A halo model approach to describe clustering and emission of the two main star-forming galaxy populations for Cosmic Infrared Background studies

Giorgia Zagatti

Erminia Calabrese, Caterina Chiocchetta, Martina Gerbino, Mattia Negrello, Luca Pagano

Based on https://arxiv.org/pdf/2406.19069























ysis ↓ cosmic variance
ysis ↓ foregrounds contamination

Results

Why should we study extragalactic foregrounds

Contaminants of the primary CMB signal

Tracers of the matter density field clustering studies and insights on formation and evolution of the LSS of the Universe

Conclusions



clustering studies and insights on formation and evolution of the LSS of the Universe



Giorgia Zagatti - CIB

University of Ferrara - Italy

Introduction	Halo Model	CIB model	Datasets	Model fit to data	Results	Conclusions

AIM: build an analytic model of the CIB emission for all frequency and multipole ranges



Introduction	Halo Model	CIB model	Datasets	Model fit to data	Results	Conclusions

AIM: build an analytic model of the CIB emission for all frequency and multipole ranges

BUT lack of theoretical predictions about the model parameters



Introduction	Halo Model	CIB model	Datasets	Model fit to data	Results	Conclusions		
AIM: build an analytic model of the CIB emission for all frequency and multipole ranges								
BUT lack of theoretical predictions about the model parameters								
Model fit to data								





- Halo Model
- CIB Model
- Datasets
- Model fit to data
- Results
- Conclusion



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions

Halo Model: tool to predict non-linear matter and galaxy power spectra



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
Halo Mo	del:toolto	predict no	on-linear m	atter and gal	axy power	spectra

HOW?



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions

Halo Model : tool to predict non-linear matter and galaxy power spectra **HOW?**

1. Halo mass function

2. Halo bias

<u> Tinker et al. (2010b)</u>

3. Halo density profile

<u>Navarro et al. (1997)</u>



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions

Halo Model : tool to predict non-linear matter and galaxy power spectra **HOW?**

1. Halo mass function

2. Halo bias

<u> Tinker et al. (2010b)</u>

3. Halo density profile

<u>Navarro et al. (1997)</u>





Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions

Halo Model : tool to predict non-linear matter and galaxy power spectra **HOW?**

- 1. Halo mass function
- 2. Halo bias

<u> Tinker et al. (2010b)</u>

3. Halo density profile

<u>Navarro et al. (1997)</u>





Introduction	n Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
Halo	Model : tool to	predict nc	n-linear m	atter and gal	axy power	spectra
		I	HOW?			
1.	Halo mass fur	nction				
2.	Halo bias	2	matter pow	ver spectrum		
3.	Halo density]	profile			s, ≡ 27 ♦	
				4 = 4 4	4 4 4	ул. ж.
				4 N T		
						la di L

10 - -

. a.

Introductio	n Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions		
Halo Model : tool to predict non-linear matter and galaxy power spectra								
		I	HOW?					
1.	Halo mass fun	iction						
2.	Halo bias	2	matter pow	er spectrum				
3.	Halo density p	profile			16			
Halo O	occupation dis	tribution				9		







Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
		Halo Occ Tinker & Wetzel (2010)	upation di	stribution		



Giorgia Zagatti - CIB





























Introduction Halo M	odel CIB model	Data set	Model fit to data	Results	Conclusions
	HM: how dark matteris distributedHOD: how galaxiesare distributed				



Introduction Halo	Model	CIB model	Data set	Model fit to data	Results	Conclusions
	HM: h is dist HOD: are dis	ow dark matter ributed how galaxies stributed	Emissivit $j_{ u}$	y function (lun (z)	ninosity func	tion)) T



Introduction H	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
	H is J J H a:	IM: how dark matter s distributed IOD: how galaxies re distributed	Emissivit $j_{ u}$	y function (lun (z)	minosity fund ET I	rtion)

 $C_{\ell,\nu\times\nu\prime}^{CIB} = C_{\ell,\nu\times\nu\prime}^{\text{clust}} + C_{\ell,\nu\times\nu\prime}^{SN}$















Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
Planck d	lata (P14)	Planck da	ata- Lenz analy	vsis (L19)	SPIRE data ((V19)

Dataset	Total area (deg ²)	Frequency channels (GHz)	Multipole range	Binning scheme	Color Correction
P14	2240	217, 353, 545, 857	150-2500	Logarithmic	1.119, 1.097, 1.068, 0.995
L19	2240	353, 545, 857	75-2500	Linear	1.097, 1.068, 0.995
V19	90	600, 857, 1200	600-11000	Linear and logarithmic	0.9988, 0.9929, 0.9957



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
Planck o	lata (P14)	Planck da	ata- Lenz analy	vsis (L19)	SPIRE data ((V19)

	ET				
Dataset	Total area (deg²)	Frequency channels (GHz)	Multipole range	Binning scheme	Color Correction
P14	2240	217, 353, 545, 857	150-2500	Logarithmic	1.119, 1.097, 1.068, 0.995
L19	2240	353, 545, 857	75-2500	Linear	1.097, 1.068, 0.995
V19	90	600, 857, 1200	600-11000	Linear and logarithmic	0.9988, 0.9929, 0.9957



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
Planck o	lata (P14)	Planck da	ata- Lenz analy	vsis (L19)	SPIRE data ((V19)

	ET		clustering			
Dataset	Total area (deg ²)	Frequency channels (GHz)	Multipole range	Binning scheme	Color Correction	
P14	2240	217, 353, 545, 857	150-2500	Logarithmic	1.119, 1.097, 1.068, 0.995	
L19	2240	353, 545, 857	75-2500	Linear	1.097, 1.068, 0.995	
V19	90	600, 857, 1200	600-11000	Linear and logarithmic	0.9988, 0.9929, 0.9957	
	ET + 1	/ .T	٤	shot noise		


Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
Planck data (P14) Planck data- Lenz analy			vsis (L19)	SPIRE data ((V19)	





How to compare model predictions and observed data



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
How to co	mpare mode observed	l predictions data	and			
$C_{\ell,\nu_1\times\nu_2}^{CIB,\mathrm{data}} = .$	$\mathcal{A}_{\nu_1 x \nu_2} \times c c_{\nu_1}$	$\times cc_{\nu_2} \times C_{\ell,\nu}^{CL}$	$B, model u_1 imes u_2$			
calibration un	certainties co	olor corrections				

 $\mathcal{A}_{\nu imes
u \prime} = \sqrt{f_{\mathrm{cal}}^{
u} imes f_{\mathrm{cal}}^{
u \prime}}$



Introduction	Halo Model	CIB model	Data set	Model fit to data	Results	Conclusions
How to con	mpare mode observed	l predictions data	and			
$C_{\ell,\nu_1\times\nu_2}^{CIB,\text{data}} = \checkmark$	$\underbrace{\mathcal{A}_{\nu_1 x \nu_2} \times c c_{\nu_1}}_{\checkmark}$	$\times cc_{\nu_2} \times C_{\ell,\nu}^{CL}$	$B, model u_1 imes u_2$			
calibration und	certainties co	olor corrections				
$\mathcal{A}_{\nu imes u \prime} = \sqrt{f_0}$	$r_{ m cal}^{ u} imes f_{ m cal}^{ u\prime}$					

- emcee sampler
- Gaussian likelihood

$$\log \mathcal{L}(C_{\ell}^{\text{data}}|C_{\ell}^{\text{obs}}) \propto -\frac{1}{2} \sum_{\ell} \frac{\left(C_{\ell}^{\text{data}} - C_{\ell}^{\text{obs}}\right)^2}{\sigma_{\ell}^2}$$





- emcee sampler
- Gaussian likelihood

$$\log \mathcal{L}(C_{\ell}^{\text{data}}|C_{\ell}^{\text{obs}}) \propto -\frac{1}{2} \sum_{\ell} \frac{\left(C_{\ell}^{\text{data}} - C_{\ell}^{\text{obs}}\right)^2}{\sigma_{\ell}^2}$$





Giorgia Zagatti - CIB

- emcee sampler
- Gaussian likelihood

$$\log \mathcal{L}(C_{\ell}^{\text{data}}|C_{\ell}^{\text{obs}}) \propto -\frac{1}{2} \sum_{\ell} \frac{\left(C_{\ell}^{\text{data}} - C_{\ell}^{\text{obs}}\right)^2}{\sigma_{\ell}^2}$$

How to deal with the shot noise term

Results

$$C_{\ell,\nu\times\nu\prime}^{SN} = \mathcal{C}_{\nu\times\nu\prime} / SN_{\nu} \times SN_{\nu\prime}$$





- emcee sampler
- Gaussian likelihood

$$\log \mathcal{L}(C_{\ell}^{\text{data}} | C_{\ell}^{\text{obs}}) \propto -\frac{1}{2} \sum_{\ell} \frac{\left(C_{\ell}^{\text{data}} - C_{\ell}^{\text{obs}}\right)^2}{\sigma_{\ell}^2}$$

How to deal with the shot noise term

$$C_{\ell,\nu\times\nu\prime}^{SN} = \mathcal{C}_{\nu\times\nu\prime} \sqrt{SN_{\nu}\times SN_{\nu\prime}}$$

Maximal correlation between the shot noise contributions

$$\mathcal{C}_{\nu \times \nu \prime} = 1$$

Free parameters in the MCMC analysis [-1, 1]





- emcee sampler
- Gaussian likelihood

$$\log \mathcal{L}(C_{\ell}^{\text{data}} | C_{\ell}^{\text{obs}}) \propto -\frac{1}{2} \sum_{\ell} \frac{\left(C_{\ell}^{\text{data}} - C_{\ell}^{\text{obs}}\right)^2}{\sigma_{\ell}^2}$$

Free parameters of the fit

Università degli Studi di Ferrara

 $M_{min}^{ET}, \alpha_{ET} \qquad SN_{\nu}, f_{cal}^{\nu} \\ M_{min}^{LT}, \alpha_{LT} \qquad (\mathcal{C}_{\nu \times \nu'})$

analysis [-1,1]

Free parameters in the MCMC



























Data set

CIB model

 5σ tension between the best fit values of the minimum mass for ET galaxies.

Model fit to data

Results





Data set

CIB model

 5σ tension between the best fit values of the minimum mass for ET galaxies.

Model fit to data

Results

All the attempts to alleviate the tension among the two datasets have failed.





Data set

CIB model

 5σ tension between the best fit values of the minimum mass for ET galaxies.

Model fit to data

Results

All the attempts to alleviate the tension among the two datasets have failed.

Where does this tension come from



Data set

 5σ tension between the best fit values of the minimum mass for ET galaxies.

Model fit to data

Results

All the attempts to alleviate the tension among the two datasets have failed.

• Where does this tension come from

• data

CIB model

- model
- both







Backup slides: halo model tools and results

$$\begin{split} P_{uv}^{1h}(k) &= \int_{0}^{\infty} \hat{W}_{u}(m,k) \hat{W}_{v}(m,k) \frac{dn}{dm} dm; \\ P_{uv}^{2h} &= P_{mm}^{lin}(k) \prod_{n=u,v} \left[\int_{0}^{\infty} \hat{W}_{n}(m,k) b(m) \frac{dn}{dm} dm \right]. \\ \bullet \quad \frac{dn}{d\ln m} &= -\frac{1}{2} f(\sigma, z) \frac{\bar{\rho}}{m} \frac{d\ln \sigma^{2}}{d\ln m}, \\ f(\sigma, z) &= A(z) \left[\left(\frac{\sigma}{b(z)} \right)^{-\alpha(z)} + 1 \right] e^{-c(z)/\sigma^{2}} \\ \bullet \quad \delta_{h}(m, z_{1}|M, V, z_{0}) &= b(m, z_{1})\delta, \\ b(m, z_{1}) &= 1 - A \frac{v^{a}}{v^{a} + \delta_{c}^{a}} + Bv^{b} + Cv^{c}, \\ \bullet \quad \rho(r) &= \frac{\rho_{s}}{r/r_{s}(1+r/r_{s})^{2}} \\ \end{split}$$

$$P_{LT}^{lh}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [2N_{cent}^{LT}N_{sat}^{LT}u(m,k) + \\ N_{sat}^{VLT}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [2N_{cent}^{LT}N_{sat}^{LT}u(m,k) + \\ N_{sat}^{VLT}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [N_{sat}^{LT}u(m,k) + \\ N_{sat}^{VLT}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [N_{sat}^{LT}u(m,k) + \\ N_{sat}^{VLT}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [N_{sat}^{LT}u(m,k) + \\ N_{sat}^{VLT}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [N_{sat}^{LT}u(m,k) + \\ N_{sat}^{LT}N_{sat}^{LT}u(m,k) + \\ N_{sat}^{LT}N_{sat}^{LT}u(m,k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} [N_{sat}^{LT}u(m,k) + \\ N_{sat}^{LT}N_{sat}^{LT}u(m,k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \\ P_{mn}^{Ln}(k) &= \frac{1}{(\overline{n}_{gal}^{LT})^{2}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{sat}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \right]^{2}. \\ P_{mix}^{Ln}(k) &= \frac{1}{n_{gal}^{LT}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{gal}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \\ P_{mix}^{Ln}(k) &= \frac{1}{n_{gal}^{LT}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{gal}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \\ P_{mix}^{Ln}(k) &= \frac{1}{n_{gal}^{LT}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{gal}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \\ N_{mix}^{LT}(k) &= \frac{1}{n_{gal}^{LT}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{gal}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \\ N_{mix}^{LT}(k) &= \frac{1}{n_{gal}^{LT}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{gal}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \\ N_{mix}^{LT}(k) &= \frac{1}{n_{gal}^{LT}} \int_{0}^{\infty} dm \frac{dn}{dm} b(m) \frac{N_{gal}^{LT}}{\overline{n}_{gal}^{LT}}u(k,m) \\$$

Non-linear galaxy power spectrum in a halo model framework



Backup slides: best up-to-date literature constraints on the clustering parameters

	X12	P11	C13
$\log(M_{\min}^{\rm ET}/M_{\odot}h^{-1})$	12.09 ± 0.06	$11.95 \pm 2.10 - 12.21 \pm 0.51$	12.00 ± 0.04
$\alpha_{\rm ET}$	1.81 ± 0.04	$1.02 \pm 0.87 - 1.30 \pm 1.16$	1.55 ± 0.05
$\log(M_{\min}^{\mathrm{LT}}/M_{\odot}h^{-1})$	≡ 10.85	-	10.85 ± 0.06
α_{LT}	≡ 1		≡ 1







The luminosity function for ET galaxies is a convolution of the <u>halo formation rate</u> (obtained as the time derivative of the hmf) and the <u>galaxy luminosity distribution</u> (log-normal distribution).

$$\frac{\mathrm{d}n}{\mathrm{d}t_{\mathrm{vir}}} = \mathrm{d}n \frac{\mathrm{d}\ln\left(\nu f(\nu)\right)}{\mathrm{d}t_{\mathrm{vir}}}. \qquad P(\log L|\log \overline{L})\mathrm{d}\log L = \frac{\exp\{-\log^2(L/\overline{L})/2\sigma^2\}}{\sqrt{2\pi\sigma^2}}\mathrm{d}\log L,$$

$$\Phi(\log L, z) = \int_{M_{\rm vir}^{\rm min}}^{M_{\rm vir}^{\rm max}} \mathrm{d}M_{\rm vir} \int_{z_{\rm vir}^{\rm min}}^{z_{\rm vir}^{\rm max}} \mathrm{d}z_{\rm vir} \left| \frac{\mathrm{d}t_{\rm vir}}{\mathrm{d}z_{\rm vir}} \right| \frac{\mathrm{d}n}{\mathrm{d}t_{\rm vir}} P(\log L, z).$$

Required an empirical parameterization for the luminosity function of LT galaxies.

NB: from bolometric to freq-dependent LF we used a data-driven approach using SEDs (different SEDs for the two different galaxy populations)





Backup slides: emissivity functions



NB: broader contribution over redshifts from ET galaxies. The fact that we don't expect any contribution from ET galaxies at very low redshifts is encoded in the emissivity functions (cut-off)



Backup slides: P14 without 217GHz frequency channel included in the MCMC analysis



The exclusion of the 217GHz frequency channel makes the results of the fits more stable among the two scenarios.

Smaller shift in the 2D posterior distributions. To understand this we need to look at the full parameter space explored by the fit. We find a tension in the two scenarios for the value of the shot noise level of the 217 GHz frequency channel. Specifically, we obtain a value for the shot noise level which is higher in the case with free correlations than in the case with fixed correlations. We also see that the correlation coefficients involving the 217 GHz frequency channel are significantly lower than unity. We explain this behavior by noting the degeneracy between the shot noise level and the correlation coefficients at 217 GHz. which P14 data are not able to break. Specifically, they are anti-correlated, meaning that a shift of the shot noise level toward a lower value, closer to the one obtained in the first scenario, leads to higher values of the corr coeff.



Backup slides: shot noise



Parameter	Prior	Results $C_{\nu_1 \times \nu_2} = 1$	Results $C_{\nu_1 \times \nu_2}$ open
$\overline{\log(M_{\min}^{\rm ET}/M_{\odot}h^{-1})}$	[10.7,12.8]	$11.45_{-0.13}^{+0.16}$	11.12 ± 0.19
$\log(M_{\min}^{\text{LT}}/M_{\odot}h^{-1})$	[10.5,12.8]	$11.18^{+0.31}_{-0.27}$	$11.56^{+0.21}_{-0.17}$
$\alpha_{\rm LT}$	[0.2,3.5]	$1.337^{+0.063}_{-0.072}$	$1.436^{+0.075}_{-0.067}$
SN ₂₁₇	[0,50]	6.72 ± 0.76	22 ± 4
SN353	[50,500]	273 ± 16	276 ± 20
SN545	[400,4000]	1296 ± 120	1247 ± 100
SN ₈₅₇	[200,8000]	1827 ± 300	1899^{+400}_{-600}
f_{cal}^{353}	1 ± 0.0156	0.999 ± 0.014	1.010 ± 0.015
f_{cal}^{545}	1 ± 0.122	1.096 ± 0.032	1.094 ± 0.032
$f_{\rm cal}^{857}$	1 ± 0.128	1.289 ± 0.076	1.200 ± 0.072
$C_{217\times353}$	[-1, 1]	-	0.648+0.068
$C_{217\times545}$	[-1, 1]	-	0.469 ± 0.054
$C_{217\times 857}$	[-1, 1]	-	0.562 ± 0.070
$C_{353\times 545}$	[-1, 1]	-	$0.946^{+0.035}_{-0.031}$
$C_{353\times 857}$	[-1, 1]	-	$0.937^{+0.056}_{-0.042}$
$C_{545\times 857}$	[-1, 1]	-	$0.934_{-0.039}^{+0.046}$

INFN CARDIFF UNIVERSITY PRIFYSGOL CARDYD

Backup slides: $\alpha_{\rm ET}$ unconstrained





Backup slides: P14 vs L19





Backup slides: L19 analysis and covariance matrix



Fig. B.3. Covariance matrix of the L19 CIB measurements.

Fig. B.4. Correlation matrix of the L19 CIB measurements.

As initial step, we implement L19's pipeline to compute the CIB power spectra and extend it to build a Gaussian covariance matrix. We do this to explore the impact of possible bin-to-bin correlations in the fit and in the chi2. From these, we note that there is no significant correlation among different multipoles, but there is very significant correlation between the same bin at different frequencies. However, re-running the analysis including this new matrix showed no impact on the chi2 of the fit.



100



 v_1 Three different datasets obtained after a v_2 re-analysis of the maps.

- The first version of the data set was obtained applying the same neutral hydrogen column density (N_{HI}) threshold of 2.5 x 10^{20} cm⁻² across all three frequency channels (blue curve).
- The second data set applied a lower $N_{\rm HI}$ threshold of 1.5 x 10²⁰ cm⁻² across all channels to explore the impact of dust and in particular of dust residuals in the analysis (green curve).
- The third data set has been obtained following the prescription of L19 and imposing a different threshold for each frequency channel. Specifically we set $N_{HI} = 2.5 \times 10^{20} \text{ cm}^{-2}$, 2.0 x 10^{20} cm⁻² and 1.8 x 10^{20} cm^{-2} for 353, 545 and 857 GHz respectively (orange curve).


Backup slides: L19 analysis and possible dust contamination



different thresholds for the three frequency channels, showed no significant differences from the second data set, except for the 353 GHz frequency channel which provided more constraining power due to more data retained.

11.0

logM^{LP}_{min}

11.5

11.4

10.8

10.6

1.5

11.0 11.5

logM^{EP}_{min}

α^{Γb}

uim M^{LM} 11.0

the two higher frequency channels result in much higher best-fit values for the Mmin of ET galaxies. Recalling that this clustering parameter acts as a re-scaling of the power spectrum, the high values could possibly be hinting at an excess of power coming from dust residuals. more aggressive mask, so smaller contamination. The behavior of the posterior distributions of the clustering parameters changes significantly. There is no tension between the minimum masses of ET galaxies, supporting the hypothesis of a **dust residual**.



1.3 1.4 1.5

 α_{IP}

353 GHz

545 GHz

857 GHz

Backup slides: P14 model-data





1000

1500 2000

Backup slides: L19 model-data



 $- C_{\ell}^{CIB}$ $C_{\ell}^{CIB, clust}$ $C_{\ell}^{CIB, shot}$ L19 data

2000

Backup slides: V19 model-data







Backup slides: full model, fisher corr, fisher SNR - varying cosmology





Focus on cosmological parameters only



Purple: fixed cosmology entering the FG modelingBlue: varying cosmology in the FG modeling

Higher Fisher Signal-to-Noise Ratio

? What's next

- Go through MCMC analysis of different setups
- Find the best configuration to constrain the kSZ amplitude
- Further investigation on the tSZxCIB cross-correlation

