Observation model of a Planck-like cluster catalogue

J. Démoclès
IRFU/Service de Physique des Particules - CEA/DSM - CNRS, Bât. 141, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France

We present a model, relying on a simulation of the millimetric sky and taking into account known clusters properties, which predicts the observed flux distribution and completeness of a Planck-like cluster catalogue. A Fisher matrix application of this model is shown here.

1 Introduction

The galaxy clusters are the largest structures formed by gravitational collapse. Their abundance is a powerful cosmological probe. In particular, the Sunyaev-Zel'dovich effect provides a way of observing galaxy cluster in millimetric wavelenghts by the characteristic deformation it induces in the CMB spectrum. Furthemore, it provides a mass proxy which is found to be robust with the dynamical state of the cluster.

Nevertheless, the exploitation of the SZ cluster catalogues requires to quantify and understand the systematic effects which affect cluster detection. This can be obtained through a Monte-Carlo based on a reliable simulation of the sky taking into account instrumental effects and on an unbiased cluster extraction algorithm. In addition, we need a reliable association procedure which allows the identification of the recovered clusters among the detected sources.

We present an observation model which directly links the theoretical expectation for cluster abundance in function of SZ flux (Y) and redshift (z) to the observed one in terms of photometry, contamination and completeness.

During all this study, we use ΛCDM model. The cosmological parameters are assumed to be $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $\sigma_8 = 0.8$ and $h = 0.75$.

2 Monte-Carlo simulation

Our simulated maps contain the four main astrophysical components of the millimetric sky which are the primordial CMB anisotropies (excluding the dipole), the bright infra-red (IR) galaxies, the IR emission of the Galaxy and SZ clusters.

The cluster abundance is computed using Jenkins et al [3] mass function. The scaling relations to link cluster mass with its observable properties and cluster profiles are derived from published relation based on X-rays observation [2]. The clusters are assumed elliptical following the prescription of Cooray (2000).

We assume white instrumental noise so that the pixel noise is : $nois_{pix} = n_{eqT} \times \sqrt{N_{BOL} \times t_{pix}}$ where n_{eqT} is the noise at each frequency band, N_{BOL} the number of bolometers at this particular frequency and t_{pix} the time spent on each pixel. These are given following the experiment

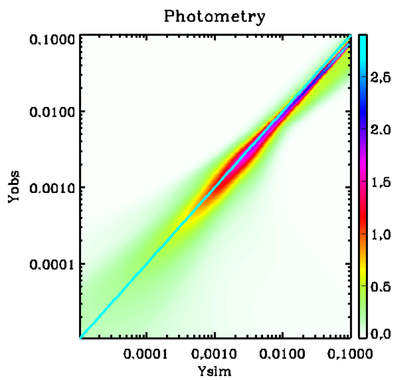


Figure 1: Model of photometry corrected from the threshold effect

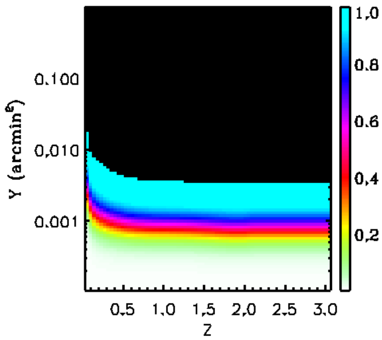


Figure 2: Model of completeness.

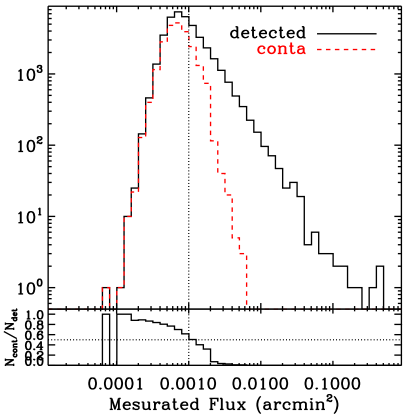


Figure 3: Contamination

characteristics as given for a Planck-like nominal survey. In particular, we simulate the sky outside the Galactic plan.

Using a wavelet ILC component separation, we recover a noisy SZ map that we convoluate with a Gaussian and apply SExtractor to extract the clusters. This provides a catalogue of detected sources without any assumption on the cluster shape or profile.

We need to identify the true detected clusters among all the sources to characterize the detected catalogue. This is done through an association algorithm based on cluster position in the map. We determine a spherical region around each cluster of the simulated catalogue starting by the brightest ones. The radius of this region d_{ASSO} varies in function of the flux and the size of the cluster. We plot the histograms of the number of associated clusters in function of their distance of association for different ranges of cluster flux and size to determine its dependancies. These distributions exhibit a global maximum before decreasing and flattening at $d_{ASSO} \approx 0.5 \times \theta_{200}$ for the resolved clusters ($\theta_{200} \geq 5$ arcmin). We fix a distance of association of 5 arcmin for the unresolved clusters. We estimate the rate of false association given by the flat tail of the distribution at 1% to 2% (depending on the cluster sizes).

3 Observation model

The observation model is a tool which allows to take into account the survey limitations in the predictions of the cosmological constraints. It quantifies the systematics in the photometry, completeness and contamination.

The photometry is the error on the measured flux. It is really important because it is the mass estimator. Therefore, this error has to be taken into account in addition to the expected dispersion of the mass-SZ flux relation when one intends to compare the observation with the theory. Then, for a theoretical number of clusters of N_{th} and given the experiment, we will observe N_{det} sources which is the sum of the true recovered clusters N_{obs} and the contamination N_{cont}. While the contamination is directly given by the association process, the number of detected cluster is obtained through the application of the selection function of the survey on the theoretical distribution of clusters. Thus, in the plan (Y, z), we can write :

$$\frac{dN_{obs}}{dzdY_{obs}}(z, Y_{obs}) = \int \frac{dN_{th}}{dzdY_{th}}(z, Y_{th}) \times C(z, Y_{th}) \times P(Y_{th}|Y_{obs}) \tag{1}$$

where $C(z, Y_{th})$ is the completeness and $P(Y_{th}|Y_{obs})$ the photometry of the observed catalogue.

Here, we computed each quantities separately.

3.1 Photometry

When we compare the reconstructed flux with the corresponding simulated one, there are mainly two aspects and each one can be described by a Gaussian :

- the first one is the increase of the dispersion. It is the global resolution of the photometry.
- the second one is the systematic effects at low and high flux. At low flux, we see a bias due to a threshold effect. As we approach to the detection threshold only the clusters for which the noise fluctuations make them brighter will be detected. This induces a bias illustrated by a systematic overestimation of the flux of the faint clusters. On the contrary at high flux, we underestimate the flux of the brightest clusters due to their extension on the sky.

In addition to the effects described above, we notice a systematic offset of the measured flux. Indeed it is 10% up to 20% under the simulated value for a measured flux spanning from $Y_{obs} \sim 9 \times 10^{-4}$ to $\sim 2 \times 10^{-1}\,\mathrm{arcmin}^2$. This indicates that the flux is not integrated over the full size of the cluster but corresponds instead to only $\approx 0.8 - 0.7 \times R_{200}$ [a].

To isolate the effect of the photometry from the selection effect of the flux limited selection, we symetrise the distribution at low flux. The value of the flux where the threshold effect starts is given by the position of the second gaussian when its mean becomes lower than the mean of the first gaussian. In the case of the Planck-like experiment it is found to be at $Y \approx 6 \times 10^{-3}\,\mathrm{arcmin}^2$. The result is shown in figure 1.

3.2 Completeness

The completeness of the catalogue comes from the efficiency of detecting a cluster depending on its characteristics, here taken as its redshift and integrated flux. Using the results of the Monte-Carlo, we estimate the completeness by computing the ratio between the distribution of the simulated clusters and the distribution of the recovered clusters in the plan (Y, z).

At each redshift, the completeness is well described by a Fermi-Dirac function :

$$P(Y, z) = 1 - \frac{1}{1 + \exp\left(Y - Y_A(z)\right)/Y_B(z))}$$

where $Y_A(z)$ is the value of the flux for which the completeness fall at 50% of the simulated sample and $Y_B(z)$ is the slope. The slope increases and the 50% threshold decreases as we go towards high redshift. Indeed, at high z, the signal becomes more concentrated whereas at low z, clusters are more extended so the signal is dimmer. Both the slope and the threshold evolution flatten around $z \sim 1$.

3.3 Contamination

The contamination is given by the number of sources detected but not associated to any simulated clusters. As expected, it is almost 100% of the sources detected at low flux ($Y \sim 10^{-4}\,\mathrm{arcmin}^2$) and fall close to zero at high flux with a 50% contamination at $Y \sim 10^{-3}$ (figure 3).

---

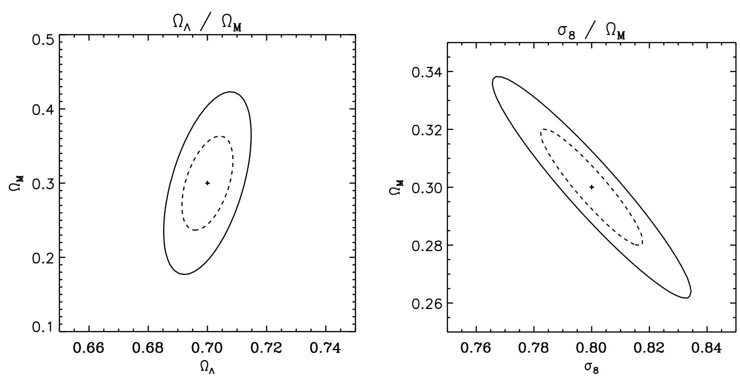
[a] The radius where the density of the cluster is 200 times larger than the critical density of the Universe $\rho_c(z)$ at the redshift of the cluster

4 Fisher analysis

The method and results presented here are general to compute the selection function of a SZ survey. Only the parameters of the fits used to model the photometry and the completeness change. We compute the observation model in the case of a smaller survey but with a higher resolution allowing a larger fraction of clusters to be resolved.

A straightforward application of this model is to use it to compute the observed abundance of clusters $dN_{obs}/dzdY$ which is used to constrain the cosmological parameters. We use the formalism of the Fisher matrix to compute the sensibility on Ω_M, Ω_Λ and σ_8. We compare the results obtained assuming a simple cut in flux at 50% of completeness at high redshift, with those using our model.

The result is shown in figure 4 a) and b). Using a simple cut in flux leads to an optimistic estimation of the constraints. The difference is mainly due to the fact that we overestimate the number of clusters especially at low redshift compared to the case based on the model of completeness.



a) **b)**

Figure 4: **Prediction of the cosmological constraints** assuming all the other parameters known : a) In the plan Ω_Λ, Ω_M. b) In the plan σ_8, Ω_M. The dashed lines are the constraints obtained with a simple cut in flux in the theoretical abundance of clusters and the solid lines are obtained with the application of the model of completeness.

5 Discussion and conclusions

We presented an observation model which offers analytical expressions to characterise SZ cluster catalogues. It includes a photometric characterisation, a completeness and contamination modeling for large SZ surveys.

It is a way to directly relate theoretical cluster distribution to the observed one taking into account all the systematics due to the reconstruction of the clusters. In particular, the angular size of the clusters is an issue for both the completeness and the photometry, extended clusters (mostly low redshift ones) being less accurately reconstructed. It is crucial to take this fact into account when exploiting SZ cluster catalogues to cosmology.

Acknowledgments

I would like to thank the organizers for a very interesting and enjoyable conference and my colleagues from the SPP and SAp at CEA-Saclay for useful discussions and suggestions.

References

1. J.B. Juin, D. Yvon, A. Réfrégier et C. Yèche, A&A **465**, 57-65, Jan. 2007
2. Arnaud M., Pratt G. W., Piffaretti R. et al , ArXiv e-prints, Oct 2009
3. A. Jenkins et al, *MNRAS* **321**, 372-384, Sep. 2001