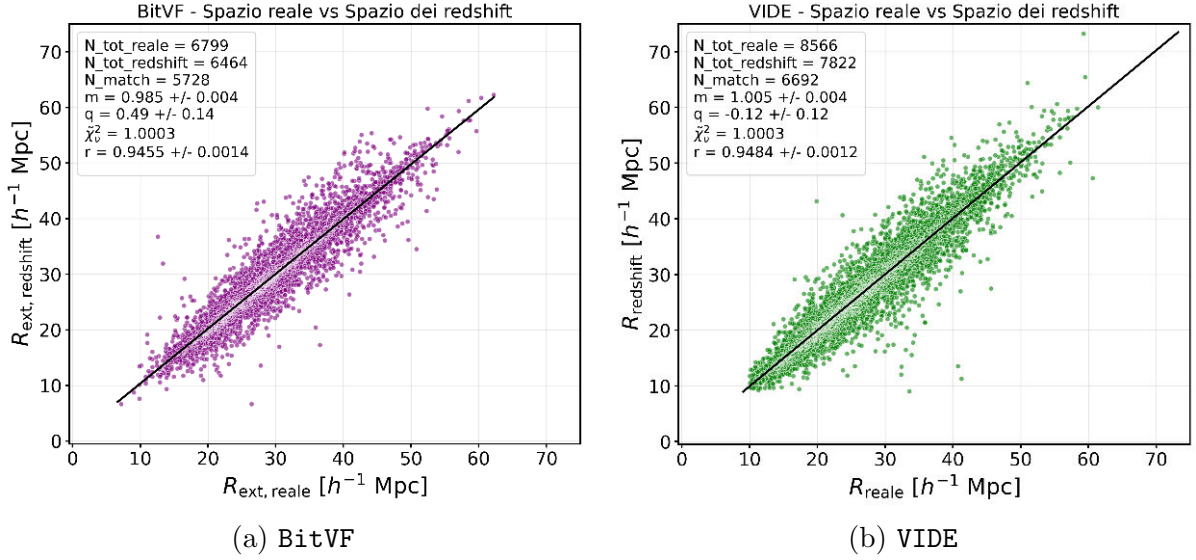


Figura 5.12: *Scatter plot* dei raggi dei vuoti identificati tra i due spazi. I pannelli di sinistra e destra mostrano rispettivamente BitVF, con definizione di raggio  $R_{\text{ext}}$ , e VIDE. Nelle legende sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di identificazioni ( $N_{\text{match}}$ ), la pendenza  $m$ , l'intercetta  $q$ , il  $\tilde{\chi}^2_v$  e il coefficiente di correlazione  $r$  del fit lineare. La linea continua rappresenta il miglior fit lineare.

La Fig. 5.12 mostra gli *scatter plot*. Per BitVF, utilizzando il raggio esteso  $R_{\text{ext}}$ , la pendenza del fit è  $m = (0.985 \pm 0.004)$ ; per VIDE è  $m = (1.005 \pm 0.004)$ , compatibile con l'unità entro  $1.3\sigma$ .

La Tabella 5.4 riporta le statistiche complete dell'analisi.

Tabella 5.4: Statistiche di identificazione tra spazio reale e dei redshift per BitVF e VIDE con metrica della somiglianza di raggio con filtro spaziale.

| Catalogo | Def.             | $N_{\text{real}}$ | $N_{\text{redshift}}$ | Match | Fraz.              | $\text{Med}[d/(R_1 + R_2)]$ | $\sigma_{\text{Med}}$ | $\langle \Delta R/R \rangle$ | $\sigma_{\Delta R/R}$ |
|----------|------------------|-------------------|-----------------------|-------|--------------------|-----------------------------|-----------------------|------------------------------|-----------------------|
| BitVF    | $R_{\text{ext}}$ | 6799              | 6464                  | 5728  | $(86.4 \pm 0.4)\%$ | $0.0226 \pm 0.0017$         | $0.1028 \pm 0.0010$   | $0.0032 \pm 0.0013$          | $0.0989 \pm 0.0009$   |
| VIDE     | $R$              | 8566              | 7822                  | 6692  | $(81.7 \pm 0.4)\%$ | $0.0921 \pm 0.0015$         | $0.0953 \pm 0.0008$   | $0.0005 \pm 0.0013$          | $0.1080 \pm 0.0009$   |

### 5.3 Analisi 3: mappatura dei volumi

L'Analisi 3 introduce un approccio complementare alle precedenti. Mentre le Analisi 1 e 2 cercano di identificare specifiche coppie di vuoti (uno per catalogo) che si corrispondono,

l'Analisi 3 si concentra sulla conservazione del volume: per ogni vuoto nel primo catalogo, detto di riferimento, si considerano tutti i centri del secondo catalogo, detto target, che si trovano all'interno del suo volume, assunto sferico, e si calcola il rapporto tra il volume del vuoto di riferimento e la somma dei volumi dei vuoti target che vi cadono dentro:

$$\mathcal{R} = \frac{V_{\text{ref}}}{\sum V_{\text{target}}}, \quad (5.3)$$

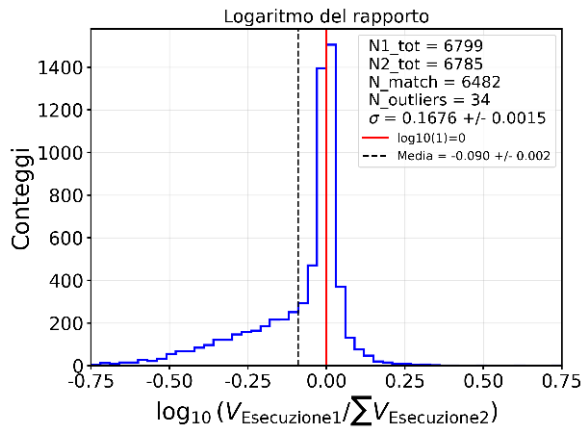
dove  $V_{\text{ref}} = \frac{4}{3}\pi R_{\text{ref}}^3$  e  $V_{\text{target}} = \frac{4}{3}\pi R_{\text{target}}^3$ .

L'interpretazione di  $\mathcal{R}$  è la seguente: se  $\mathcal{R} = 1$ , il volume di riferimento è perfettamente riprodotto dalla somma dei volumi dei vuoti nello spazio dei redshift che vi cadono dentro; se  $\mathcal{R} > 1$ , i vuoti nello spazio dei redshift sono più piccoli e frammentati (occorrono più vuoti per coprire lo stesso volume); se  $\mathcal{R} < 1$ , i vuoti in spazio dei redshift sono più grandi o si sovrappongono (la somma dei volumi supera il volume di riferimento). In questo senso, l'Analisi 3 non fornisce una corrispondenza biunivoca, ma piuttosto una misura di come il volume di un vuoto venga mappato dall'insieme dei vuoti nell'altro catalogo.

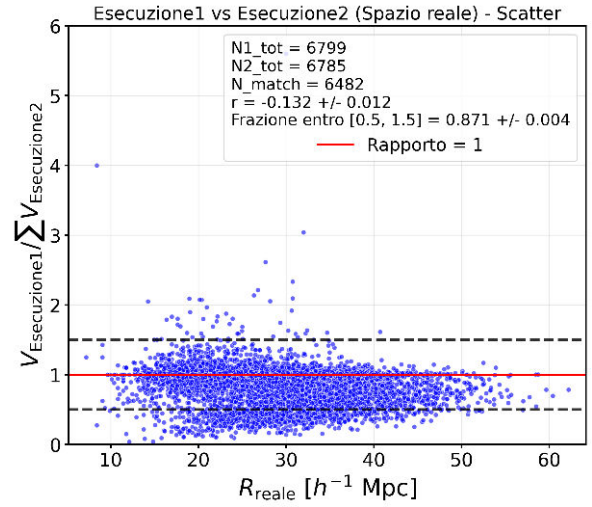
Con questa analisi abbiamo prodotto sei grafici, organizzati in tre blocchi da due ciascuno: il primo blocco riguarda il test di coerenza della mappatura dei volumi tra le due esecuzioni indipendenti di **BitVF** in spazio reale; il secondo quella tra spazio reale e dei redshift per una esecuzione di **BitVF**; il terzo quella tra spazio reale e dei redshift per una esecuzione di **VIDE**. Per ciascun blocco vengono generati un istogramma dei rapporti di volume e uno *scatter plot* del rapporto in funzione del raggio di riferimento. Per le motivazioni già discusse, si è scelto di utilizzare unicamente la definizione di raggio  $R_{\text{ext}}$  per le analisi condotte su qualsiasi esecuzione di **BitVF**.

### 5.3.1 Esecuzione1 vs Esecuzione2: coerenza della mappatura dei volumi in spazio reale

In primo luogo, si è applicata la mappatura dei volumi alle due esecuzioni indipendenti di **BitVF** in spazio reale. In questo caso, ci si aspetta che per ogni vuoto della Esecuzione1, la somma dei volumi dei vuoti della Esecuzione2 che cadono dentro il suo raggio sia circa uguale al volume del vuoto stesso, poiché le due esecuzioni dovrebbero identificare gli stessi vuoti. Il rapporto  $\mathcal{R}$  dovrebbe quindi essere prossimo a 1.



(a) Istogramma logaritmico dei rapporti di volume  $\mathcal{R} = V_{\text{Esecuzione1}} / \sum V_{\text{Esecuzione2}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,\text{tot}}$ ,  $N_{2,\text{tot}}$ ), il numero di match ( $N_{\text{match}}$ ), il valore medio  $\langle \mathcal{R} \rangle$ , la deviazione standard  $\sigma_{\mathcal{R}}$  e la frazione di punti entro  $[0.5, 1.5]$ . Le linee verticali indicano il valore  $\mathcal{R} = 1$  (rossa continua) e il valore medio (nera tratteggiata) della distribuzione.



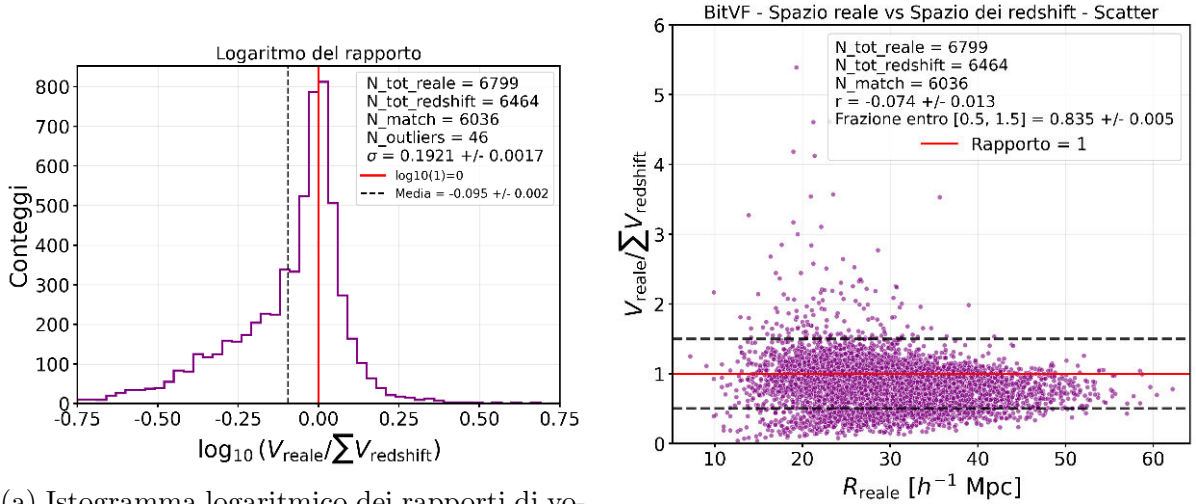
(b) *Scatter plot* del rapporto  $\mathcal{R}$  in funzione del raggio di riferimento  $R_{\text{ref}}$ . Nella legenda sono riportati i numeri totali di vuoti ( $N_{1,\text{tot}}$ ,  $N_{2,\text{tot}}$ ), il numero di match ( $N_{\text{match}}$ ), il coefficiente di correlazione e la frazione di punti entro  $\pm 0.5$  dal valore ideale. Le linee orizzontali indicano i valori  $\mathcal{R} = 1$  (rossa continua) e  $\mathcal{R} = 0.5, 1.5$  (nere tratteggiate).

Figura 5.13: Mappatura dei volumi tra le due esecuzioni indipendenti di BitVF in spazio reale.

La Fig. 5.13 mostra i risultati del test. La distribuzione dei rapporti è centrata su  $\langle \mathcal{R} \rangle = (0.866 \pm 0.004)$  con deviazione standard  $\sigma_{\mathcal{R}} = (0.355 \pm 0.003)$ , indicando che il volume medio è riprodotto in maniera robusta tra le due esecuzioni. Lo *scatter plot* mostra che la maggior parte dei punti giace entro il 50% dal valore ideale, con una frazione entro  $[0.5, 1.5]$  pari al 87.1% e coefficiente di correlazione  $r = (-0.132 \pm 0.012)$ , che indica coerentemente assenza di dipendenze sistematiche. Questi risultati sono un buon punto di partenza per provare la coerenza della mappatura dei volumi come metodo di confronto fra cataloghi di vuoti in uno stesso spazio.

### 5.3.2 BitVF: mappatura dei volumi fra spazio reale e dei redshift

Applicando ora l'analisi al confronto tra spazio reale e dei redshift per BitVF, si ottengono i risultati mostrati in Fig. 5.14.



(a) Istogramma logaritmico dei rapporti di volume  $\mathcal{R} = V_{\text{real}}/\sum V_{\text{redshift}}$  per BitVF. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di match ( $N_{\text{match}}$ ), il valore medio  $\langle \mathcal{R} \rangle$ , la deviazione standard  $\sigma_{\mathcal{R}}$  e la frazione di punti entro  $[0.5, 1.5]$ . Le linee verticali indicano il valore  $\mathcal{R} = 1$  (rossa continua) e il valore medio (nera tratteggiata) della distribuzione.

(b) *Scatter plot* del rapporto  $\mathcal{R}$  in funzione del raggio di riferimento  $R_{\text{ref}}$  per BitVF. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di match ( $N_{\text{match}}$ ), il coefficiente di correlazione e la frazione di punti entro  $\pm 0.5$  dal valore ideale. Le linee orizzontali indicano i valori  $\mathcal{R} = 1$  (rossa continua) e  $\mathcal{R} = 0.5, 1.5$  (nere tratteggiate).

Figura 5.14: Mappatura dei volumi per BitVF tra spazio reale e dei redshift.

La Fig. 5.14 mostra che per **BitVF** il rapporto  $\mathcal{R}$  è centrato su  $\langle \mathcal{R} \rangle = (0.884 \pm 0.011)$  con deviazione standard  $\sigma_{\mathcal{R}} = (0.866 \pm 0.008)$ . Il valore medio inferiore a 1 indica che, in media, la somma dei volumi dei vuoti in spazio dei redshift sovrastima il volume reale di circa il 12%, suggerendo che gli effetti RSD alterano leggermente la distribuzione dei volumi portando a una sovrastima sistematica. La dispersione, sebbene non trascurabile ( $\sigma_{\mathcal{R}} \simeq 0.87$ ), è minore di quella osservata per **VIDE**, come si vedrà nel seguito. La frazione di punti entro  $[0.5, 1.5]$  è dell'83.5%, indicando che nella maggioranza dei casi il volume reale è riprodotto entro il 50% dalla somma dei volumi in spazio dei redshift. Il coefficiente di correlazione  $r = (-0.074 \pm 0.013)$  mostra una debole dipendenza negativa dal raggio, che indica una tendenza, seppur modesta, a sottostimare il rapporto per raggi più grandi. L'entità di questa correlazione è comunque ridotta, suggerendo che la distorsione indotta dagli RSD agisca in modo sostanzialmente uniforme su tutte le scale.

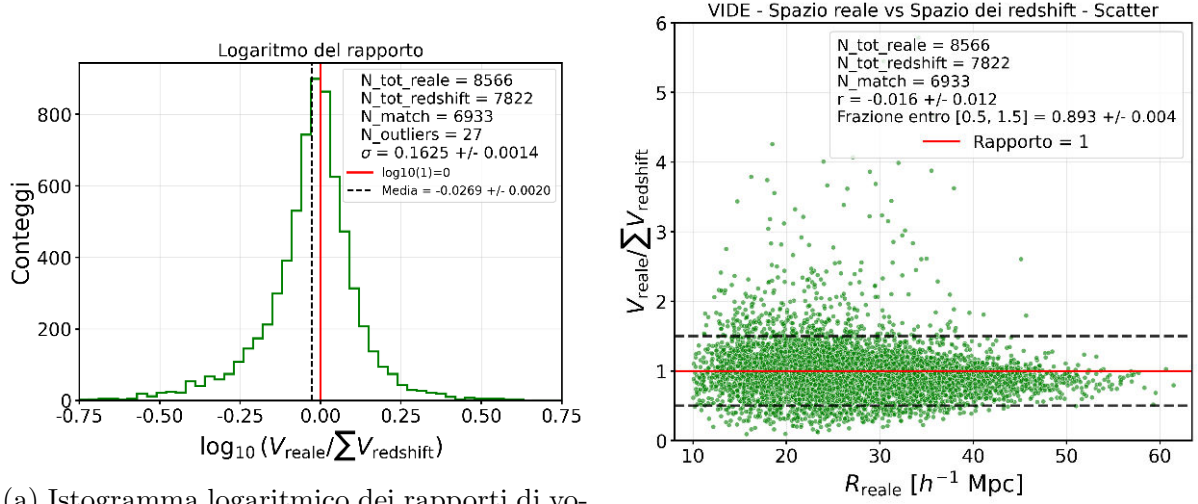
### 5.3.3 VIDE: mappatura dei volumi fra spazio reale e dei redshift

Con analoga analisi per **VIDE**, si ottengono i risultati mostrati in Fig. 5.15.

La Fig. 5.15 mostra che per **VIDE** il rapporto  $\mathcal{R}$  è centrato su  $\langle \mathcal{R} \rangle = (1.027 \pm 0.012)$  con deviazione standard  $\sigma_{\mathcal{R}} = (1.002 \pm 0.009)$ . Il valore medio è compatibile con 1 entro gli errori, indicando che, in media, il volume reale è ben riprodotto dalla somma dei volumi in spazio dei redshift. La dispersione è maggiore rispetto a **BitVF** ( $\sigma_{\mathcal{R}} \simeq 1.00$  contro  $\simeq 0.87$ ), suggerendo una corrispondenza caso per caso più dispersa. Tuttavia, la frazione di punti entro  $[0.5, 1.5]$  è del 89.3%, superiore a quella di **BitVF**, indicando che, sebbene più disperso, **VIDE** riesce a mantenere una frazione maggiore di vuoti entro questo intorno del valore ideale. Il coefficiente di correlazione  $r = (-0.016 \pm 0.012)$  è quasi compatibile con zero, indicando assenza di dipendenze sistematiche dal raggio.

La Tabella 5.5 riassume le statistiche principali della mappatura dei volumi per i tre casi analizzati.

I risultati mostrano un comportamento differenziato tra i due algoritmi. **BitVF** presenta una minore dispersione ( $\sigma_{\mathcal{R}} \simeq 0.87$  vs  $\simeq 1.00$ ), indicando una maggiore coerenza caso per caso nella mappatura dei volumi tra i due spazi, a fronte di una leggera sovrastima sistematica del volume reale da parte dei vuoti in spazio dei redshift ( $\langle \mathcal{R} \rangle \simeq 0.88$ ). **VIDE**, invece, mantiene un valore medio più vicino a 1 e una frazione di punti entro l'intervallo  $[0.5, 1.5]$  maggiore, ma con una dispersione più elevata. In conclusione, **BitVF**



(a) Istogramma logaritmico dei rapporti di volume  $\mathcal{R} = V_{\text{reale}}/\sum V_{\text{redshift}}$  per VIDE. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di match ( $N_{\text{match}}$ ), il valore medio  $\langle \mathcal{R} \rangle$ , la deviazione standard  $\sigma_{\mathcal{R}}$  e la frazione di punti entro  $[0.5, 1.5]$ . Le linee verticali indicano il valore  $\mathcal{R} = 1$  (rossa continua) e il valore medio (nera tratteggiata) della distribuzione.

(b) *Scatter plot* del rapporto  $\mathcal{R}$  in funzione del raggio di riferimento  $R_{\text{ref}}$  per VIDE. Nella legenda sono riportati i numeri totali di vuoti ( $N_{\text{tot, reale}}$ ,  $N_{\text{tot, redshift}}$ ), il numero di match ( $N_{\text{match}}$ ), il coefficiente di correlazione e la frazione di punti entro  $\pm 0.5$  dal valore ideale. Le linee orizzontali indicano i valori  $\mathcal{R} = 1$  (rossa continua) e  $\mathcal{R} = 0.5, 1.5$  (nere tratteggiate).

Figura 5.15: Mappatura dei volumi per VIDE tra spazio reale e dei redshift.

Tabella 5.5: Statistiche della mappatura dei volumi per i tre casi analizzati. Per ciascun caso sono riportati i numeri totali di vuoti nei due cataloghi ( $N_{1, \text{tot}}$ ,  $N_{2, \text{tot}}$ ), il numero di match ( $N_{\text{match}}$ ), il valore medio del rapporto  $\langle \mathcal{R} \rangle$ , la sua deviazione standard  $\sigma_{\mathcal{R}}$  e la frazione di punti con rapporto nell'intervallo  $[0.5, 1.5]$ .

| Caso                               | $N_{1, \text{tot}}$ | $N_{2, \text{tot}}$ | $N_{\text{match}}$ | $\langle \mathcal{R} \rangle$ | $\sigma_{\mathcal{R}}$ | Frazione entro $[0.5, 1.5]$ |
|------------------------------------|---------------------|---------------------|--------------------|-------------------------------|------------------------|-----------------------------|
| Esecuzione1 vs Esecuzione2 (Reale) | 6799                | 6785                | 6482               | $0.866 \pm 0.004$             | $0.356 \pm 0.003$      | $(87.1 \pm 0.4)\%$          |
| BitVF Reale vs Redshift            | 6799                | 6464                | 6036               | $0.884 \pm 0.011$             | $0.866 \pm 0.008$      | $(83.5 \pm 0.5)\%$          |
| VIDE Reale vs Redshift             | 8566                | 7822                | 6933               | $1.027 \pm 0.012$             | $1.002 \pm 0.009$      | $(89.3 \pm 0.4)\%$          |

risulta più coerente nella corrispondenza caso per caso, mentre **VIDE** garantisce una migliore riproduzione del volume medio e una maggiore frazione di vuoti ben mappati.

Si noti infine che, in questo contesto, il termine *match* assume un significato diverso rispetto alle analisi precedenti: un vuoto viene considerato identificato quando esiste almeno un centro nel catalogo di target che cade all'interno del suo volume sferico, portando al calcolo di un rapporto  $\mathcal{R}$ . Se invece non esiste alcun centro, il vuoto viene considerato non identificato e viene escluso dall'analisi. Il numero di questi vuoti, sebbene significativamente inferiore a quello dei *match*, non è trascurabile e, come si evince dalla Tabella 5.5, rappresenta una frazione non irrilevante del campione totale. La natura di questi vuoti isolati, che compaiono anche nelle Analisi 1 e 2, e le ragioni per cui non vengono identificati da alcun vuoto nell'altro catalogo meriterebbe un'indagine più approfondita, come si discuterà nel Capitolo successivo.

### 5.3.4 Confronto tra le metriche e prospettive per un'analisi combinata

Le tre metriche di identificazione presentate in questo capitolo forniscono informazioni complementari sulla corrispondenza tra cataloghi di vuoti in spazi diversi. L'Analisi 1 quantifica lo spostamento dei centri, l'Analisi 2 valuta la persistenza dei raggi tra vuoti spazialmente vicini, mentre l'Analisi 3 misura la conservazione dei volumi in modo aggregato. Tuttavia, ciascuna metrica identifica un diverso insieme di *match* di vuoti: non è detto che un vuoto identificato con il criterio del primo vicino lo sia anche con quello della somiglianza di raggio con filtro spaziale o della mappatura dei volumi.

Questo solleva una serie di domande aperte che meriterebbero un'indagine approfondita. In primo luogo, sarebbe interessante quantificare quante coppie di vuoti persistono indipendentemente dal metodo utilizzato, ovvero quante vengono identificate da tutti e tre i criteri. Queste coppie rappresenterebbero il nucleo "robusto" di vuoti che si corrispondono in modo inequivocabile tra i due spazi, e potrebbero costituire un campione privilegiato per analisi cosmologiche, essendo meno affette dalle distorsioni indotte dalle RSD.

In secondo luogo, i vuoti che vengono identificati da un criterio ma non da un altro presentano probabilmente caratteristiche distintive. Ad esempio, un vuoto potrebbe essere identificato dal primo vicino, perché i centri dei due vuoti si discostano di poco,

ma non dalla somiglianza di raggio, perché il raggio cambia significativamente, oppure viceversa. Studiare le proprietà di questi sottoinsiemi — in termini di distribuzione spaziale, raggi, profili di densità e posizione all'interno della *box* simulata — potrebbe rivelare informazioni sulla natura delle distorsioni indotte dalle RSD e su come i diversi algoritmi rispondono a tali effetti. In particolare, sarebbe utile verificare se i vuoti non matchati tendono a trovarsi ai bordi della simulazione, dove gli effetti di bordo potrebbero giocare un ruolo importante, oppure se presentano sistematicamente raggi estremi (molto grandi o molto piccoli).

In terzo luogo, una naturale estensione di questo lavoro consiste nel combinare le tre metriche in un unico criterio di identificazione. Ad esempio, si potrebbe definire una metrica composita che bilanci distanza spaziale, somiglianza in raggio e conservazione del volume, oppure adottare un approccio a votazione: una coppia di vuoti viene considerata identificata solo se soddisfa almeno due dei tre criteri, o addirittura tutti e tre. Questo permetterebbe di selezionare un campione di vuoti "altamente affidabili" per le analisi cosmologiche, riducendo le sistematiche introdotte dalla scelta di una particolare metrica di identificazione.

Infine, una domanda cruciale è se sia possibile identificare tipi di vuoti da escludere per minimizzare gli effetti RSD sul catalogo risultante. Ad esempio, vuoti molto grandi potrebbero essere più suscettibili alla frammentazione nello spazio dei redshift, mentre vuoti molto piccoli potrebbero essere più influenzati dallo *shot noise*. Se si riuscisse a individuare delle soglie (in raggio, densità centrale, o altre proprietà) oltre le quali i vuoti sono sistematicamente non matchati o mal mappati, si potrebbe definire una procedura di *cleaning* che selezioni solo i vuoti "robusti" alle RSD, migliorando l'affidabilità delle analisi cosmologiche basate sui vuoti.

L'analisi di queste questioni, sebbene non affrontata in questa tesi, rappresenta una direzione naturale per miglioramenti futuri di questo lavoro.

# Capitolo 6

## Conclusioni e prospettive future

In questo capitolo vengono riassunti i principali risultati ottenuti nel corso di questa tesi, con particolare attenzione alla validazione dell'algoritmo **BitVF** e al confronto delle sue prestazioni con quelle dell'algoritmo **VIDE**. Vengono inoltre discusse le implicazioni dei risultati per l'utilizzo dei vuoti cosmici come sonde cosmologiche, le principali limitazioni del presente lavoro e infine alcune possibili direzioni per sviluppi futuri.

### 6.1 Riassunto dei risultati principali

Questa tesi si è posta l'obiettivo di testare e validare l'algoritmo **BitVF**, un identificatore di vuoti cosmici basato su criteri dinamici, sviluppato da Sartori [2022], Sartori et al. [2026]. L'analisi è stata condotta su cataloghi di vuoti identificati a partire da simulazioni N-corpi di aloni di materia oscura, confrontando i cataloghi prodotti da **BitVF** con quelli ottenuti da **VIDE**, un algoritmo geometrico consolidato in letteratura.

L'analisi si è articolata in due fasi principali. Nel Capitolo 4 sono state presentate le proprietà statistiche dei cataloghi di vuoti, con un'analisi preliminare di una esecuzione isolata di **BitVF** e un confronto sistematico delle VSF, dei monopoli e dei quadrupoli tra **BitVF** e **VIDE**. I risultati (Figg. 4.3, 4.5 e 4.4) hanno mostrato che **BitVF** produce cataloghi di vuoti con profili di densità più profondi rispetto a **VIDE** ( $\xi_0 \simeq -0.978$  per **BitVF** contro  $\xi_0 \simeq -0.843$  per **VIDE**) e una minore frazione di vuoti a parità di raggio, con la configurazione utilizzata. **BitVF** si è rivelato significativamente più stabile nel passaggio da spazio reale a quello dei redshift: per quanto riguarda la VSF, la frazione

di punti (cioè i bin in raggio) per cui il residuo  $(VSF_Z - VSF_R)/\sigma$  cade entro  $\pm 2\sigma$  è del  $(92 \pm 7)\%$  per **BitVF**, contro il  $(57 \pm 13)\%$  di **VIDE**. Analogamente, per i monopoli della funzione di cross-correlazione vuoto-tracciante, la deviazione standard dei residui  $(MP_Z - MP_R)/\sigma$  è  $\sigma_\delta = 0.049$  per **BitVF**, circa 2.5 volte più piccola di quella di **VIDE** ( $\sigma_\delta = 0.121$ ). Questi risultati suggeriscono che **BitVF** risente meno degli effetti indotti dalle RSD nell'identificazione di vuoti rispetto a **VIDE**, poiché le sue quantità statistiche variano meno nel passaggio da uno spazio all'altro.

L'analisi del parametro `merging_threshold` per i cataloghi **VIDE** (Fig. 4.6 e Tabella 4.1) ha rivelato che una soglia 0 produce vuoti su scale molto grandi, fino a  $240 h^{-1}\text{Mpc}$ , ma con profili di densità meno profondi ( $\xi_0 \simeq -0.45$ ), mentre una soglia  $10^{-9}$  restituisce vuoti più piccoli e meglio definiti ( $\xi_0 \simeq -0.85$ ) ma più sensibili agli effetti RSD. Per il confronto con **BitVF** è stata scelta la configurazione con soglia  $10^{-9}$ , in quanto identifica vuoti di dimensioni confrontabili.

Per ultimo, si è presentato il risultato della procedura di *cleaning* per una esecuzione di **BitVF** (Fig. 4.9) e si è visto che la procedura di accoppiamento vuoto a vuoto, presentata nel Capitolo 5, rimane robusta anche dopo l'applicazione dell'algoritmo di *cleaning*. La frazione di accoppiamenti tra le due esecuzioni indipendenti, nello specifico quella stimata attraverso la metrica del primo vicino, si mantiene elevata ( $> 85\%$ ) sia prima che dopo la procedura, indicando che il *cleaning* dei cataloghi non introduce distorsioni sistematiche nella corrispondenza tra le due. Questo risultato suggerisce che le proprietà statistiche dei vuoti e la loro riproducibilità siano sostanzialmente indipendenti dai filtri applicati a posteriori, rafforzando la robustezza delle conclusioni ottenute.

Nel Capitolo 5 sono state condotte tre analisi di validazione basate su metriche di identificazione diverse: il criterio del primo vicino (Analisi 1), la somiglianza di raggio con filtro spaziale (Analisi 2) e la mappatura dei volumi (Analisi 3). Le tre analisi hanno fornito una visione complementare del comportamento dei due algoritmi. **BitVF** si è dimostrato complessivamente più stabile nell'identificazione dei vuoti tra spazio reale e dei redshift, come evidenziato dalla maggiore frazione di match e dalle minori distanze normalizzate, mentre **VIDE** ha mostrato una migliore riproduzione del volume medio nella mappatura dei volumi. Si sono inoltre ottenuti buoni risultati a riguardo della riproducibilità di esecuzioni indipendenti di **BitVF** (Tabelle 5.1 e 5.3).

Il risultato più rilevante emerso dall'intera analisi di validazione è la straordinaria precisione con cui **BitVF** identifica i centri dei vuoti nello spazio dei redshift rispetto

a VIDE. Come mostrato nella Fig. 5.4, la distribuzione delle distanze normalizzate per BitVF è estremamente piccata a zero, con una mediana  $\text{Med}[d/(R_{\text{reale}} + R_{\text{redshift}})] = (0.0078 \pm 0.0006)$ , mentre per VIDE la distribuzione è notevolmente più ampia e spostata verso valori più elevati, con mediana  $\text{Med}[d/(R_{\text{reale}} + R_{\text{redshift}})] = (0.0557 \pm 0.0005)$ . La differenza tra le mediane dei due algoritmi è di un fattore  $\sim 7$ , a testimonianza della capacità di BitVF di tracciare i centri dei vuoti tra i due spazi con una precisione che VIDE non riesce a eguagliare. Questo risultato costituisce la prova più forte della minore sensibilità di BitVF agli effetti RSD rispetto a un algoritmo geometrico consolidato come VIDE, e rappresenta il risultato più interessante e significativo emerso da questa tesi.

Nell'Analisi 1, l'identificazione per primo vicino ha mostrato che BitVF presenta una frazione di match più elevata (56.4% per  $R_v$ , 64.9% per  $R_{\text{ext}}$ ) rispetto a VIDE (44.3%), e distanze normalizzate significativamente più piccole ( $\langle d/(R_1 + R_2) \rangle \sim 0.02 - 0.025$  per BitVF contro  $\sim 0.056$  per VIDE), indicando che BitVF è in grado di identificare un campione più ampio di vuoti mantenendo una maggiore precisione spaziale. La pendenza del fit lineare per BitVF,  $m = (0.993 \pm 0.005)$ , è vicina alla compatibilità con 1, mentre per VIDE  $m = (1.025 \pm 0.003)$  indica una leggera sovrastima sistematica dei raggi in spazio dei redshift con questa metrica. Si noti tuttavia che la distribuzione delle differenze relative dei raggi per VIDE risulta più stretta:  $\sigma_{\Delta R/R} \simeq 0.05$  contro  $\simeq 0.11$  per BitVF; riflettendo una maggiore selettività dell'identificazione. In altre parole, l'analisi mostra che i vuoti di VIDE conservati nello spazio dei redshift mantengono raggi molto simili ( $\sigma \simeq 0.05$ ), indicando che la corrispondenza è più selettiva ma con meno match e minore precisione centro-centro. BitVF, invece, è meno sensibile alle RSD e più stabile nell'identificazione dei centri, consentendo più identificazioni e distanze normalizzate minori, a costo di una maggiore dispersione nei raggi ( $\sigma \simeq 0.11$ ), conseguenza fisica delle distorsioni indotte dal campo di densità.

Nell'Analisi 2, l'identificazione per somiglianza di raggio con filtro spaziale ha confermato i risultati precedenti: BitVF mantiene una frazione di match dell'86.4% contro l'81.7% di VIDE, e distanze normalizzate  $\text{Med}[d/(R_{\text{real}} + R_{\text{redshift}})] = 0.0226$  contro 0.0921 (Tabella 5.4). Seppur in maniera meno evidente, anche in questo caso la maggiore selettività di VIDE si riflette in una distribuzione delle differenze relative leggermente più stretta:  $\sigma_{\Delta R/R} = 0.1080$  per VIDE contro 0.0989 per BitVF.

La riproducibilità di BitVF tra due esecuzioni indipendenti (Tabelle 5.1 e 5.3) è risultata eccellente per entrambi i criteri, con frazioni di identificazione superiori all'83%

e valori medi delle distanze normalizzate  $\langle d/(R_1 + R_2) \rangle \sim 0.008$ , indipendentemente dalla metrica utilizzata.

Nell'Analisi 3, la mappatura dei volumi ha rivelato un diverso comportamento dei due algoritmi. È importante notare che il rapporto  $\mathcal{R} = V_{\text{reale}} / \sum V_{\text{redshift}}$  è atteso essere sistematicamente inferiore a 1, sia per la costruzione stessa della metrica (la somma dei volumi dei vuoti in spazio dei redshift che cadono all'interno del raggio del vuoto reale tende a sovrastimare il denominatore per fluttuazioni statistiche), sia per un effetto fisico: i vuoti in spazio dei redshift presentano mediamente un volume maggiore rispetto a quelli in spazio reale, a causa dell'elongazione lungo la linea di vista indotta dalle RSD. **BitVF** presenta una minore dispersione ( $\sigma_{\mathcal{R}} \simeq 0.87$  vs  $\simeq 1.00$  per **VIDE**), indicando una maggiore coerenza caso per caso nella mappatura dei volumi tra i due spazi, nonostante un valore medio  $\langle \mathcal{R} \rangle \simeq 0.88$  che riflette il comportamento atteso. **VIDE**, invece, mantiene un valore medio più vicino a 1 ( $\langle \mathcal{R} \rangle \simeq 1.027$ ) e una frazione di punti entro  $[0.5, 1.5]$  maggiore (89.3% contro 83.5%), ma con una dispersione più elevata. Il test di coerenza fra esecuzioni indipendenti nello spazio reale (Tabella 5.5) ha confermato la validità del metodo, con  $\langle \mathcal{R} \rangle = (0.866 \pm 0.004)$  e frazione entro  $[0.5, 1.5]$  del 87.1%.

## 6.2 Implicazioni per l'uso dei vuoti cosmici come sonde cosmologiche

I risultati esposti vogliono fornire supporto all'utilizzo dei vuoti cosmici come sonde cosmologiche, in linea con quanto discusso ad esempio da Sheth and van de Weygaert [2004], Jennings et al. [2013], Hamaus et al. [2014]. La maggiore stabilità di **BitVF** nel passaggio da spazio reale a spazio dei redshift, evidenziata sia dalle VSF che dai profili di densità, rende questo algoritmo particolarmente adatto per minimizzare possibili sistematiche nella modellizzazione delle RSD. L'elevata riproducibilità di **BitVF** tra esecuzioni indipendenti, con frazioni di match superiori all'83% e coefficienti di correlazione  $r \simeq 1$ , suggerisce l'affidabilità dei risultati e una loro non dipendenza dalla particolare realizzazione del catalogo iniziale.

L'analisi di identificazione ha mostrato che **BitVF** identifica i medesimi centri nei due spazi, reale e dei redshift, con spostamenti trascurabili, come evidenziato dalle distribuzioni delle distanze normalizzate estremamente piccate a zero. Questa caratteristica

rende **BitVF** uno strumento ideale per misurare il quadrupolo della cross-correlazione vuoto-tracciante, che è direttamente sensibile alle distorsioni lungo la linea di vista e può essere utilizzato per vincolare parametri cosmologici.

Inoltre, la minore dispersione nella mappatura dei volumi per **BitVF** ( $\sigma_{\mathcal{R}} \simeq 0.87$  contro  $\simeq 1.00$  per **VIDE**) indica che **BitVF** preserva una corrispondenza più coerente caso per caso dei volumi tra gli spazi. Questo è un vantaggio significativo per applicazioni che richiedono una mappatura precisa dei volumi, come lo studio della funzione di correlazione vuoto-materia o la calibrazione di modelli teorici di VSF [Contarini et al., 2019, 2022].

### 6.3 Limitazioni del presente lavoro

Nonostante i risultati generalmente positivi, questo lavoro presenta alcune importanti limitazioni di cui è opportuno discutere. In primo luogo, tutte le analisi sono state condotte su un singolo catalogo di traccianti a un redshift fissato  $z = 1$ , e in configurazione a geometria cubica. Sebbene la scelta della simulazione sia motivata dalla sua risoluzione e dal volume sufficiente a garantire un campione statisticamente significativo di vuoti, l'assenza di una geometria a cono luce limita la diretta confrontabilità con cataloghi osservativi reali, che presentano effetti di profondità variabile lungo la linea di vista. Sarebbe pertanto necessario estendere l'analisi a diversi redshift e a diverse tipologie di traccianti, nonché a geometrie più realistiche, per verificare la generalità dei risultati e la robustezza di **BitVF** in condizioni osservative più vicine alla realtà.

In secondo luogo, l'assunzione di sfericità dei vuoti, adottata in tutte le analisi dei capitoli 4 e 5, costituisce un'approssimazione della loro morfologia reale. Essi, infatti, presentano forme complesse e irregolari e l'uso di una geometria sferica potrebbe influenzare i risultati delle analisi, in particolare per i vuoti più piccoli o in regioni ad alta densità. Un'estensione naturale di questo lavoro consisterebbe nell'utilizzo di tecniche che tengano conto della forma reale dei bacini di sottodensità.

In terzo luogo, la procedura di *cleaning* è stata applicata ai cataloghi di **BitVF**, mostrando di non introdurre distorsioni sistematiche nella corrispondenza tra le due esecuzioni; al contrario, i risultati suggeriscono un leggero miglioramento della coerenza tra i cataloghi. Tuttavia, i cataloghi puliti non sono stati utilizzati per le analisi statistiche presentate in questo lavoro, né per la validazione dei modelli teorici. Di conseguenza,

sebbene i risultati sulle identificazioni siano convincenti e rappresentino un buon supporto a questa tesi, non è possibile affermare con certezza che le proprietà statistiche dei cataloghi e la loro sensibilità agli effetti RSD rimangano invariate dopo il *cleaning*. Una verifica rigorosa richiederebbe la ripetizione sistematica delle analisi sui cataloghi puliti; questo, insieme all'estensione dell'analisi a geometrie a cono luce, rappresenta una naturale e necessaria direzione per sviluppi futuri.

In quarto luogo, in tutte e tre le metriche di identificazione presentate nel Capitolo 5, una frazione non trascurabile di vuoti rimane senza corrispondenza nell'altro catalogo. La natura di questi vuoti non identificati non è stata analizzata in dettaglio: essi potrebbero essere vuoti che non trovano un candidato sufficientemente vicino o sufficientemente simile in raggio (vuoti isolati), vuoti che vengono persi perché il loro miglior candidato è già stato assegnato a un altro vuoto, oppure vuoti che subiscono trasformazioni significative nel passaggio da uno spazio all'altro, come la frammentazione di un vuoto grande in più vuoti più piccoli nello spazio dei redshift, o viceversa la fusione di più vuoti in uno solo. L'analisi dettagliata di queste popolazioni di vuoti non identificati, in termini di distribuzione spaziale, raggi e proprietà di densità, potrebbe fornire ulteriori informazioni sulla natura delle distorsioni indotte dalle RSD e sulla diversa risposta dei due algoritmi a tali effetti. Questa limitazione, sebbene non infici la validità delle conclusioni principali, rappresenta un aspetto che meriterebbe un'indagine più approfondita in lavori futuri.

Infine, l'analisi si è concentrata sulle proprietà statistiche dei cataloghi di vuoti (VSF, profili di densità, metriche di identificazione), ma non ha esplorato l'uso dei vuoti per vincolare parametri cosmologici. Questo rappresenta un passo successivo naturale per lavori più strutturati.

## 6.4 Prospettive future e considerazioni conclusive

I risultati ottenuti in questa tesi aprono diverse direzioni per possibili sviluppi. Per verificare la robustezza di **BitVF** in diversi regimi cosmologici e osservativi, si potrebbe estendere l'analisi a cataloghi di traccianti diversi, come galassie o aloni di diverse masse, e a diversi redshift. In particolare, l'uso di geometrie a cono luce, più vicine alle reali condizioni osservative, consentirebbe di testare l'algoritmo in scenari più realistici e di

valutarne le prestazioni in presenza di effetti di selezione e profondità variabile lungo la linea di vista.

In secondo luogo, come già accennato, l'utilizzo della forma reale dei vuoti, piuttosto che dell'approssimazione sferica, migliorerebbe l'accuratezza delle analisi di identificazione e mappatura dei volumi. Questo approccio, che richiederebbe l'implementazione di tecniche basate sulla morfologia effettiva dei bacini di sottodensità, potrebbe rivelare differenze più sottili nel comportamento dei diversi algoritmi di identificazione e fornire una descrizione più accurata delle proprietà dei vuoti.

In terzo luogo, un'analisi dedicata dei vuoti non identificati in ciascuna delle metriche presentate potrebbe fornire informazioni preziose sulla natura delle distorsioni indotte dalle RSD. In particolare, sarebbe interessante studiare se i vuoti non identificati abbiano caratteristiche sistematicamente diverse da quelli identificati: ad esempio, se siano tendenzialmente più grandi o più piccoli, se si trovino in regioni a diversa densità di traccianti, o se mostrino profili di densità anomali. Questo tipo di analisi potrebbe aiutare a comprendere quali vuoti sono più suscettibili alle distorsioni indotte dalle RSD e come i diversi algoritmi rispondono a tali effetti, fornendo indicazioni utili per la scelta dell'algoritmo più appropriato a seconda dell'applicazione cosmologica desiderata.

Parallelamente, i cataloghi di vuoti sottoposti a *cleaning* potrebbero essere utilizzati per calibrare modelli teorici come, ad esempio, il Vdn per la VSF, consentendo di verificare se la procedura di *cleaning* migliori l'accordo con le predizioni teoriche. Analogamente, l'uso dei profili di densità dei vuoti per testare modelli di gravità modificata rappresenta una direzione promettente. Come già accennato, un'estensione di questo lavoro consiste nell'utilizzo dei cataloghi di vuoti prodotti da BitVF per vincolare parametri cosmologici, ad esempio  $\Omega_m$  e  $\sigma_8$ , attraverso la VSF e i profili di densità. La minore sensibilità agli effetti RSD di BitVF, emersa dalle analisi presentate, potrebbe tradursi in vincoli più stringenti, riducendo le incertezze associate alla modellizzazione di questi effetti e migliorando la precisione delle stime di questi parametri.

Infine, come discusso nella Sezione 5.3.4 del Capitolo 5, un'estensione naturale di questo lavoro consiste nel combinare le tre metriche di identificazione in un unico criterio. Si potrebbe definire una metrica composita che bilanci distanza spaziale e somiglianza in raggio, oppure adottare un approccio a votazione in cui una coppia di vuoti viene considerata identificata solo se soddisfa almeno due dei tre criteri. Questo permetterebbe di selezionare un campione di vuoti "altamente affidabili" per le analisi cosmologiche,

riducendo le sistematiche introdotte dalla scelta di una particolare metrica. Inoltre, l'analisi di quei vuoti che vengono identificati da tutte e tre le metriche, contrapposti a quelli identificati da una sola, potrebbe rivelare quali proprietà dei vuoti li rendono più o meno suscettibili alle distorsioni indotte dalle RSD, fornendo indicazioni utili per definire criteri di selezione ottimali.

In conclusione, i risultati ottenuti sono in linea con l'ipotesi, avanzata da Sartori [2022], Sartori et al. [2026], secondo cui la metodologia dinamica di **BitVF**, basata sulla ricostruzione del campo di spostamento attraverso l'approssimazione di Zel'dovich, possa offrire vantaggi rispetto agli algoritmi puramente geometrici nel tracciare gli effetti RSD. In particolare, emergono evidenze di una notevole mitigazione delle sistematiche relative all'identificazione dei vuoti in spazio dei redshift da parte di **BitVF**, come suggerito dalla maggiore stabilità delle proprietà statistiche e dalla maggiore coerenza nelle metriche di identificazione e mappatura dei volumi. Il risultato più eclatante è la precisione con cui **BitVF** identifica i centri dei vuoti nello spazio dei redshift, con una mediana delle distanze normalizzate di un fattore  $\sim 7$  inferiore a quella di **VIDE**, a testimonianza della sua straordinaria capacità di tracciare le posizioni dei vuoti nonostante le distorsioni indotte dalle RSD. Tuttavia, rimangono aperte questioni relative all'impatto di diverse configurazioni di **BitVF** sulle analisi cosmologiche, dove i guadagni potrebbero essere più o meno significativi a seconda del contesto applicativo. Future indagini saranno necessarie per valutare appieno il potenziale di **BitVF** in questo senso.

**BitVF** si configura dunque come uno strumento promettente e valido per l'analisi dei vuoti cosmici in situazioni in cui la capacità di tracciare gli effetti RSD e di fornire cataloghi di vuoti riproducibili sia di rilevante importanza. L'estensione dell'analisi a geometrie a cono luce, l'utilizzo dei cataloghi puliti per la calibrazione di modelli teorici e l'applicazione a cataloghi osservativi reali, come quelli di *Euclid* [Euclid Collaboration et al., 2022], rappresentano le direzioni naturali per sviluppi futuri.

# Bibliografia

- Ravi K. Sheth and Rien van de Weygaert. A hierarchy of voids: much ado about nothing. *Monthly Notices of the Royal Astronomical Society*, 350:517–538, 2004. doi: 10.1111/j.1365-2966.2004.07661.x.
- Nico Hamaus, Benjamin D. Wandelt, P. M. Sutter, Guilhem Lavaux, and Michael S. Warren. Cosmology from void-galaxy correlations. *Physical Review Letters*, 112:041304, 2014. doi: 10.1103/PhysRevLett.112.041304.
- Euclid Collaboration et al. Euclid. i. overview of the euclid mission. *Astronomy & Astrophysics*, 662:A112, 2022. doi: 10.1051/0004-6361/202141348.
- Y. B. Zel’dovich. Gravitational instability: An approximate theory for large density perturbations. *Astronomy and Astrophysics*, 5:84–89, 1970.
- Albert Einstein. Die feldgleichungen der gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, pages 778–786, 1915.
- Edwin Hubble. A spiral nebula as a stellar system, messier 31. *The Astrophysical Journal*, 69:103–158, 1929a. doi: 10.1086/143167.
- Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929b. doi: 10.1073/pnas.15.3.168.
- Planck Collaboration, Y. Akrami, et al. Planck 2018 results. vii. isotropy and statistics of the cmb. *Astronomy & Astrophysics*, 641:A7, 2020a. doi: 10.1051/0004-6361/201935201.

- Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, and et al. Planck 2018 results. vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, 2020b. doi: 10.1051/0004-6361/201833910.
- Eugenio Bianchi and Carlo Rovelli. Why all these prejudices against a constant? *arXiv e-prints*, 2010.
- S. Perlmutter, G. Aldering, G. Goldhaber, et al. Measurements of omega and lambda from 42 high-redshift supernovae. *The Astrophysical Journal*, 517:565, 1999. doi: 10.1086/307221.
- A. G. Riess, A. V. Filippenko, P. M. Challis, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116:1009, 1998. doi: 10.1086/300499.
- S. Taubenberger, S. H. Suyu, E. Komatsu, I. Jee, S. Birrer, V. Bonvin, F. Courbin, C. E. Rusu, A. J. Shajib, and K. C. Wong. The hubble constant determined through an inverse distance ladder including quasar time delays and type ia supernovae. *Astronomy & Astrophysics*, 628:L7, 2019. doi: 10.1051/0004-6361/201935980.
- Adam G. Riess, Stefano Casertano, Wenlong Yuan, Lucas M. Macri, and Dan Scolnic. Large magellanic cloud cepheid standards provide a 1% foundation for the determination of the hubble constant and stronger evidence for physics beyond  $\Lambda$ cdm. *The Astrophysical Journal*, 876(1):85, 2019. doi: 10.3847/1538-4357/ab1422.
- James E. Brau. Evolution of the energy density of the universe. University of Oregon, Department of Physics, 2011. URL [https://pages.uoregon.edu/jimbrau/BrauImNew/Chap27/7th/AT\\_7e\\_Figure\\_27\\_01.jpg](https://pages.uoregon.edu/jimbrau/BrauImNew/Chap27/7th/AT_7e_Figure_27_01.jpg). Accessed: 2026-04-20.
- Fritz Zwicky. Die rotverschiebung von extragalaktischen nebeln. *Helvetica Physica Acta*, 6:110–127, 1933. doi: 10.1007/s10714-008-0707-4.
- Vera C. Rubin, W. Kent Jr. Ford, and Norbert Thonnard. Rotational properties of 21 sc galaxies with a large range of luminosities and radii, from ngc 4605 ( $r=4\text{kpc}$ ) to ugc 2885 ( $r=122\text{kpc}$ ). *The Astrophysical Journal*, 238:471–487, 1980. doi: 10.1086/158003.

- Alexandre Refregier. Weak gravitational lensing by large-scale structure. *Annual Review of Astronomy & Astrophysics*, 41:645–668, 2003. doi: 10.1146/annurev.astro.41.111302.102207.
- Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, 2010. doi: 10.1088/0034-4885/73/8/086901.
- Eric V. Linder. Cosmic growth history and expansion history. *Physical Review D*, 72:043529, 2005. doi: 10.1103/PhysRevD.72.043529.
- Y. B. Zeldovich. A hypothesis, unifying the structure and the entropy of the universe. *Monthly Notices of the Royal Astronomical Society*, 160:1–3, 1972. doi: 10.1093/mnras/160.1.1P.
- B. L. Falck, M. C. Neyrinck, M. A. Aragon-Calvo, G. Lavaux, and A. S. Szalay. Example article title here. *The Astrophysical Journal*, 745:1–6, 2012. doi: 10.1088/0004-637X/745/1/17.
- N. Kaiser. On the spatial correlations of abell clusters. *The Astrophysical Journal Letters*, 284:L9–L12, 1984. doi: 10.1086/184341.
- T. Ishiyama, F. Prada, A. A. Klypin, G. Sato-Polito, A. V. Macciò, et al. The uchuu simulations: Data release 1 and dark matter halo concentrations. *Monthly Notices of the Royal Astronomical Society*, 506(3):4210–4228, 2021. doi: 10.1093/mnras/stab1755.
- F. Marulli, A. Veropalumbo, M. Moresco, et al. Cosmobolognalib: C++/python libraries for cosmological calculations. <https://gitlab.com/federicomarulli/CosmoBolognaLib>, 2016–2026. Accessed: 2026-04-20.
- J. M. Colberg et al. The aspen-amsterdam void finder comparison project. *Monthly Notices of the Royal Astronomical Society*, 387:933–944, 2008. doi: 10.1111/j.1365-2966.2008.13307.x.
- Guilhem Lavaux and Benjamin D. Wandelt. Precision cosmography with stacked voids. *Monthly Notices of the Royal Astronomical Society*, 403:403–412, 2009. doi: 10.1111/j.1365-2966.2009.16121.x.

- G. Kauffmann and A. P. Fairall. Voids in the distribution of galaxies: an assessment of their significance and derivation of a void probability function. *Monthly Notices of the Royal Astronomical Society*, 248:313–324, 1991. doi: 10.1093/mnras/248.2.313.
- Mark C. Neyrinck. Zobov: a parameter-free void-finding algorithm. *Monthly Notices of the Royal Astronomical Society*, 386:2101–2109, 2008. doi: 10.1111/j.1365-2966.2008.13180.x.
- P. M. Sutter et al. Vide: The void identification and examination toolkit. *Astronomy and Computing*, 9:1–9, 2014. doi: 10.1016/j.ascom.2014.10.002.
- Erwin Platen, Rien Van De Weygaert, and Bernard J. T. Jones. A cosmic watershed: the wvf void detection technique. *Monthly Notices of the Royal Astronomical Society*, 380(2):551–570, 2007. doi: 10.1111/j.1365-2966.2007.12125.x.
- J. E. Forero-Romero, Y. Hoffman, S. Gottlöber, A. Klypin, and G. Yepes. The dynamics of the cosmic web. *Monthly Notices of the Royal Astronomical Society*, 391:759–766, 2008. doi: 10.1111/j.1365-2966.2008.13910.x.
- A. Elyiv et al. Uncorrelating and lagrangian zel’dovich void finders. *Astronomy & Astrophysics*, 584:A82, 2015. doi: 10.1051/0004-6361/201526458.
- Kai Lehman, Luisa Lucie-Smith, Nico Hamaus, Christopher T. Davies, and Klaus Dolag. Cosmological inference with cosmic voids and neural network emulators. *arXiv e-prints*, 2025.
- J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical gaussian fluctuations. *The Astrophysical Journal*, 379:440–460, Oct 1991. doi: 10.1086/170520.
- Elise Jennings, Yin Li, and Wayne Hu. The abundance of voids and the excursion set formalism. *Monthly Notices of the Royal Astronomical Society*, 434:2167–2181, 2013. doi: 10.1093/mnras/stt1162.
- Sofia Contarini et al. Cosmic voids in modified gravity models with massive neutrinos. *Monthly Notices of the Royal Astronomical Society*, 488:3526–3542, 2019. doi: 10.1093/mnras/stz1857.

- Sofia Contarini et al. The bias of cosmic voids in the presence of massive neutrinos. *Monthly Notices of the Royal Astronomical Society*, 511:2585–2604, 2022. doi: 10.1093/mnras/stac184.
- Simone Sartori. Back-in-time void finder: dynamical detection of cosmic voids for precision cosmology. Master’s thesis, Università di Bologna, 2022.
- Simone Sartori, Sofia Contarini, Elena Sarpa, Giulia Degni, Federico Marulli, Stephanie Escoffier, and Lauro Moscardini. The back-in-time void finder: dynamical identification of cosmic voids through optimal transport reconstruction. *arXiv e-prints*, 2026.
- Rupert A. C. Croft and Enrique Gaztañaga. The dynamics of the cosmic web: A new way to reconstruct the linear density field. *Monthly Notices of the Royal Astronomical Society*, 285(4):793–812, Mar 1997. doi: 10.1093/mnras/285.4.793.
- Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l’Académie Royale des Sciences de Paris*, pages 666–704, 1781.
- Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on Pure and Applied Mathematics*, 44(4):375–417, 1991. doi: 10.1002/cpa.3160440402.
- Yann Brenier. Extended monge-kantorovich theory. 1813:91–121, 2003. doi: 10.1007/978-3-540-44857-0\_5.
- T. Ronconi and F. Marulli. Cosmological exploitation of cosmic void statistics - new numerical tools in the cosmobolognalib to extract cosmological constraints from the void size function. *Astronomy & Astrophysics*, 607:A24, 2017. doi: 10.1051/0004-6361/201730852.

# Appendice A

## Statistiche utilizzate nell'analisi

In questa appendice vengono riportate le principali definizioni statistiche utilizzate nell'analisi dei cataloghi di vuoti, con le relative formule e i metodi di propagazione degli errori.

### A.1 Residuo in sigma

Il residuo tra due misurazioni  $x_i$  e  $y_i$  con errori  $\sigma_{x_i}$  e  $\sigma_{y_i}$  è definito come:

$$\delta_i = \frac{y_i - x_i}{\sqrt{\sigma_{x_i}^2 + \sigma_{y_i}^2}}. \quad (\text{A.1})$$

L'errore sul residuo, ottenuto per propagazione, è:

$$\sigma_{\delta_i} = \sqrt{\left(\frac{\sigma_{y_i}}{\sqrt{\sigma_{x_i}^2 + \sigma_{y_i}^2}}\right)^2 + \left(\frac{\sigma_{x_i}}{\sqrt{\sigma_{x_i}^2 + \sigma_{y_i}^2}}\right)^2 + \left(\frac{\delta_i \sqrt{\sigma_{x_i}^2 + \sigma_{y_i}^2}}{\sigma_{x_i}^2 + \sigma_{y_i}^2}\right)^2}. \quad (\text{A.2})$$

### A.2 Frazione di match

Per il confronto tra due cataloghi di vuoti, la frazione di match è definita come:

$$f = \frac{2N_{\text{match}}}{N_1 + N_2}, \quad (\text{A.3})$$

con errore:

$$\sigma_f = \sqrt{\frac{2f(1-f)}{N_1 + N_2}}. \quad (\text{A.4})$$

### A.3 Correlazione pesata

Il coefficiente di correlazione di Pearson pesato è:

$$r = \frac{\sum_i w_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i w_i (x_i - \bar{x})^2 \sum_i w_i (y_i - \bar{y})^2}}, \quad (\text{A.5})$$

dove  $w_i = (\sigma_{x_i}^2 + \sigma_{y_i}^2)^{-1}$  e  $\bar{x}$ ,  $\bar{y}$  sono le medie pesate. L'errore su  $r$  è stimato tramite la formula di Fisher:

$$\sigma_r = \frac{1 - r^2}{\sqrt{N - 3}}. \quad (\text{A.6})$$

### A.4 Chi-quadro ridotto

Per i fit lineari presentati nell'analisi, il chi-quadro ridotto è:

$$\tilde{\chi}_\nu^2 = \frac{1}{N - 2} \sum_i \frac{(y_i - y_{\text{fit},i})^2}{\sigma_i^2}. \quad (\text{A.7})$$

### A.5 Errori sulle statistiche

| Statistica          | Errore                 | Formula                                    |
|---------------------|------------------------|--------------------------------------------|
| Media               | Errore standard        | $\sigma_{\bar{x}} = \sigma_x / \sqrt{N}$   |
| Deviazione standard | Errore sulla dev. std. | $\sigma_\sigma = \sigma / \sqrt{2(N - 1)}$ |
| Percentuale         | Errore binomiale       | $\sigma_p = \sqrt{p(1 - p)/N}$             |