

ANALISI DATI DEL PROFILO DI DENSITÀ DI MATERIA OSCURA

Guglielmo Barbanti

Marzo 2024



1 Obiettivi

Questa relazione vuole descrivere e analizzare i risultati ottenuti dall'analisi dati eseguita sul profilo di densità degli aloni di materia oscura.

2 Introduzione teorica

Lo studio e la caratterizzazione del profilo di densità degli aloni di materia oscura giocano un ruolo di primo rilievo per la stima della massa di un ammasso di galassie e nel comprendere le proprietà della materia oscura stessa.

Tale profilo di densità proiettato degli aloni è descritto dal *fit* analitico (NFW profile):

$$\Sigma(R, M, c, z) = 2r_s \rho_s f(x)$$

dove R, M, c, z sono parametri liberi. Nello specifico R è la distanza dal centro dell'alone di materia oscura (misurata in Mpc), M è la massa dell'alone di materia oscura, c è la sua concentrazione e infine la quantità z è il *redshift*.

La funzione r_s è il raggio caratteristico dell'ammasso, ed è definita come:

$$r_s = \frac{1}{c} \left(\frac{3M}{4\pi 200 \rho_m (1+z)^3} \right)^{1/3}$$

dove ρ_m è la densità media dell'Universo, pari a:

$$\rho_m \simeq 0.286 \cdot 0.7^2 \cdot 2.7751428946 \cdot 10^{11} [M_\odot / \text{Mpc}^3]$$

La funzione ρ_s è la densità caratteristica dell'ammasso:

$$\rho_s = \frac{\rho_m (1+z)^3 200 c^3}{3 f_c}$$

con

$$f_c = \ln(1+c) - \frac{c}{1+c}$$

Infine si ha che la $f(x)$, dove $x = \frac{R}{r_s}$, è definita a tratti:

$$f(x) = \begin{cases} \frac{1}{x^2-1} \left(1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}} \right) & \text{se } x < 1 \\ \frac{1}{3} & \text{se } x = 1 \\ \frac{1}{x^2-1} \left(1 - \frac{2}{\sqrt{x^2-1}} \operatorname{arctan} \sqrt{\frac{x-1}{1+x}} \right) & \text{se } x > 1 \end{cases} \quad (1)$$

3 Analisi dei dati

3.1 Prima richiesta

Dato R nell'intervallo $[0.03, 1.0] \text{ Mpc}$, e dati i seguenti valori di massa $M = [10^{13.5}, 10^{14.0}, 10^{15.0}] M_\odot$, di concentrazione $c = [2, 5, 10]$ e redshift $z = 0.0$ si è definita la funzione NFW $\Sigma(R, M, c, z)$ per calcolare i profili di densità nelle 9 combinazioni precedenti.

Nello specifico si sono prima definite le functions che vanno a formare la Σ , e poi si è partizionato l'intervallo di R in un numero arbitrario di valori (in questo caso 150).

Si ha così a disposizione 150 valori di R su cui calcolare, nei 9 casi possibili di combinazioni di massa e concentrazione, il profilo di densità.

Tramite *subplots* sono stati creati i 9 pannelli contenenti l'andamento, che di seguito riportiamo.

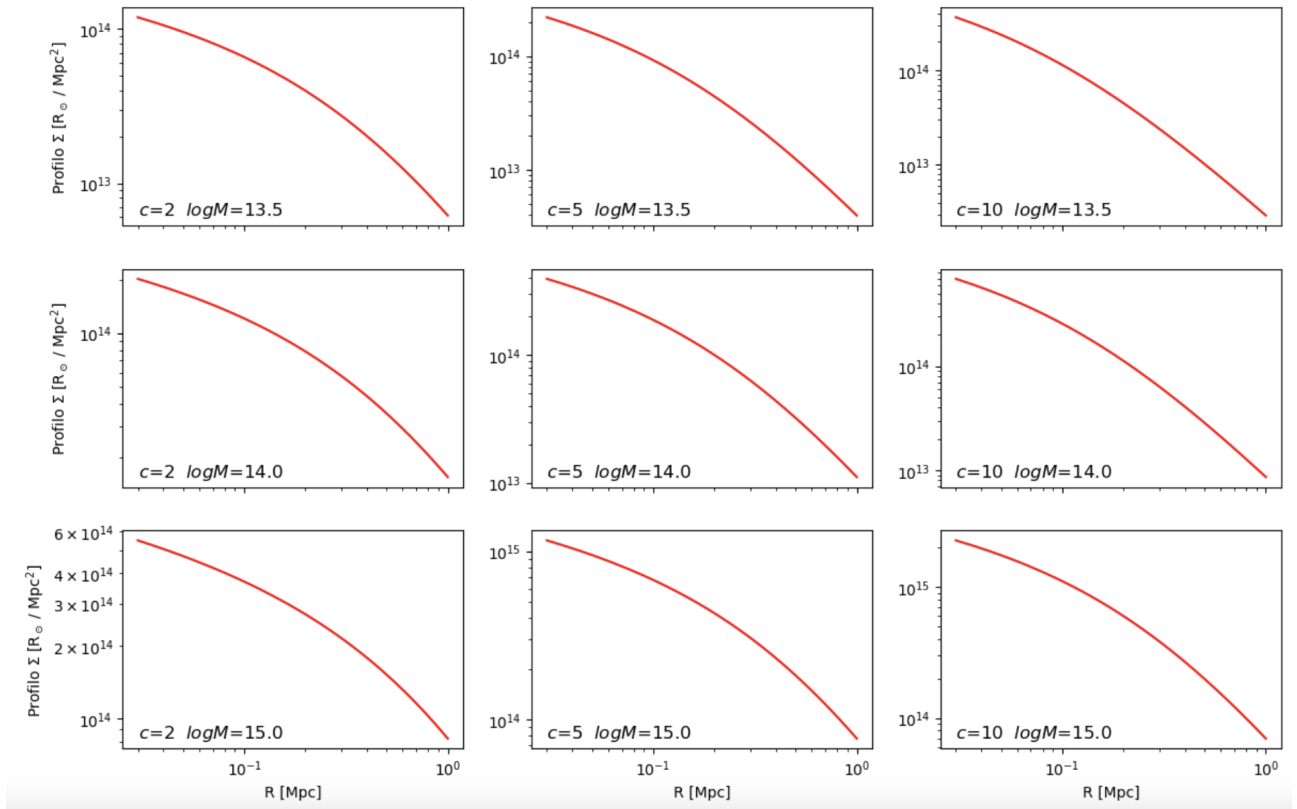


Figura 1: Andamenti in scala log-log del profilo Σ per le diverse coppie di c e M

Si può notare un andamento simile, di decrescenza, per le coppie di valori considerate. Inoltre all'aumentare della massa le curve sono meno piatte di quelle a massa minore.

3.2 Seconda richiesta

Dal file "halo-catalog.fit", aperto in lettura con il modulo "fits" dalla libreria "astropy.io", sono stati selezionati i 40103 profili di densità di aloni di materia, ottenuti da simulazione, i loro redshift e la ricchezza (una somma pesata del numero di galassie dell'alone).

I valori di R in cui tali profili sono stati calcolati sono stati aperti in lettura dal file "R-values.npy", tramite la funzione *load* di numpy.

Sono stati suddivisi i dati in base ai valori assunti dalla ricchezza, con estremi di tali intervalli definiti da una lista: [15, 20, 30, 45, 60, 200].

Si hanno così a disposizione 5 intervalli di ricchezza in cui è stato calcolato il redshift medio, il profilo di densità medio e la sua deviazione standard.

Per ogni intervallo è stato possibile graficare il profilo di densità in funzione dei valori di R forniti dal file "R-values.npy", che ne contiene 8. Per questo motivo sono state allocate due matrici 5×8 , una per il profilo di densità medio e una per la deviazione standard.

Sempre grazie a *subplots* è stato possibile graficare i valori di Σ in funzione degli R per i 5 intervalli di ricchezza.

Si può notare come tali andamenti richiamino quelli analizzati in precedenza, e siano tra di loro simili. Si vede inoltre che all'aumentare della ricchezza (gli ultimi intervalli) i valori assunti dal profilo di densità sono più grandi, in quanto è maggiore il numero di galassie nell'ammasso e di conseguenza la massa, che va così ad accrescere la densità. Di seguito il grafico.

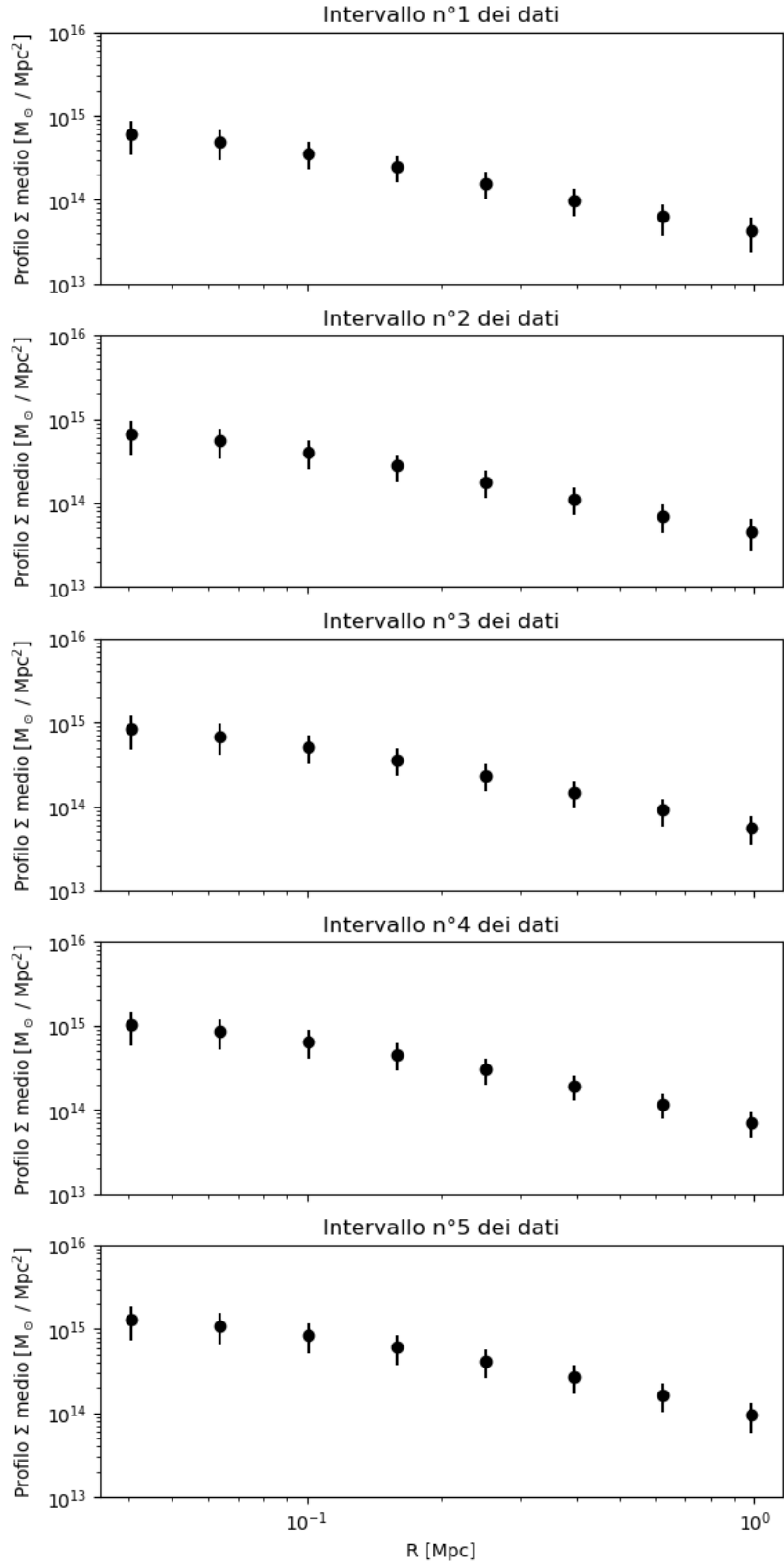


Figura 2: Andamenti in scala log-log del profilo Σ con relativa deviazione standard in funzione di R per i 5 intervalli di ricchezza

3.3 Terza richiesta

E' stato possibile stimare i valori di $\log M$ e concentrazione c per ogni intervallo di ricchezza tramite l'algoritmo di *emcee*.

A tal scopo sono state importate le matrici di covarianza dal file "covariance-matrices.npy" che vanno a definire la funzione di log-Likelihood, supposta Gaussiana. Come valore atteso nella definizione generale della gaussiana è stata ovviamente scelta la funzione di fit NFW valutata opportunamente. La funzione di log-Prior è definita uniforme se $\log M$ appartiene all'intervallo $[13, 16]$ e se c appartiene all'intervallo $[0.2, 20]$, altrimenti è nulla.

La scelta di lavorare con i logaritmi spezza i prodotti in somme, rendendo più agevole il codice.

Per cui, dopo aver definito la log-Posterior (la somma della log-Prior e la log-Likelihood) è stata allocata una matrice 5×2 di *bestfit* che andrà a contenere per i 5 intervalli i dati di best fit stimati tramite il software *emcee*, e un'array tridimensionale per salvare la chain.

Dopo l'esecuzione di *emcee*, di cui riportiamo in appendice il comportamento dei walkers in funzione del numero di passi per ambo i parametri, vengono mostrati a schermo i valori stimati di best fit, che riportiamo:

Valori di $\log M$ e c per l'intervallo n°1:	[14.33028327	2.67712052]
Valori di $\log M$ e c per l'intervallo n°2:	[14.40885302	2.79772732]
Valori di $\log M$ e c per l'intervallo n°3:	[14.56837255	3.06711335]
Valori di $\log M$ e c per l'intervallo n°4:	[14.72062524	3.31349494]
Valori di $\log M$ e c per l'intervallo n°5:	[14.92337943	3.57161715]

Figura 3: Valori stimati di best fit tramite *emcee*

3.4 Quarta richiesta

Grazie al modulo *pygtc* è stato possibile graficare e visualizzare le le posterior distributions dei walkers, ottenendo delle stime di distribuzioni di probabilità dei parametri $\log M$ e c . Di seguito i grafici a triangolo:

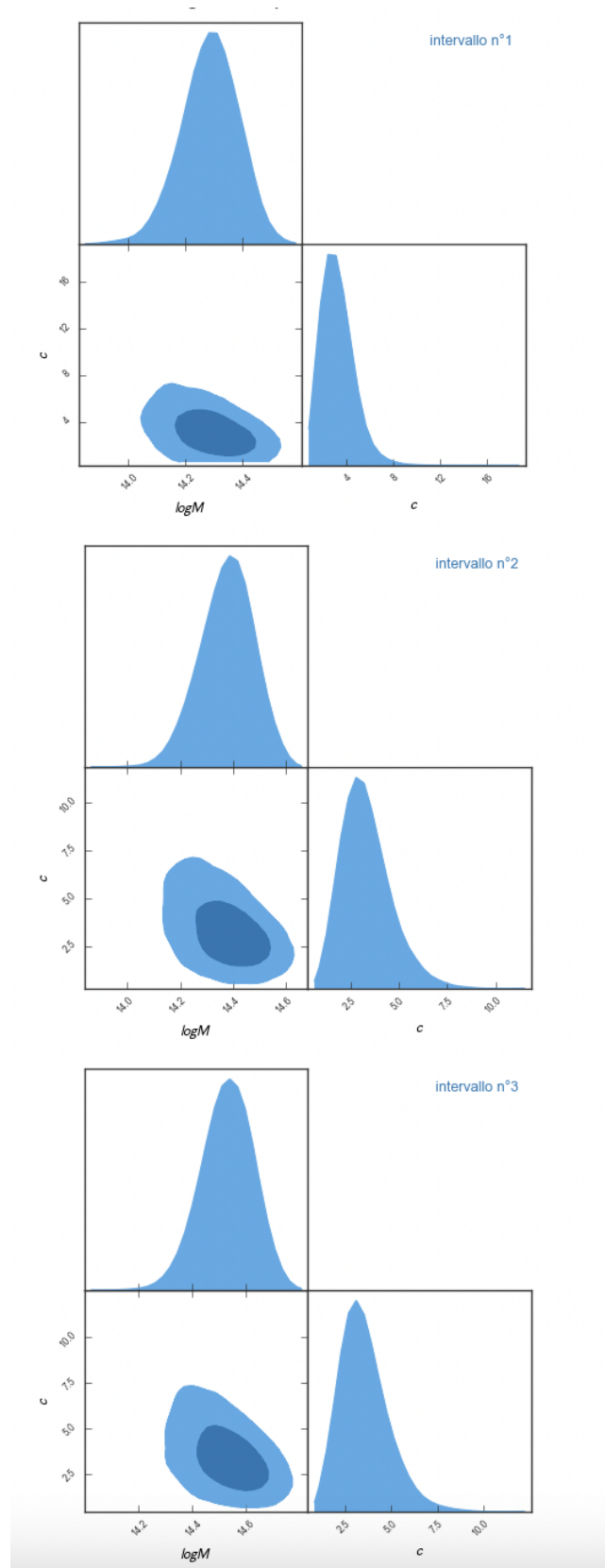


Figura 4: Posterior distributions per i primi 3 intervalli

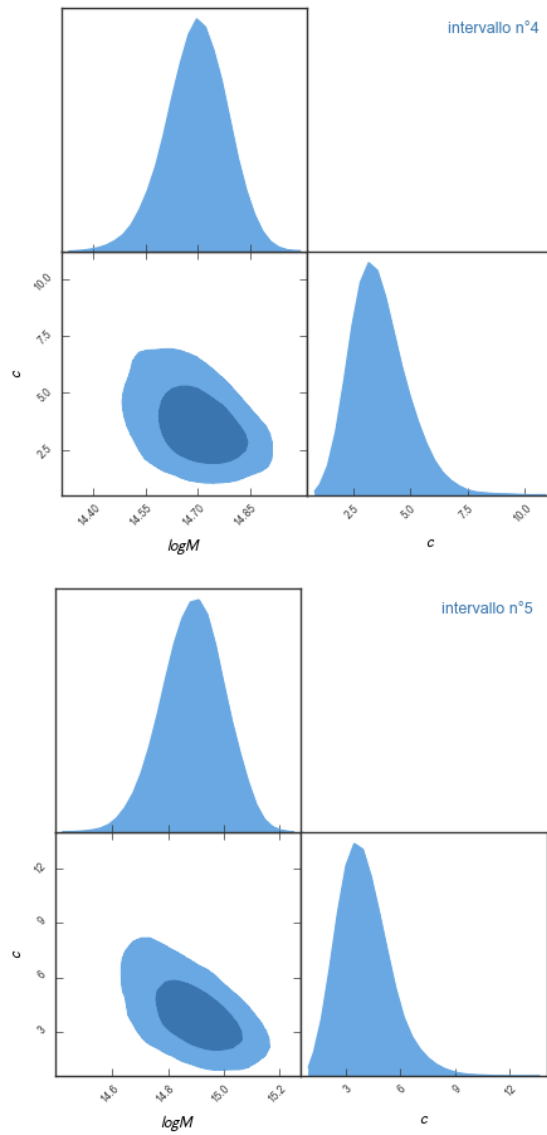


Figura 5: Posterior distributions per gli ultimi 2 intervalli

Ciò che si evince dai grafici è che la distribuzione di $\log M$ appare Gaussiana, mentre quella di c è un po' asimmetrica. Analogamente a prima i valori che assumono $\log M$ e c crescono negli intervalli finali, a ricchezza maggiore.

Si sono calcolati poi il valor medio e relativa deviazione standard dei dati dei parametri delle chain precedenti, con le funzioni di *mean* e *std* della libreria *Numpy*, per andare a confrontarli con i valori di best fit ottenuti.

I valori medi e relativa deviazione standard sono stampati a schermo:

```

Intervallo n°1
Valor medio e dev. std. di logM: 14.285604059667177 0.09787627764429908
Valor medio e dev. std. di c: 3.209140159476092 1.348953059189867

Intervallo n°2
Valor medio e dev. std. di logM: 14.374672781280136 0.09829328699514746
Valor medio e dev. std. di c: 3.293463210632859 1.2762257606078442

Intervallo n°3
Valor medio e dev. std. di logM: 14.53076931155583 0.09519935838391543
Valor medio e dev. std. di c: 3.5216469159507042 1.311482419928884

Intervallo n°4
Valor medio e dev. std. di logM: 14.694451813900688 0.08579821141674168
Valor medio e dev. std. di c: 3.6660669810438513 1.2037159412282155

Intervallo n°5
Valor medio e dev. std. di logM: 14.888045608253687 0.10710284747371275
Valor medio e dev. std. di c: 4.02218107139974 1.394844097199033

```

Figura 6: Valori medi e σ di $\log M$ e c dalla chain per i 5 intervalli di ricchezza

E' stato pertanto possibile confrontarli con i dati di best fit andando a plottare, grazie a *subplots*, per ogni intervallo di ricchezza il rispettivo valore di best fit dei parametri e il valor medio del parametro con relativa deviazione standard. Riportiamo di seguito tale confronto, in cui è possibile apprezzare la compatibilità di tutti i dati entro 1 deviazione standard.

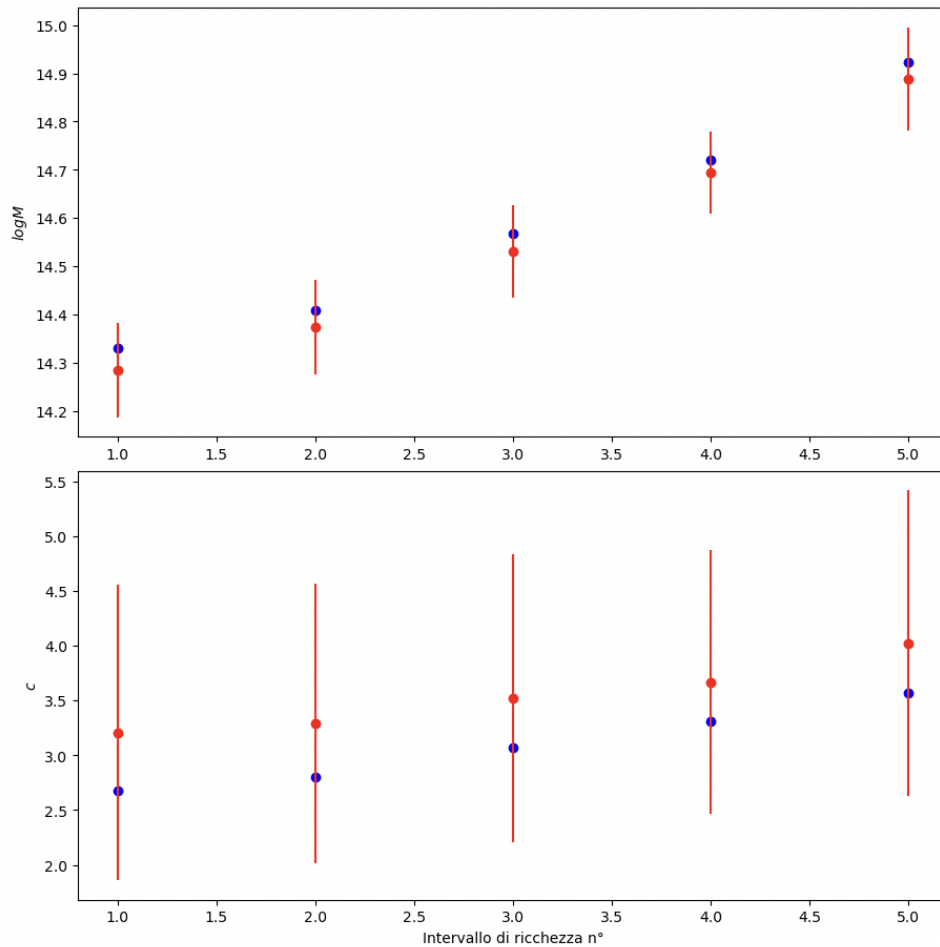


Figura 7: Per ogni intervallo di ricchezza in blu il best fit e in rosso il valor medio, con relativa σ

Anche qui si apprezza bene l'aumento di $\log M$ e c per gli intervalli più ricchi.

Si è svolto il test d'ipotesi per la funzione di fit NFW per ogni intervallo di ricchezza. Nello specifico si è eseguito un test di χ^2 , definendo come statistica di test:

$$t = \sum_i^5 \frac{(\bar{x}_i - \mu_i)^2}{\sigma_{\bar{x},i}^2}$$

dove \bar{x}_i è il profilo di densità medio per l'intervallo i -esimo, μ_i è il valore atteso per l'intervallo i -esimo, ossia la funzione NFW valutata sugli 8 valori di R con i parametri $\log M$ e c di best fit e a redshift medio. Mentre $\sigma_{\bar{x},i}^2$ è l'errore associato a \bar{x}_i .

Il χ^2 è a 6 gradi di libertà, perchè sugli 8 di partenza (i valori di R in cui calcolo la NFW) devo sottrarre 2 gradi di libertà in quanto ho stimato 2 parametri ($\log M$ e c).

Sono stati dunque calcolati i χ^2 per ogni intervallo e salvati in un array, mostrato a schermo:

Valori dei Chi quadri: [0.46890866 0.37731816 0.21717306 0.13750378 0.05396714]

Figura 8: Valori di χ^2 per i 5 intervalli

Ponendo un livello di significatività al 5%, si trova sulle tabulazioni degli integrali di χ^2 che l'ipotesi è rigettata se $t > 12.592$ per 6 gradi di libertà.

La nostra statistica di test ha prodotto un valore, in ogni intervallo, ben inferiore a quello che ci farebbe entrare nella regione critica. Possiamo considerare tale test superato, ulteriore conferma che la funzione Σ non è da rigettare come descrizione dei profili di densità medi degli aloni di materia oscura.