Diploma Thesis

Bence Kocsis

THE DETECTION OF CLUSTER EDGES WITH THE SUNYAEV ZEL'DOVICH EFFECT

Thesis advisors

Zsolt Frei and Zoltán Haiman

Eötvös Loránd University of Sciences, Hungary
June 2004

Contents

1 Magyar nyelvű kivonat

A galaxisklaszterek megfigyelése a Sunyaev-Zel'dovich effektus segítségével

A Sunyaev-Zeldovich (SZ) effektus egyre értékesebb megfigyelési lehetőséget nyújt az asztrofizika számára. A jelen technológiai fejlődésnek köszönhetően az SZ effektus immár alkalmas a masszív klaszterok feltérképezésére, az intraklaszter közeg (ICM) vizsgálatára és a kozmológiai paraméterek megszorítására.

A termikus SZ effektus a kozmikus mikrohullámú háttérsugárzás (CMB) másodlagos fluktuációja, ami forró intergalaktikus gázzal való kölcsönhatás során jön létre. Az utolsó szórási felületről érkező hideg CMB fotonok inverz Compton-szóródnak, ami kicsiny intenzitás gyengülést okoz kis illetve erősödést okoz nagy energián ($\nu \lesssim 218$GHz ill. $\nu \gtrsim 218$GHz). Az SZ effektus a CMB anizotrópia domináns forrása kis szögskálán.

Jelen műszerek jelentős jel-zaj aránnyal detektálnak klasztereket. A következő generációs műszerek által alkotott galaxis klaszter térképek az égbolt jelentős részét le fogják fedni. A klaszterek számosságát illetve ezek szögfüggő teljesítményspektrumát többen vizsgálták (lásd pl. [3] [4] [5] [6]). A felmérések eredménye meglehetősen lenyűgöző, a következő generációs SZ műszerek várhatóan naponta több klasztert kell, hogy észleljenek.

Az SZ effektus megfigyelése napjainkban kezdi feltárni a benne rejlő lehetőségeket. A jelenleg működő SZ megfigyelések közé tartozik a BIMA, a Diabolo, a SuZIE, a Ryle Távcső, az OVRO és a CBI. Ezeken kívül a következő generációs bolometrikus és interferometrikus műszereinek építése a végső stádiumhoz közelít. Ilyenek az SZA, az AMI és az AMiBA teleszkóprendszerek. Ezen műszerek több nagyságrenddel lesznek érzékenyebbek, mint az elődjeik. A jelen évtized második felében tovább folytatódik az SZ műszerek fejlesztése, hisz az elmúlt évben számos hosszú távú program nyert támogatást: az APEX, az ACT, az SPT, a Planck és az ALMA.

A dolgozat legfőbb célkitűzése annak a vizsgálata, hogy az említett következő generációs műszerek közül az ALMA milyen mértékben képes a galaxishalmazokról alkotott bizonyos ismereteinket javítani. A klaszterek belsejében levő anyag SZ lenyomatának mérésével, az ionizált gáz sűrűségeloszlását lehet meghatározni és ebből következtetéseket lehet levonni a sötét anyag struktúrára vonatkozóan.

A galaxisfejlődés klasszikus elmélete szerint [14] a klaszterek önhasonló akréción keresztül jönnek létre. A kis méretű sűrűsödési perturbációk felé gömbszimmetrikus anyagáramlás indul. Az elmélet jóslata szerint, tiszta ütközéses vagy ütközésmentes anyag esetén, – vagy akár a kettő együttes keveréke mellett – a folyamat végén stacionárius önhasonló egyensúly áll be az anyagtöbblettel rendelkező sűrűségingadozás körül. A klasztergáz beesését megállítja egy, a középpontból kifelé haladó, lökéshullám. A gáz sugárirányú sűrűségfüggvényében a lökés közelében egy ugrás alakul ki, ahol a sűrűség a belső oldalon lényegesen megnövekedhet. Vannak bizonyos megfigyelések, amelyek szerint együttes optikai-, rádió- és röntgenmegfigyelések segítségével [15] ezt a sűrűség lökést már meg is figyelték bizonyos keletkező klaszter prototípusok körül. Elméletileg ezen diszkontinuitás léte a SZ effektus mérésével is bizonyítható.

Mindezidáig a Sunyaev-Zel'dovich effektusra vonatkozó tanulmányok nem vették figyelembe a klaszterek peremén megjósolt gázsűrűség radiális levágását. A leggyakrabban használt egyszerű leírás az ún. „β-modell" (lásd Cavalier, [17]). Ezen modell egy egyszerű empírikus illesztési formulát jelent a gázsűrűség profilra vonatkozóan, amely független a háttérben rejlő sötét anyag eloszlástól, mindazonáltal nem mond ellent a jelen SZ és röntgen megfigyelés-eknek (lásd Mohr, [18]). Ettől részletesebb modellek megpróbálják levezetni a gáz sűrűség profilt azáltal, hogy izotróp hidrosztatikai egyensúlyt feltételeznek a gáz és sötét anyag eloszlás között. Az öngravitáló ütközésmentes sötét anyag sűrűségprofilját numerikus N-test szimulációk segítségével hozzák létre. A két csoport Navarro et al. [21] illetve Moore et al. [22] megegyeznek abban, hogy a profilt egy univerzális függ-

vény írja le, amely a külső tartományban r^{-3} szerint csökken. A belső tartományokra vonatkozó hatványkitevőt tekintve azonban eltér a véleményük, -1 ill. -1.5 a két vitatott érték. A gázsűrűségre vonatkozó megoldást ezen sötét anyag háttéren adja meg például az ún. „levágott izoterm gömb modell" (lásd Shapiro, [16]). Az izoterm elképzelést módosítja az „előfűtött modell" (lásd Holder és Carlstrom [19]) egy hőmérsékletfüggő, nem gravitációs eredetű entrópia alapot feltételezve a klaszter magjának közelében. Egy következő megközelítés szerint (lásd Komatsu és Seljak, [20]), – a sötét anyaghoz hasonló módon – a gáz sűrűség is egy univerzális önhasonló alakot vesz fel. A legkorszerűbb hidrodinamikai numerikus szimulációk [7] összhangban vannak a fenti modellekkel. A klaszter számláló statisztikára és a szögfüggő teljesímény spekt-rumra vonatkozó röntgen és SZ effektus jóslatok némileg eltérnek, azonban a megfigyelési mélység és érzékenység mindezidáig nem jutott el arra a szintre, hogy dönteni tudjon a különböző modellek között.

A dolgozatban Komatsu és Seljak izotróp, univerzális, önhasonló gázra vonatkozó feltételezését követtem [20], de bevezettem egy lineáris levágást a gázsűrűség eloszlására vonatkozóan egy véges sugáron kívül. A sötét energiát is tartalmazó, lapos geometriájú, hideg sötét anyag modellt (ΛCDM) használtam a jelen WMAP méréseknek megfelelő paraméterekkel: $(\Omega_\Lambda, \Omega_{\rm dm}, \Omega_{\rm b}, H_0) = (0.7, 0.253, 0.047, 70 \times {\rm km/sMpc})$.

A dolgozatban megmutattam hogy a közel-jövőben megvalósuló Atacama Large Millimeter Array (ALMA) a Sunyaev Zel'dovich effektus vizsgálatával elsőként alkalmas lehet a klaszterek pontos gázeloszlásának vizsgálatára. A program realitását jelzi, hogy az ALMA prototípus teleszkóp, – egy 12m átmérőjű, multifrekvenciás rádióteleszkóp,– az APEX, már várhatóan idén, 2004 második felére elkészül. A végső állapotát az ALMA rendszer 2012-ben éri el. A cél hatvannégy APEX teleszkóp összehangolt üzemeltetése. A műszer minimális illetve maximális karhosszúsága 150 m ill. 10 km. Ezen műszer tervezett érzékenységére és lineáris méretére vonatkozó számítások alapján reális feltételezés 10μK-nél kisebb SZ hőmérsékleti zajt tekinteni 2" szögfelbontás és 100GHz munkafrekvencia mellett. Ezt az érzékenységet

körülbelül 30 óra megfigyelési idővel lehet elérni. A megfigyelési időt tovább növelve a detektor bizonyta-lanság tovább csökkenthető.

A gömbszimmetrikusnak feltételezett, sztatikus esetre vonatkozó, hidrodinamikai egyenleteket önhasonló gázra oldottam meg, a megfelelő sötét anyag hátterének figyelembevételével. A sötét anyag profilt a szokásos módon α-val paramétereztem, és minden számolást α mindkét értékére elvégeztem, azaz $\alpha = 1$ ill. 1.5-re. Ebben a tárgyalásban a sűrűség a sugár függvényében kicsiny, többszörös viriálsugáron kívül is véges járulékot ad. Ezeb kiinduló modellhez képest a sűrűségprofilban bevezettem egy hipotetikus lineáris radiális levágást, ahol a sűrűség hirtelen zérusra csökken. Ez a feltételezés konzisztens a klaszterek fejlődésének elméletével (lásd [14] and [15]).

Erre az eloszlásra vonatkozóan kiszámoltam az SZ kép megváltozását az ALMA detektor esetén. A detektor és a háttér együttes hatására 10μK Gauss fehér zajt feltételeztem. Megkonstruáltam a korábbiakban kiszámolt önhasonló sűrűségprofilhoz illesztett optimális szűrőt. A megfelelő likelihood függvényt használtam a levágásos és a levágás nélküli SZ hipotézisprofilok közötti döntés elvégzésére illetve a döntés szignifikancia szintjének vizsgálatára. Fontos hangsúlyozni, hogy ezen teszt gyakorlatilag modellfüggetlen, hiszen csak kis mértékben függ a klaszter belső tartományaitól azonban a külső régióban –ahol az irodalomban található különböző modellek azonos eloszlást jósolnak – szignifikáns függés adódik. Kétféle módon számoltam ki a jel-zaj hányadost, négy illetve kettő szabad paraméterrel. Az előbbit jelöljük Model I-nek az utóbbit pedig Model II-nek. Ezek a paraméterek az I esetben a peremhez tartozó radiális távolság, $x_{\rm max}$, a levágás „vastagsága", D, a koncentráció paraméter, c, és a középponti SZ hőmérséklet csökkenés $T(0)$. A II esetben a c és a $T(0)$ paramétereket a klaszterek gömb-szimmetrikus összeomlására vonatkozó elméletéből kaptam meg, és ez körül semmilyen Δc ill. $\Delta T(0)$ szórást nem engedtem meg. Egyéb paraméterek, mint a viriál tömeg, $M_{\rm vir}$, a vöröseltolódás, z, a belső sötét anyag exponens, α, és a klaszter szögmérete,Θ, lényegesen nagyobb pontossággal mérhetők, ezért ezeket nem változtattam. Az eredmények 10%-tól

kevésbé térnek el a Modell I és a Modell II esetén. A jel-zaj hányados egy nagy tömegű klaszter esetén ($M_{\rm vir} = 10^{15} \times M_\odot$) $S/N = 470$, egy normál klaszterre pedig ($M_{\rm vir} = 10^{14} \times M_\odot$) $S/N = 16$. Ebben a számolásban a c és a $T(0)$ paraméterekre feltételeztem, hogy megegyeznek a korrekt értékekkel, mind a levágásos mind a levágás nélküli esetben.

Ezután, egy realisztikusabb esetet tekintettem, amelyben a klaszter paramétereit, – levágással vagy anélkül, – külön-külön illesztettem egy adott megfigyelés esetén. Ebben az esetben az derült ki, hogy a határ nélküli esetben jelentős szisztematikus hiba adódhat a c illetve $T(0)$ értékeire vonatkozóan. Ennek feltétele, hogy a klaszternek valóban létezzék határa, mégpedig legfeljebb kb. 1 virál sugár körül. A jel-zaj arány mindazonáltal, a – arra a kérdésre, hogy létezik-e a klaszternek pereme, – legfeljebb 56%-al csökken le az eredeti egyszerűbb számoláshoz képest.

Végül megvizsgáltam a paraméterbecslés pontosságát a peremet tartalmazó modell esetén. Eredménynek a következő értékeket kaptam:

$$\begin{aligned}(\Delta c/c, \Delta T(0)/T(0), \Delta x_{\rm max}/c, \Delta D/c) &= (2.8\%, 3.7\%, 2.9\%, 13\%) \\ &= (0.45\%, 0.2\%, 0.1\%, 0.5\%)\end{aligned}$$

normál illetve masszív klaszter esetén. Mivel α és c között létezik egy egy-egy értelmű megfeleltetés, $c(\alpha = 1)/c(\alpha = 1.5) = 1.7$, az SZ effektus képes eldönteni, hogy melyik α modell írja le jobban a valóságot. A sötét anyagra vonatkozó vitát ezzel a jövőben biztosan meg lehet oldani. Az itt feltüntetett számokat a Fisher mátrixra vonatkozó közelítéssel kaptam. A dolgozatban megvizsgáltam az egzakt valószínűségi eloszlást is, amik nagy pontossággal (1%-on belül) megegyeztek a Fisher mátrixból származtatott értékekkel.

A dolgozatban minden számolást 18 féle paraméterválasztás esetén végeztem el. Ez $3 \times 2 \times 3$ blokkokra bomlik: a valódi klaszter viriáltömegét 10^{13}, 10^{14} és $10^{15} \times M_\odot$, a valódi D értéket 0.01 és $0.5c$, a valódi $x_{\rm max}$ értéket 1, 1.5 és $2c$ között változtattam. Ezen kívül a legegyszerűbb esetekben, $(M, D, x_{\rm max}) = (10^{15} M_\odot, 0.01c, c)$ leellenőriztem a vöröseltolódás és α függést is. Általában azt kaptam, hogy a klaszter széle nagy tömeg és kis $x_{\rm max}$

értékek esetén detektálható legpontosabban. A vöröseltolódás függvényében 1 nagyságrend változás jöhet létre a jel-zaj arányban, és annak $z \approx 0.4$ körül minimuma van. Az α ill. D konkrét értéke viszont a hatékonyság szempontjából közömbös.

Végeredményben tehét az SZ effektus mérésével az ALMA redszer nem csak a klaszter szélének jelenlétét fogja tudni kimutatni, hanem annak helyét is nagy pontossággal meg fogja majd állapítani. Ezen ismeretek egészen biztosan elő fogják segíteni a klaszterek belsejében lévő anyag szerkezetének pontosabb megértését.

2 Introduction

The Sunyaev-Zel'dovich (SZ) effect has recently become a valuable observational tool [1]. Observational programs are beginning to measure the SZ effect to map out massive clusters of galaxies, study the intracluster medium (ICM), and constrain cosmological parameters [2].

The thermal SZ effect is a secondary distortion of the cosmic microwave background (CMB) spectrum caused by hot intergalactic gas along the line of sight to the surface of last scattering. The cool CMB photons undergo inverse Compton scattering on the hot electrons, gaining on average a small amount of energy in the process, creating an intensity decrement at low frequencies ($\nu \leq 218$GHz) and an increment at high frequencies. The SZ effect is the dominant source of CMB anisotropy at small angular scales. Figure 1 displays the CMB anisotropy and the SZ component.

Current instruments are now regularly detecting and imaging clusters at high signal-to-noise, and the next generation of instruments should be capable of mapping fairly large portions of the sky as a means of finding clusters of galaxies. Several works (e.g., [3] [4] [5] [6]) have predicted the number of clusters that could be expected in future SZ surveys and their angular power spectrum. The survey yields are quite impressive; the next generation of SZ instruments should be able to detect several clusters per day.

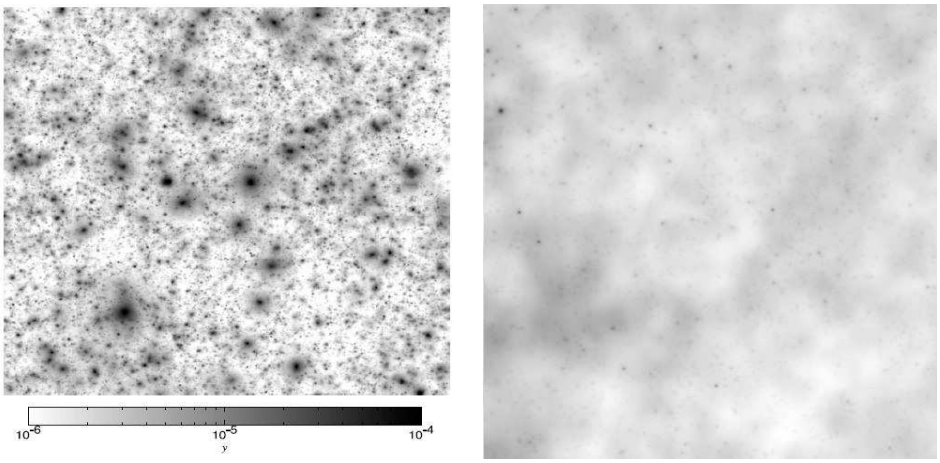

Figure 1: Maps of the simulated thermal SZ effect adopted from [7]. The left panel shows only the SZE contribution, while the right shows both the primary and the secondary CMB anisotropy. The field of view is roughly $1^\circ \times 1^\circ$ on both plots.

Observations of the SZE are just beginning to demonstrate their poten-

tial. Currently commissioning SZE observations include the BIMA[1] [8], the Diabolo [9], the SuZIE[2] [10], the Ryle Telescope[3] [12], the OVRO[4] [13], and the CBI[5]. New generation bolometric and interferometric instruments are approaching the final stages of development, such as the BOLOCAM[6] [11], the SZA[7], the AMI[8], and the AMiBA[9]. These instruments are anticipated to be over an order of magnitude more sensitive and will soon begin routine observations. The SZ instruments will continue increasing its capabilities and performance throughout the decade, as several additional new generation SZE instruments have recently been funded. These are the APEX[10], the ACT[11], the SPT[12], the Planck[13], and the ALMA[14].

The present study aims to analyze the extent that this near-future precision technology will improve our understanding on galaxy cluster physics. In particular we make predictions for the ALMA system. Measuring the SZ imprint of the ICM, it is possible to map the ionized gas density distributions and make predictions on the underlying dark matter structure.

The classical theory of cluster evolution [14] states that the clusters form by self-similar accretion and secondary infall on a compact density perturbation. The theory predicts that collisional or collisionless matter, or the mixture of the two evolve to a stationary self-similar equilibrium within a neighborhood of the overdense density perturbation. The infall of cluster gas is halted by a shock wave which propagates outward from the center. The gas density has a discontinuity at the shock, and the density is much larger in the inner region. There are indications that this density shock has already

[1] An array of ten 6.1m mm-wave telescopes located at Hat Creek, CA, and operated by the Berkeley-Illinois-Maryland-Association.

[2] Sunyev-Zel'dovich Infrared Experiment, a six-element 140 GHz bolometer array.

[3] An array of eight 13m telescopes used in interferometric mode, located in Cambridge, England.

[4] An array of six 10.4m telescopes located in Owens Valley, CA, and operated by Caltech.

[5] Cosmic Background Imager, an array of 13 0.9m telescopes, sensitive to 3'-20'.

[6] A 151-element bolometer array, operated by the Caltech Submillimeter Observatory.

[7] Sunyaev-Zel'dovich Array, an array of eight 3.5m telescopes which will be deployed with the existing OVRO and BIMA arrays.

[8] ArcMinute Micro-Kelvin Imager, an array of ten 3.7m telescopes, located in Cambridge, England.

[9] Array for Microwave Background Anisotropy, 19 1.2m telescopes operated at 90 GHz.

[10] Atacama Pathfinder Experiment, a 12m telescope ALMA prototype, end of 2004

[11] Atacama Cosmology Telescope, 6m telescope with a 1024-element bolometer array, 3 colors, planned to start running by November 2006.

[12] South Pole Telescope, 8m telescope with a 1000-element bolometer array, operational by 2008.

[13] Planck Explorer satellite, launch planned for 2008

[14] Atacama Large Millimeter Array, an array of sixty four 12m telescopes, planned for 2012

been observed around forming cluster prototypes with combined optical, radio, and X-ray analysis [15]. In principle, the existence of the discontinuity can also be detected using the SZ effect.

Previous studies on cluster phyisics with an emphasis on the SZ effect do not account for the anticipated cutoff in the cluster gas density profiles. The simplest popular discription is the "β model" [17]. The model provides a simple fitting formula for the gas density profile, independent of the underlying dark matter profile, and it is in accord with present SZ and X-ray observations [18]. More sophisticated models derive the gas density profile by assuming isotropic equilibrium between the gas and dark matter distributions. The dark matter density profile is obtained by numerical N-body simulations. The two groups [21] and [22] agree that the profile is universal with an r^{-3} ringdown in the outer regions but argue on the exponent in the inner regions between -1 or -1.5. A solution for the gas density on this dark matter background [16] is the truncated isothermal sphere model. The "preheated model" of [19] assumes a temperature dependent nonzero entropy floor in the cluster core, modifying the isotherm model. Another approach is presented in [20] where the gas density assumes a universal self-similar form. State-of-the-art hydrodynamic numerical simulations [7] are in accord with the above models. The implications on the X-ray and SZ signatures regarding the cluster number count statistic and the angular power spectrum are somewhat different in the various models, but as of now the observation depth and sensitivity has not reached the level to decide which model comes closest to reality.

In this study we follow the approach of [20] for isotropic self-similar equilibrium, but introduce a linear cutoff in the gas density distribution at a finite radius. For the cosmological model, we use the ΛCDM cosmology. We approximate the angular power spectrum of the unresolved CMB anisotropy as a white gaussian noise. We calculate the SZ surface density profile for the models with and without the density cutoff, and examine if these models can be distinguished with the future ALMA system in the noise background. For the comparison we derive the optimal matched filter, and analyze the signal to noise ratio at which the distinction can be made between the two models. We then calculate the probability distribution and the level of significance of the comparison. The score of the test is dominated by the contribution of the region around the cutoff in the outer region. Since all other models agree on the SZ imprint of this domain, our result for detecting cluster edges is mostly model independent. Finally we examine the precision at which the various parameters of the self-similar model – including the radial distance and the "width" of the cutoff,– can be measured.

3 The proposal for the Atacama Large Millimeter Array

This study offers a possible application to the ALMA project regarding the observation of the gas density cutoff. Thus, it is necessary to review the expected technical specifications of this instrument.

ALMA will be comprised of some sixty four 12-meter sub-millimeter quality antennas, with baselines extending from 150 m up to 10 km. Its receivers will cover the range from $\nu = 70$ to 900 GHz. Anticipated SZ temperature sensitivity for a $\Delta t_s \approx 60s$ integration period is between $\sigma_N \approx 100\mu$K and 24 mK for the compact configuration depending on the frequency of the detection.

We assume antennas defined by the request for proposals for construction of the prototype antennas. The expected aperture efficiency has been taken from [26]. The receiver sensitivity is assumed to be 40K - 50K at 100GHz, $6\frac{h\nu}{k_B} + 4$K up to 650GHz and $8\frac{h\nu}{k_B}$ at 850GHz. The atmosphere is assumed to be characterized by 1mm of precipitable water vapor, and opacities are calculated at the various frequencies according to [27].

Generally a baseline B corresponds to a beam diameter

$$\Delta\phi = \frac{c}{\nu B} = 2\text{"} \times \left(\frac{\nu}{100\text{GHz}}\right)^{-1} \left(\frac{B}{300\text{m}}\right)^{-1} \tag{1}$$

The resulting temperature sensitivity for the compact configuration ($B =$ 150m) according to [28] is given in table 1. The left panel shows the beam diameter and the sensitivity with integration time 60s. On the right panel we calculated the baseline needed to have $\Delta\phi \leq 2$" and the corresponding sensitivity with the same integration time, $60s$. In the last column of the right panel we calculated the corresponding integration time, which is needed to reduce the noise variance to 10μK. This was achieved by using $\sigma_D \propto \sqrt{B^2 \Delta t_{int}}$. This integration time is below 2 days for frequencies below 300GHz.

For the present study we will always approximate the sensitivity of the detector with $\sigma_D \lesssim 10\mu$K at a resolution of $\Delta\phi \approx 2$". This assumption is realistic for an observation with ALMA. For even larger observation times, the detector noise can be neglected compared to the statistical fluctuations of astrophysical origin.

The design of the ALMA system was not related to the aspect considered in this study. It was intended to detect and study the earliest and most distant galaxies, the epoch of the first light in the Universe. It will also look deep into the dust-obscured regions where stars are born to examine

ν [GHz]	$\Delta\phi$ [$arcsec$]	$(\sigma_D)_{60s}$ [mK]
35	11.79	0.1079
90	4.584	0.1871
140	2.947	0.2518
230	1.794	0.4317
345	1.196	1.007
409	1.009	1.799
650	0.6347	13.67
850	0.4853	24.46

ν [GHz]	B [m]	$(\sigma_D)_{60s}$ [mK]	$(\Delta t_{int})_{10\mu\mathrm{K}}$ [hr]
35	884	3.747	67
90	344	0.983	31
140	221	0.547	23
230	150	0.432	31
345	150	1.007	169
409	150	1.799	539
650	150	13.67	3.11×10^4
850	150	24.46	9.97×10^4

Table 1: Temperature sensitivities and integration times

the details of star and planet formation. In addition to these two main science drivers the array will make major contributions to virtually all fields of astronomical research.

4 Self-similar surface density profiles

Many high-resolution N-body simulations suggest that the dark matter density profile $\rho_{\rm dm}(r)$ is well described by a self-similar form: $\rho_{\rm dm}(r) = \rho_s y_{\rm dm}(r/r_s)$ where ρ_s is the mass density normalization factor, r_s is a length scale, and $y_{\rm dm}(x)$ is a non-dimensional function representing the profile. The density normalization, ρ_s, is determined to yield mass $M_{\rm vir}$ when $\rho_{\rm dm}(r)$ is integrated within the virial radius, $r_{\rm vir}$. The length scale r_s is defined as $r_s = r_{\rm vir}/c$, where c is the concentration parameter. We calculate $r_{\rm vir}(M_{\rm vir}, z)$ and c according to a spherical top hat model and an empirical fitting formula found in the literature by equations (16) and (12) below.

The dark-matter profile is approximated with the following analytic form

$$y_{\rm dm}(x) = \frac{1}{x^\alpha (1+x)^{3-\alpha}} \tag{2}$$

where the parameter α is assumed to be 1 or 3/2 (see [21] and [23]).

The gas density profile can be obtained using a polytropic model in hydrostatic equilibrium with the dark matter background. In this case, the gas density and temperature profiles assume a self-similar form,

$$\rho_{\rm gas}(r/r_s) = \rho_{\rm gas}(0) y_{\rm gas}(r/r_s) \tag{3}$$

$$T_{\rm gas}(r/r_s) = T_{\rm gas}(0) y_{\rm gas}^{\gamma-1}(r/r_s) \tag{4}$$

where $P_{\rm gas} \propto \rho_{\rm gas} T_{\rm gas} \propto \rho^\gamma$ has been used. The hydrostatic equilibrium equation

$$\rho_{\rm gas}^{-1} \frac{{\rm d}P_{\rm gas}}{{\rm d}r} = -G \frac{M(\leq r)}{r^2} \tag{5}$$

can be solved [24] for $y_{\rm gas}(x)$ for a fixed dark matter mass distribution.

$$y_{\rm gas}^{\gamma-1}(x) = 1 - 3\eta^{-1} \frac{\gamma-1}{\gamma} \frac{c}{m(c)} \int_0^x {\rm d}u \frac{m(u)}{u^2} \tag{6}$$

where η is an integration constant, and $m(u)$ is the dimensionless mass within a distance u from the center

$$m(x) \approx m_{\rm dm}(x) = \int_0^x {\rm d}u u^2 y_{\rm dm}(u) \tag{7}$$

These integrals can be evaluated analytically [24] for the particular cases $\alpha = 1$ and $\alpha = 3/2$.

The mass-temperature normalization is given by the virial theorem

$$\eta^{-1} = \frac{G\mu m_{\rm p} M_{\rm vir}}{3 r_{\rm vir} k_{\rm B} T_{\rm gas}(0)} \tag{8}$$

Both theoretical and numerical studies assert that the gas density profile traces the dark matter density profile in the outer region of the halo (see [20] and references therein). Therefore, the slopes of these two profiles are assumed to be the same for $c/2 < x < 2c$. x^* shall denote the x value where the slopes equal. This requirement fixes the η normalization in eq. (6) and the polytropic index γ (see [20] for details). The normalization is given by

$$\eta = \frac{-3}{\gamma s^*}\frac{c}{x^*}\frac{m(x^*)}{m(c)} + 3\frac{\gamma-1}{\gamma}\frac{c}{m(c)}\int_0^{x^*} \mathrm{d}u \frac{m(u)}{u^2} \tag{9}$$

$$s^* = -\left[\alpha + (3-\alpha)\frac{x^*}{1+x^*}\right] \tag{10}$$

A good approximation for γ is

$$\gamma = 1.15 + 0.01(c_{NFW} - 6.5) \tag{11}$$

where the parameter c_{NFW} is an α independent concentration parameter defined below.

The concentration parameters, c and c_{NFW} can be written in terms of the virial mass and redshift for a given cosmological model. An empirical fitting formula [29] is

$$c(M_{\rm vir}, \alpha, z) = \begin{cases} c_{NFW} \text{ if } \alpha = 1 \\ c_{NFW}/1.7 \text{ if } \alpha = 1.5 \end{cases} \tag{12}$$

$$c_{NFW} = \frac{6}{1+z}\left(\frac{M_{\rm vir}}{10^{14}h^{-1}M_\odot}\right)^{-1/5} \tag{13}$$

Specifically $c(10^{15}M_\odot, 1, 0) = 4.1$ and $c(10^{13}M_\odot, 1, 0) = 10.4$. Eq. (12) supplies a one-to-one correspondence between c and $M_{\rm vir}$, these constants carry the same information about the cluster. With this identification, in the remaining of this study, we shall use c and $M_{\rm vir}$ interchangeably.

The self-similar model defined above does not assign a value for the gas density normalization $\rho_{\rm gas}(0)$. It's value can be calculated by requiring that the ratio of the total dark matter and gas masses' ratio within $x_{\rm max}$ attain $\Omega_{\rm dm}/\Omega_{\rm B}$. Simulations without feedback from galaxy formation typically find values for the cluster gas mass fraction that are only slightly lower than the input global baryon fraction [32]. We shall adopt $(\Omega_{\rm dm}, \Omega_{\rm B}) = (0.253, 0.047)$ according to recent WMAP measurements. Thus

$$\rho_{\rm gas}(0) = \frac{m_{\rm dm}(x_{\rm max})}{m_{\rm gas}(x_{\rm max})}\frac{\Omega_{\rm B}}{\Omega_{\rm dm}} \tag{14}$$

where $m_{\rm gas}(x)$ is the accumulated dimensionless mass of the gas defined as

$$m_{\rm gas}(x) = \int_0^x {\rm d}u u^2 y_{\rm gas}(u) \tag{15}$$

and $m_{\rm dm}$ was defined similarly by (7). The value of $x_{\rm max}$ cannot be increased without bounds, since $m_{\rm gas}(x)$ and $m_{\rm dm}(x)$ are unbounded for large x with different distributions. Moreover, since $y_{\rm dm}(x)$ tends to zero and $y_{\rm gas}(x)$ approaches a finite nonzero value while $x \to \infty$, the dark matter-gas mass fraction $\lim_{x\to\infty} m_{\rm dm}(x)/m_{\rm gas}(x) = 0$ implying that $\rho_{\rm gas}(0) = 0$. Therefore $x_{\rm max} = \infty$ is inadequate for calculating the normalization. Nevertheless $\rho_{\rm gas}(0)$ has only a mild $x_{\rm max}$ dependence for reasonably low $x_{\rm max}$ values. Using $x_{\rm max} = 5c$, yields a maximum error of 10% for $x_{\rm max} \in [c, 10c]$ (see fig. 5 below).

We calculate the virial radius, $r_{\rm vir}$, with the spherical top hat collapse model using $(\Omega_m, \Omega_\Lambda, h) = (0.30, 0.70, 0.70)$ flat universe

$$r_{\rm vir} = \left[\frac{M_{\rm vir}}{(4\pi/3)\Delta_c(z)\rho_c(z)}\right]^{1/3} \tag{16}$$

$$\rho_c(z) = \frac{3H_0^2}{8\pi G}[\Omega_m(1+z)^3 + \Omega_\Lambda] \tag{17}$$

$$\Delta_c = 18\pi^2 + 82d - 39d^2 \tag{18}$$

$$d(z) = \frac{-\Omega_\Lambda}{\Omega_m(1+z)^3 + \Omega_\Lambda} \tag{19}$$

Instead of the concentration parameter, the observable quantity is the angular diameter of the cluster. The angular radius corresponding to the virial radius is

$$\Theta = \frac{r_{\rm vir}}{d_A(z)} = r_{\rm vir}\frac{(1+z)^2}{d_L(z)} \tag{20}$$

The luminosity distance is

$$d_L(z) = (1+z)\frac{c_0}{H_0}\int_{(1+z)^{-1}}^1 \frac{{\rm d}a}{a^2\sqrt{\Omega_m a^{-3} + \Omega_\Lambda}} \tag{21}$$

This can be expressed [30] with the universal fitting formula

$$d_L(z) = (1+z)\frac{c_0}{H_0}\left[f(1) - f((1+z)^{-1})\right] \tag{22}$$

$$f(a) = 2\sqrt{s^3+1}\left[\frac{1}{a^4} - 0.1540\frac{s}{a^3} + 0.4304\frac{s^2}{a^2} + 0.19097\frac{s^3}{a} + 0.066941 s^4\right]^{-1/8} \tag{23}$$

$$s^3 = \Omega_\Lambda/\Omega_m \tag{24}$$

With this choice, the only free parameters remaining in the gas density, $\rho_{\rm gas}(x)$, are $M_{\rm vir}$ and α.

Let us introduce a modification of the density profile (2) by accounting for the cutoff of the cluster on the edge. As it was pointed out above (14) the truncation is mandatory when calculating the gas density normalization. Let us assume that the density has a linear cutoff between radii $x_{\rm max} - D$ and $x_{\rm max} + D$. At $x_{\rm max} + D$ the density attains the background value. Since this value is negligible compared to $y_{\rm dm}(c)$ for realistic clusters, it will be approximated by 0. Since the gas density profile traces the dark matter profile, the linear density cutoff can be imposed on directly the gas density. In what follows, the original density function (6) will be referred to as $y_{\rm gas_0}(x)$, and $y_{\rm gas}(x)$ shall denote the density profile with a cutoff.

$$y_{\rm gas}(x) = y_{\rm gas_0}(x) W(x) \tag{25}$$

where

$$W(x) = \begin{cases} 1 & \text{if } x < x_{\rm max} - D \\ \frac{(x_{\rm max}+D)-x}{2D} & \text{if } x_{\rm max} + D < x < x_{\rm max} - D \\ 0 & \text{if } x_{\rm max} + D < x \end{cases} \tag{26}$$

For consistency, the $\rho_{\rm gas}(0)$ gas density normalization has to be recalculated from eq. (14) using the proper $m(x)$ (eq. (7) and (15)) accordingly. Practically $\rho_{\rm gas}(0)$ is altered by $x_{\rm max}$ only, the D dependance is negligible. Thus, we recalculate $\rho_{\rm gas}(0)$ with (14), and will not worry about the $D \neq 0$ value.

5 SZ surface brightness profiles

The SZ surface brightness profile is given by

$$I(x) = i_0 g(\theta) y_C(x) \tag{27}$$

where $\theta = \frac{h\nu}{k_{\rm B} T_{\rm CMB}}$ is the dimensionless frequency, $i_0 = 2(k_{\rm B}T)^3/(hc)^2$ is the intensity scale, and y_C is the Compton parameter.

$$g(\theta) = \frac{\theta^4 e^\theta}{(e^\theta - 1)^2} \left(\theta \frac{e^\theta + 1}{e^\theta - 1} - 4 \right) (1 + \delta_{SZE}(\theta, T_e)) \tag{28}$$

The $\delta_{SZE}(x, T_e)$ term is the relativistic correction to the frequency dependence. The Compton y_C-parameter is defined as

$$y_C(x) = \int n_e \frac{k_B T_e}{m_e c_0^2} \sigma_T {\rm d}l \tag{29}$$

where σ_T is the Thompson cross-section, n_e is the electron number density, T_e is the electron temperature, k_B is the Boltzmann constant, $m_e c_0^2$ is the electron rest mass energy, and the integration is along the line of sight (i.e. along $r = \sqrt{x^2 + l^2}$ for a given x). We calculate the electron temperature, T_e, with $T_e = T_{\rm gas}$ using eq. (8), and the number density is given by

$$n_e(x) = \frac{\rho_{\rm gas}(0) y_{\rm gas}(x)}{\mu m_{\rm p}} \tag{30}$$

where $m_{\rm p}$ is the proton rest mass, and $\mu = 0.59$ for an ionized H-He plasma with 25% Helium abundance by mass.

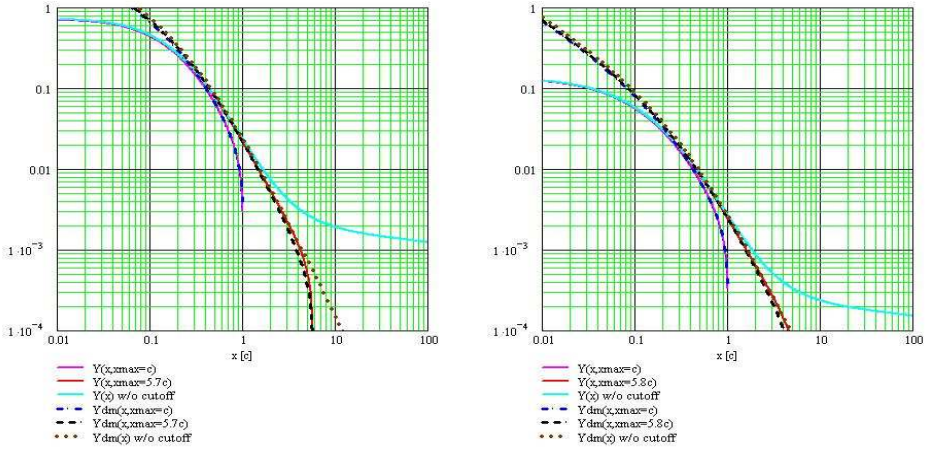

Figure 2: Predicted SZ surface brightness profiles for $\alpha = 1$ (*left*) and $\alpha = 1.5$ (*right*). The other parameters were $M_{\rm vir} = 10^{15} M_\odot$ and $z = 0.1$. The solid curves show $Y(x)$ profiles with various $x_{\rm max}$ choices. The dotted, dashed and dadotted curves were obtained by replacing $y_{\rm gas}(x)$ with $y_{\rm dm}(x) y_{\rm gas}(x^*)/y_{\rm dm}(x^*)$ in equation (35), with $x^* = c/2$ for $x_{\rm max} = c$ and $x^* = c$ for $x_{\rm max} > c$.

Substituting (4) and (30) in (29) the SZ surface brightness profile (27) is separable into a dimensionless integral with spacial dependence and a constant coefficient

$$I(x) = I_s Y(x) \tag{31}$$

where

$$Y(x) = \int_{-\infty}^{\infty} {\rm d}l \, [y_{\rm gas}(\sqrt{x^2 + l^2})]^\gamma \tag{32}$$

$$I_s = g(\theta) i_0 n_e(0) \frac{k_B T_{\rm gas}(0)}{m_e c_0^2} \sigma_T r_s \tag{33}$$

Substituting (8) and (30) gives

$$I_s = g(\theta) i_0 \sigma_T \frac{1}{m_e c_0^2} \frac{G M_{\rm vir}}{3c} \rho_{\rm gas}(0) \eta \tag{34}$$

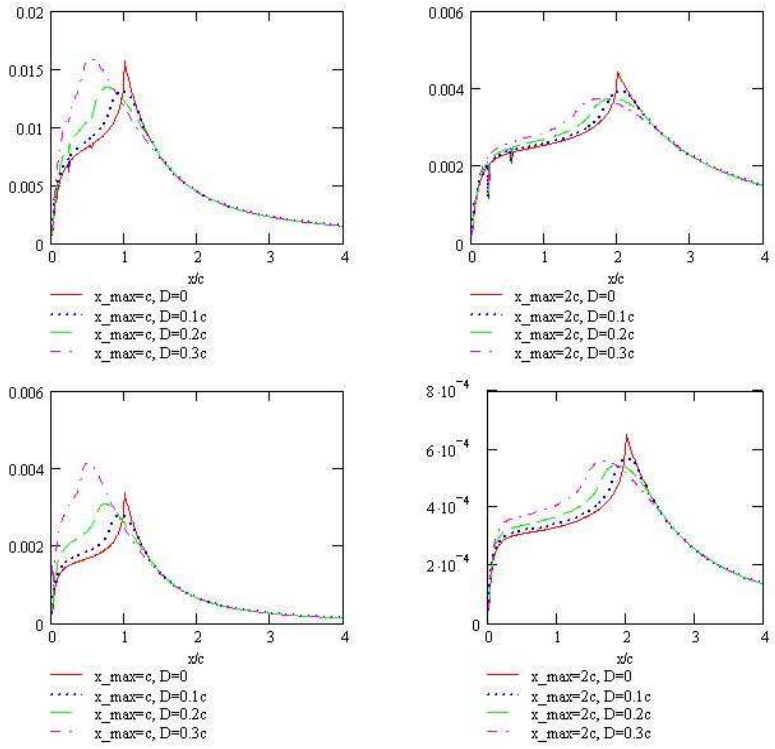

Figure 3: The difference between the normalized brightness profiles with and without a cutoff, for various D values. The difference is taken between $Y(x)/Y(0)$ profiles with different $x_{\rm max}$ and D parameters, but identical $M_{\rm vir}, \alpha, z$ values. $\alpha = 1.5$, $z = 0$, $x_{\rm max2} = 100$ was assumed on all plots. (The z dependence is implicit in the concentration parameter c.) The top row uses $M_{\rm vir} = 10^{15} M_\odot$, the bottom row has $M_{\rm vir} = 10^{13} M_\odot$. The left panels have $x_{\rm max1} = c$, the right panels have $x_{\rm max1} = 2c$. The SZ brightness profiles were normalized with $1/Y(0)$. This allows the profile with $x_{\rm max2} = 100$ to be nonzero. Unnormalized SZ profiles' differences are plotted on fig. 4.

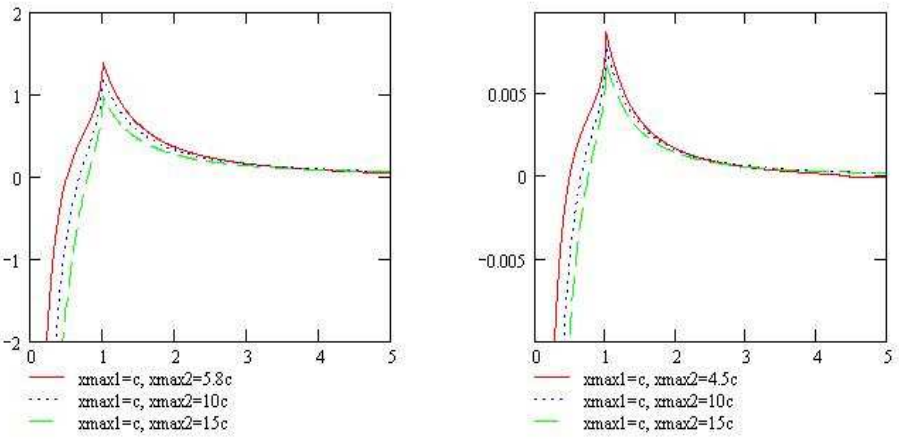

Figure 4: The difference between SZ brightness profiles with a cutoff at $x_{\max 1} = c$ and at various $x_{\max 2}$. The scale on the y-axis is 10μK, which is the typical noise power. The graphs are therefore related to the signal to noise ratio for deciding between $x_{\max 1}$ and $x_{\max 2}$. The left plot has $M_{\rm vir} = 10^{15} M_\odot$, and right has $M_{\rm vir} = 10^{13} M_\odot$. The α and z values are fixed at 1.5 and 0.1. Notice that it is irrelevant to choose $x_{\max 2}$ to approach infinity, since then the corresponding I_s would vanish.

The domain of the integral in $Y(x)$ can be truncated to the maximal radius where the density is nonzero.

$$Y(x) = 2 \int_0^{\sqrt{(x_{\max}+D)^2 - x^2}} {\rm d}l \, [y_{\rm gas}(\sqrt{x^2 + l^2})]^\gamma \tag{35}$$

Fig. 2, 3, and 4 plot the predicted SZ brightness profiles' spacial dependence with and without the cutoff. Fig. 2 depicts the $Y(x)$ profile for $\alpha = 1$ and $\alpha = 1.5$, and shows that a higher α leads to a smaller $Y(0)$ value. Unless a particular $x_{\max}$ cutoff is introduced, the $y_{\rm gas}(x)$ and $Y(x)$ profile has a nonzero limit for large x. Various choices of $x_{\max}$ can be compared by analyzing the difference between the associated intensity profiles (see below). Such differences between the $Y(x)$ and $I(x)$ profiles are plotted on figures 3 and 4. Note that the $I(x)$ profile is obtained from $Y(x)$ with multiplying by the I_s coefficient. Various $x_{\max}$ choices generally lead to unequal I_s values, implying a nonzero difference between the $I(x)$ profiles at $x = 0$.

The SZE intensity I(x) given by (31) is the deviation of the intensity spectrum from the black-body radiation.

$$I_{CMB} = \frac{2h}{c_0^2} \frac{\nu^3}{e^\theta - 1} \tag{36}$$

The SZE intensity can be expressed conveniently as a small temperature change of the CMB spectrum. Let us introduce the SZE temperature $T(x)$ increment by

$$I(x) = \frac{\partial I_{CMB}}{\partial T_{CMB}} T(x) \tag{37}$$

Solving for the SZE temperature

$$T(x) = f(\theta) y_C(x) T_{CMB} \tag{38}$$

where $f(\theta)$ is given by

$$f(\theta) = \left(\theta \frac{e^\theta + 1}{e^\theta - 1} - 4 \right) (1 + \delta_{SZE}(\theta, T_e)) \tag{39}$$

Therefore the SZE temperature profile assumes

$$T(x) = T_s Y(x) \tag{40}$$

where

$$T_s = f(\theta) T_{CMB} \sigma_T \frac{1}{m_e c_0^2} \frac{G M_{\rm vir}}{3c} \rho_{\rm gas}(0) \eta \tag{41}$$

$M_{\rm vir}/M_\odot$	α	100 Hz	150 Hz	200 Hz	250 Hz
10^{13}	1	-13	-8	-2	5
10^{13}	1.5	-19	-12	-3	7
10^{14}	1	-75	-47	-13	26
10^{14}	1.5	-119	-75	-20	40
10^{15}	1	-420	-260	-73	141
10^{15}	1.5	-688	-435	-120	233

Table 2: Calculated central SZE temperatures, $T(0)$, in μK for various $M_{\rm vir}$, α, and ν values. ($x_{\rm max} = 5c$)

Table 2. shows the calculated central $T(0)$ values for various clusters' SZE signatures for various detecting frequencies. Figure 5. shows the slight $x_{\rm max}$ dependance of the central SZE temperature. The $x_{\rm max}$ dependence is implicit in $\rho_{\rm gas}(0)$ (eq. (14)) and also in $Y(x)$ (eq. (35)). The former dominates the $x_{\rm max}$ dependence. Choosing $x_{\rm max} = 5c$ produces at most a 10% error in the central SZE temperature compared to the value with $x_{\rm max} \in [c, 10c]$.

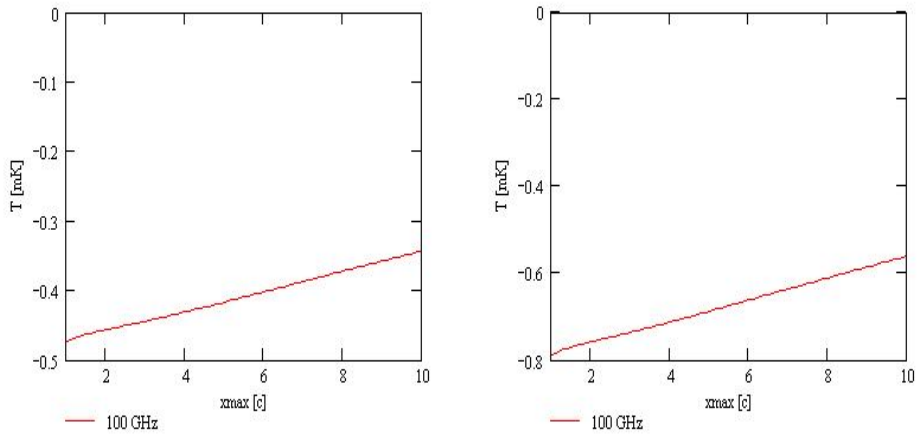

Figure 5: The central SZE temperature $T(0)$ when varying $x_{\max}$. Both figures have $M_{\rm vir} = 10^{15} M_{\odot}$, $z = 0$, $\nu = 100\,{\rm GHz}$, while $\alpha = 1$ left and $\alpha = 1.5$ right. Observe that choosing $x_{\max} \approx 5c$ for the normalization of $\rho_{\rm gas}(0)$ in eq. (14) is adequate with a maximum error of 10% for $x_{\max} \in [c, 10c]$.

Observe, that the $\alpha = 1.5$ dark matter profile yields $\propto 1.6\times$ larger $T(0)$ values than the $\alpha = 1$ model. We can elucidate the source of this increment by tracking the differences between the two models in equations (40) and (41). The difference is caused by the change in the $\rho_{\rm gas}(0)\eta Y(0)/c$ product. Since the dark matter model with $\alpha = 1.5$ is concentrated more in the central region, the gas density and the gas temperature is expected to be higher in the center for $\alpha = 1.5$. Indeed, $\rho_{\rm gas}(0) = 8.7 \cdot 10^4 \rho_{BG}$ for $\alpha = 1.5$, while it is $2.0 \cdot 10^4 \rho_{BG}$ for $\alpha = 1$. The η parameter is simply proportional to the $T_{\rm gas}(0)$ gas temperature (8), which is somewhat higher (22%) for $\alpha = 1.5$. Third, the $1/c$ factor further increases the difference by 70%. Finally, the increment caused by variables localized to the center is smeared by Y, which accounts for the fact that the observation measures the projection of the intensity along the line of sight. In particular, $Y(0) = 0.14$ for $\alpha = 1.5$ and 0.73 for $\alpha = 1$. Therefore, the resulting increase in the central SZE temperature is 60% for $\alpha = 1.5$.

A given detector measures the intensity at various θ frequencies and x radii. The SZE measurement is hampered by the secondary CMB anisotropies. These are the unresolved thermal SZ effect and the Ostriker-Vishniac (OV) effect. For the OV contribution we simply assume a flat band power contribution of $1\mu K$ [19]. The unresolved thermal SZE contamination is strongly non-Gaussian, and there is a significant amount of uncertainty in the small-scale power, due to unknown effects such as gas cooling and feedback from star formation. It is estimated [19] that the rms power on arcminute scales

are between 1 and $6\mu K$. For a conservative estimate we shall assumes

$$\sigma_N = 10\mu K \tag{42}$$

white Gaussian noise. Note that intensity can be converted to temperature with eq. (37).

6 Detecting cluster edges

The response of a detector with a given window function W(x) associated with its angular resolution for input $T(x)$ is

$$h(x) = \int T(x) W_D(x - x') \mathrm{d}^2 x' \tag{43}$$

The window function of the detector will be approximated by a Dirac-δ, so that the hypotheses are the temperature profiles with no distortion. The observation is taken over K distinct directions corresponding to the resolution of the measurement. The measurement frequency is assumed to be a fixed value.[15]

The model with a cutoff $h_1(x)$ is set uniquely in terms of the parameters $(M_{\mathrm{vir}}, z, \alpha, x_{\mathrm{max1}}, D)$. Since $h_2(x)$ is the hypothesis without a cutoff, it seems plausible to use the same definition with the distinction $x_{\mathrm{max2}} = \infty$. However, the density is unnormalizable in this case. For these reasons, instead of this choice, we shall adopt the following definitions of $h_2(x)$.

1. The central gas density of $h_2(x)$ is chosen to equal the central density of $h_1(x)$. Equivalently, the $T(0)$ temperatures are assumed to be the same for the two hypotheses.

2. The $h_2(x)$ model is chosen to equal the self similar model within a radius x_{max2} where the gas density drops to the background barion density value for the given redshift. This value is between $4.5 \leq x_{\mathrm{max2}} \leq 5.5$ for typical clusters with virial mass $10^{13} M_\odot \leq M_{\mathrm{vir}} \leq 10^{15} M_\odot$, at any redshift and α (either 1 or 1.5). The gas density normalization is calculated consistently with eq. (14). The profile outside x_{max2} is truncated with $D = 0$. Note, that there is a 5% variance in the central temperature with this definition (see fig. 5.). The central temperatures of h_1 and h_2 are different in this case, the deviation is around 10% for $x_{\mathrm{max1}} = c$.

These $h_2(x)$ definitions will be referred to as Model I and Model II, respectively. Although Model II has the advantage of being a closed and consistent theoretical description where $\rho_{\mathrm{gas}}(0)$ and thus T_s is derived from theory, Model I is favorable for practical reasons. The SZ temperature decrement scale T_s can be fitted without constraining the cluster barion mass fraction, which can vary among different observations. Similarly, the concentration parameter c can be set according to the observation, without strictly applying the $c(M_{\mathrm{vir}}, z)$ relation. For a comparison, the following statements

[15] In the numerical calculations we shall generally use 100 Hz.

describe Model I and II.

Model I

- $h_1(x)$ is described by parameters $(M_{\rm vir}, z, \alpha, T_s, c, x_{\rm max1}, D)$. The centeral gas density $\rho_{\rm gas}(0)$ is irrelevant for the comparison of the SZ profiles, but can be calculated from these parameters if needed.

- $h_2(x)$ is described by parameters $(M_{\rm vir}, z, \alpha, T_s, c)$. The integration bound $x_{\rm max2}$ in eq. (35) is taken to approach ∞. Practically, $x_{\rm max2} = 10c$ is an adequate choice.

- The corresponding parameters of $h_1(x)$ and $h_2(x)$ need not equal, they are fitted independently for a mock observation.

- For a given observation, the parameters $M_{\rm vir}$, c, and T_s are expected to approximately satisfy eq. (12) of $c(M_{\rm vir}, z)$, eq. (41) for T_s, and eq. (14) for $\rho_{\rm gas}(0)$. The deviation from these values are assumed to be caused by statistical rather than systematic errors. Our fiducial model will have input signals with these values.

- $M_{\rm vir}$ and z are not free parameters, but are determined from a comparison with X-ray and gravitational lensing and spectroscopy observations.

Model II

- $h_1(x)$ is described by parameters $(M_{\rm vir}, z, \alpha, x_{\rm max1}, D)$. The parameters $\rho_{\rm gas}(0)$, T_s, and c are derived from theory and are not allowed to deviate from those values.

- $h_2(x)$ is described by parameters $(M_{\rm vir}, z, \alpha)$. The cutoff distance $x_{\rm max2}$ is calculated to have $\rho_{\rm gas}(x_{\rm max2}) = \rho_{BG}$.

- $M_{\rm vir}$ and z are not free parameters, as in Model I.

The expected output for the SZ intensity input of a cluster with or without a cutoff, $h_1(x)$ or $h_2(x)$, can be calculated from eq. (31). Therefore the measured output is to be compared with these two signal hypotheses. The detector "sees" the cutoff if it can make a significant decision between the two choices.

6.1 Heuristic S/N calculation

We shall now derive a simple estimate for the signal to noise ratio for choosing between profiles with and without a cutoff. For this estimate, the distinction between Model I and II need not be utilized. The cutoff radius is taken $x_{\max 1} = c$ and $x_{\max 2} = \infty$ respectively. The effective signal to noise ratio is calculated by comparing the $\Delta h(x) = h_2(x) - h_1(x)$ power to the noise power in the regular two-dimensional angular space.

The best decision rule for using only a single direction is obtained when only the x radius corresponding to the maximum of $h_2(x) - h_1(x)$ is used.

$$\left(\frac{S}{N}\right)^2_{K=1} = \frac{\max_x (h_2(x) - h_1(x))^2}{\sigma_N^2} \tag{44}$$

$$= \left(\frac{T(0)}{\sigma_N}\right)^2 \frac{\max_x (Y_2(x) - Y_1(x))^2}{Y(0)^2} \tag{45}$$

$$= \left(\frac{T(0)}{\sigma_N}\right)^2 \frac{(Y_2(x_{\max}) - Y_1(x_{\max}))^2}{Y(0)^2} \tag{46}$$

A typical value for a massive cluster $M_{\rm vir} = 10^{15} M_\odot$, $\alpha = 1.5$ and $z = 0.3$ is

$$\frac{(Y_2(x_{\max}) - Y_1(x_{\max}))^2}{Y(0)^2} = 6.1 \cdot 10^{-4} \tag{47}$$

$$T(0)^2/\sigma_N^2 = 5900 \tag{48}$$

implying that $S/N \approx 1.9$ for the best pixel of the survey. These values increase when increasing z.

Now if $\Delta\phi$ denotes the smallest angle corresponding to the maximum resolution of the detection, the number of pixels the cluster covers is approximately

$$K = \left(\frac{\Theta}{c\Delta\phi}\right)^2 \tag{49}$$

Since $\Theta = 8.1'$ for $z = 0.3$ for a massive cluster, we expect $K = 1.8 \cdot 10^4$ pixels for $\Delta\phi = 2$" of ALMA.

For an estimate,

$$\frac{S}{N} \approx \sqrt{K} \left(\frac{S}{N}\right)_{K=1} \tag{50}$$

Explicitly,

$$\frac{S}{N} \approx \frac{\Theta}{c\Delta\phi} \frac{T(x_{\max 2})}{\sigma_N} \tag{51}$$

This approximation presumes two things: 1. the signal strength (i.e. the difference between the profiles with and without a cutoff) in the relevant pixels

can be estimated with the maximal signal strength, and 2. the number of such pixels is proportional to the area of the cluster within $r_s = 1$ perpendicular to the line of sight. The radius $r_s = 1$ (i.e. $r = r_{\rm vir}/c$) is special in the sense that the dark matter overdensity is $\Delta_c(z)$. It is to be emphasized that eq. (50) is only a rough estimate, since the deviation of the two profiles are really sensitive to the region near the edge. Nevertheless, we approximate the area of this ring around the edge with the area within $r_s = 1$.

For a massive cluster $M_{\rm vir} = 10^{15} M_\odot$ at $z = 0.3$ the detection of the edge at $x_{\rm max} = c$ with ALMA has a signal to noise ratio $S/N \approx 250$. If $\alpha = 1$ for the same $M_{\rm vir}$ and z values, the parameter c is $1.7\times$ larger, $T(c)$ is about the same as $\alpha = 1.5$ (while Θ and $\Delta\phi$ are naturally invariant), implying that signal to noise ratio decreases by 70%, which is still $S/N \approx 150$. Larger z values decrease Θ but increase $T(c)$ in a way that there is a minimum at $z = 0.28$. Smaller $M_{\rm vir}$ has a larger c but a smaller Θ and $T(c)$, so that S/N decreases with $M_{\rm vir}$. The edge is most visible for clusters that are closeby $z \ll 0.28$ or faraway $z \gg 0.28$, and that have a large $M_{\rm vir}$ and α.

To improve this estimate we should acknowledge that the typical one-pixel signal strength in eq. (50) is not the maximum given by (44), but the average of the various 1 pixel signal to noise ratio over the cluster. This leads to the following approximation

$$\left(\frac{S}{N}\right)^2_{K=1} = \left(\frac{T(0)}{\sigma_N}\right)^2 \frac{\int_0^x {\rm d}x'\, 2\pi x' (Y_2(x') - Y_1(x'))^2}{\pi x^2 Y(0)^2} \tag{52}$$

The average is now taken over a radius range $x' \in [0, x]$. Let us label the x dependent part by $H(x)$, so that $(S/N)^2_{K=1} = (T(0)/\sigma_N)^2 H(x)$. Obviously, the signal to noise ratio depends on the radius range where the surface brightness is measured. Figure 6 shows this dependence for a cluster with $\alpha = 1.5$ and $M_{\rm vir} = 10^{15} M_\odot$ (*left*) and $M_{\rm vir} = 10^{13} M_\odot$ (*right*). Substituting the x value of the maximum of $H(x)$, yields $\max_x H(x) = 3.0 \cdot 10^{-4}$ and $1.2 \cdot 10^{-5}$ respectively. Using $\alpha = 1$ similarly yields $8.2 \cdot 10^{-4}$ and $2.5 \cdot 10^{-5}$. Therefore, eq. (52) leads to approximately the same S/N ratio as the original eq. (44).

Interestingly, $H(x)$ does have a maximum, indicating that the consideration of the pixels beyond the maximum might not increase the precision of the fit, unless the increase in $\sqrt{K}$ compensates the decrement. In the following section we will prove that all pixels have to be considered for an optimal edge detection.

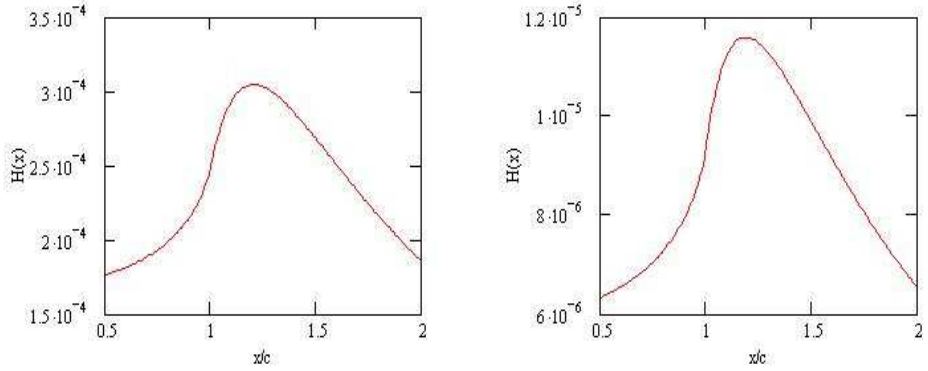

Figure 6: $H(x)$ is plotted with parameters $M_{\rm vir} = 10^{15} M_\odot$ (*left*), $10^{13} M_\odot$ (*right*), and $x_{\rm max2} = 100$, $x_{\rm max1} = c$, $z = 0.3$, $\alpha = 3/2$ for both cases.

These estimates all indicate, that the S/N ratio will be significant for regular clusters for an observation with ALMA.

6.2 S/N with the optimal filter

A measurement of the function $y(x)$ in x_k distinct directions yields a discrete sample y_k where $k \in [1, K]$.

$$y_k = \int \mathrm{d}^2 x \; y(x) W_D(x - x_k) \tag{53}$$

The set $\{y_k\}$ is an element of a K-dimensional vector space and will be denoted by $\boldsymbol{y}$. Similarly, let $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$ denote the discrete sample of the hypotheses functions $h_1(x)$ and $h_2(x)$. In particular, $\boldsymbol{y}$ denotes the measured SZE intensity profile, whereas $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$ denote the calculated brightness profiles with and without a cutoff (using eq. (31) with $x_{\rm max1} \neq \infty$ and $x_{\rm max2} = \infty$ respectively). The vectors $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$ depend on the parameters describing the cluster and the cutoff. We now derive the signal to noise ratio of the test for a given fixed choice of parameters.

If the real signal arriving to the detector was $s(x)$, then the detector measures the data

$$\boldsymbol{y} = \boldsymbol{s} + \boldsymbol{n} \tag{54}$$

where $\boldsymbol{n}$ is the probability variable corresponding to the noise.

Let us approximate the window function with a Dirac-δ, i.e. $W(x) = \delta_D(x)$, and let us assume white Gaussian noise of variance σ_N. In this case, the probability of detecting $\boldsymbol{y}$, given that the incoming signal is $\boldsymbol{s}$, is the

following

$$P(\boldsymbol{y}\,|\,\boldsymbol{s}) = \frac{1}{(2\pi)^{K/2}(\sigma_N)^K}\exp\left[-\frac{(\boldsymbol{y}-\boldsymbol{s})^2}{2\sigma_N^2}\right] \tag{55}$$

where the notion of the K-dimensional vectors' differences and scalar product was used.

$$(\boldsymbol{y}-\boldsymbol{s})^2 = \sum_{k=1}^{K}(y_k - s_k)^2 \tag{56}$$

The detection probability involves $P(\boldsymbol{s}\,|\,\boldsymbol{y})$, which can be obtained using the a priori probabilities $P(\boldsymbol{s})$ and the Bayes-theorem for conditional probabilities. Most common decision rules, such as the maximum posteriori, the Neyman-Pearson, or the minimax decisions involve constraints on the likelihood ratio.

$$L(\boldsymbol{y}) = \frac{P(\boldsymbol{y}\,|\,\boldsymbol{h_1})}{P(\boldsymbol{y}\,|\,\boldsymbol{h_2})} \tag{57}$$

Substituting (55) in (57) the log-likelihood becomes

$$\ln L(\boldsymbol{y}) = -\frac{1}{\sigma_N^2}\left(\boldsymbol{y} - \frac{\boldsymbol{h_1}+\boldsymbol{h_2}}{2}\right)\cdot(\boldsymbol{h_2}-\boldsymbol{h_1}) \tag{58}$$

$$= -\frac{1}{\sigma_N^2}(\boldsymbol{h_2}-\boldsymbol{h_1})\cdot(\boldsymbol{y}-\boldsymbol{h_2}) - \frac{1}{2\sigma_N^2}(\boldsymbol{h_2}-\boldsymbol{h_1})^2 \tag{59}$$

The likelihood depends on the $\boldsymbol{y}-\boldsymbol{h_2}$ component along $\boldsymbol{h_2}-\boldsymbol{h_1}$. The term $\boldsymbol{h_2}-\boldsymbol{h_1}$ is referred to as the matched filter.

The decision rule can be outlined as follows:

1. Obtain the hypothesis vectors $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$ from $h_1(x)$, $h_2(x)$ with $x = x_k$ defining the particular directions of the measurement. The parameters implicit in the functions $h_1(x)$ and $h_2(x)$ are given prior to the measurement.

2. Set up the desired decision margin for the likelihood ratio. For example, a natural choice is to choose the margin at $L(\boldsymbol{y}) = 0$.

3. Obtain $\boldsymbol{y}$ from the measurement.

4. Evaluate $L(\boldsymbol{y})$ from eq. (59), and choose the decision in favor of h_1 if this value is greater than the bound given in 2. above.

Since the noise distribution $P(\boldsymbol{n})$ is spherically symmetric in the K-dimensional vector space, it has the same power σ_N^2 along any basis. This

infers that the noise power for the likelihood detection rule is the noise power for a single bin, i.e.

$$N^2 = \sigma_N^2 \tag{60}$$

Thus increasing the sample size increases the signal power, but leaves the relevant noise contribution the same.

The signal power is

$$S^2 = (\boldsymbol{h_2} - \boldsymbol{h_1})^2 = \sum_k (h_{2,k} - h_{1,k})^2 = \frac{1}{\Delta x^2} \int \mathrm{d}^2x \; \Delta h(x)^2 \tag{61}$$

where $\Delta h(x) = h_2(x) - h_1(x)$ and Δx^2 denotes the area enclosed by the neighboring x_k points. We have assumed that the resolution of the measurement is fine enough to approximate the sum with the integral[16,17].

The signal to noise ratio from eq. (60) and (61)

$$\frac{S^2}{N^2} = \frac{(\boldsymbol{h_2} - \boldsymbol{h_1})^2}{\sigma_N^2} \tag{62}$$

For a given detector with angular resolution $\Delta\phi$ and a cluster of apparent angular virial radius Θ the Δx uncertainty is $c\Delta\phi/\Theta$. Substituting eq. (61) in (62) and taking the square root yields

$$\frac{S}{N} = \frac{\Theta}{\Delta\phi c \sigma_N} \sqrt{\int \mathrm{d}x^2 [h_2(x) - h_1(x)]^2} \tag{63}$$

We evaluate eq. (63) for Model I and Model II below.

6.2.1 Model I

Here we evaluate an approximation to the signal to noise ratio for Model I. The signal to noise ratio is calculable with eq. (63) if the parameters $(M_{\mathrm{vir}}, z, \alpha, T_s, c, x_{\max}, D)$ are given for $h_1(x)$ and $(M_{\mathrm{vir}}, z, \alpha, T_s, c)$ for $h_2(x)$. We choose realistic parameters, which are substituted for $h_1(x)$ to generate the fiducial model. For this section, the parameters used for both hypotheses are assumed to equal the correct values. This is only an approximation, nevertheless we anticipate the difference in T_s and c of $h_2(x)$ and $h_1(x)$ will

[16] **The Δh peak will be rounded off for finite resolutions (fig. ??) leading to a decreased signal power value!**

[17] The integral domain is approximated by a plane, since typically $\Theta \ll 1^\circ$.

be sufficiently small for this assumption to be reasonable. The general case with unequal parameters for h_1 and h_2 will be discussed in section 6.3.

$$\frac{S}{N} = \frac{\Theta(M_{\rm vir}, z)}{c(M_{\rm vir}, z)\Delta\phi}\frac{T_s}{\sigma_N}\sqrt{\int \mathrm{d}x^2[Y_2(x) - Y_1(x)]^2} \tag{64}$$

$M_{\rm vir}$ $[M_\odot]$	z	α	Θ $[']$	c	T_s $[-\mu\mathrm{K}]$	$T(0)$ $[-\mu\mathrm{K}]$	$x_{\rm max2}$ $[c]$	$x_{\rm max}$ $[c]$	D $[c]$	S/N I	S/N II
10^{13}	0.3	1	1.8	7.9	18	22	4.3	1	0.01	0.53	0.71
10^{13}	0.3	1.5	1.7	4.6	16	27	4.4	1	0.01	0.56	0.77
10^{14}	0.3	1	3.8	5.0	97	130	4.9	1	0.01	16	17
10^{14}	0.3	1.5	3.8	2.9	1100	160	5.0	1	0.01	16	19
10^{15}	0.3	1	8.1	3.1	560	780	5.5	1	0.01	470	440
10^{15}	0.3	1.5	8.1	1.8	7000	910	5.6	1	0.01	480	470

Table 3: Typical cluster parameters and S/N ratios for detecting the edge

We evaluated eq. (64) for various $(M_{\rm vir}, z, \alpha, T_s, c, x_{\rm max}, D)$. The results are listed in table 3 and plotted on fig. 7. The concentration parameter c and the angular radius Θ were chosen in accord with eq. (12) and (20) in a particular cosmology. Notice, that the signal to noise ratio obtained here equals the result of the heuristic calculation eq. (50) with the average one-pixel signal to noise eq. (52). Now, the result was obtained from eq. (62), which shows clearly that all pixels of the survey has to be considered, even though the low power pixels reduce the one-pixel average.

Table 3 gives a list of the signal to noise ratios for the typical cases. The variation caused by changing $x_{\rm max}$, D, and z are plotted on fig. 7. The main feature is that the S/N quickly increases with $M_{\rm vir}$ but is practically invariant for different α. The figures show that S/N decreases with $x_{\rm max}$, it is nearly constant when changing D, and it has a minimum in terms of z. The last two observations are slightly counterintuitive: cluster edges are equally visible for smeared edges, and cluster edges are more visible for the farthest clusters even though they are smaller.

The former can be explained with recalling the definition of the cutoff. Increasing the D parameter does not change the total cluster SZ intensity. If the cutoff shape is known exactly prior to the observation, as it is assumed in this section, increasing D (while fixing T_s) does not come closer on average to the profile without a cutoff.

To explain the second peculiarity, the S/N increase with z, we plotted the z dependence of the relevant factors on fig. 8. Although the angular

radius Θ decreases *(bottom right)*, the central temperature $T(0)$ *(top right)* and the concentration parameter *(bottom left)* reciprocal $1/c$ increases. Note that eq. (64) uses the temperature scale $T_s = T(0)/Y(0)$ *(top left)* instead of $T(0)$, but this also increases with z. The increasing terms dominate for large redshifts. In other words, the faraway clusters are more intense and have a higher contrast, which makes the edge visible even though the cluster appears smaller. For even larger redshifts, the signal to noise ratio is expected to have a cutoff, since then the observation window function can no longer be approximated by a Dirac-δ.

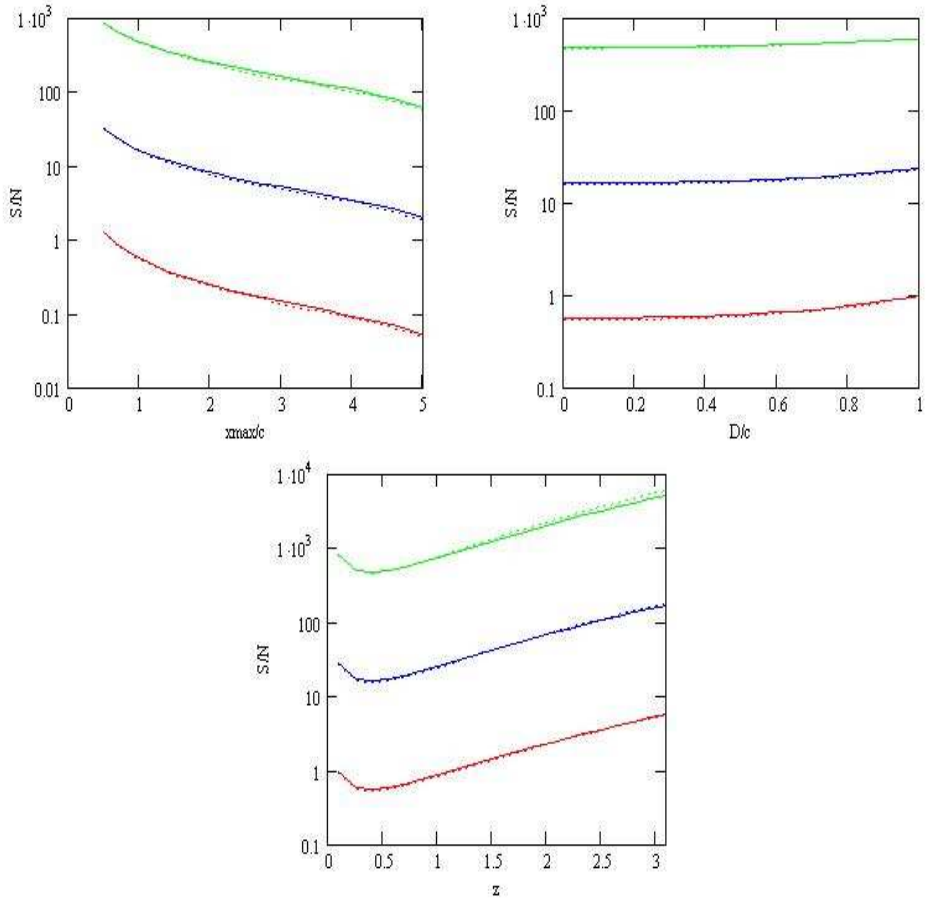

Figure 7: The signal to noise ratio for various cluster masses as a function of $x_{\max}$ *(top left)*, D *(top right)*, and z *(bottom center)*. Each curve has only one parameter changing, the other fixed parameters correspond to those listed in the first six lines of table 3. The solid lines trace $\alpha = 1.5$, the dotted lines show $\alpha = 1$. The signal to noise ratio is mostly independent of α and D, it increases with $M_{\rm vir}$ and z and decreases with $x_{\max}$.

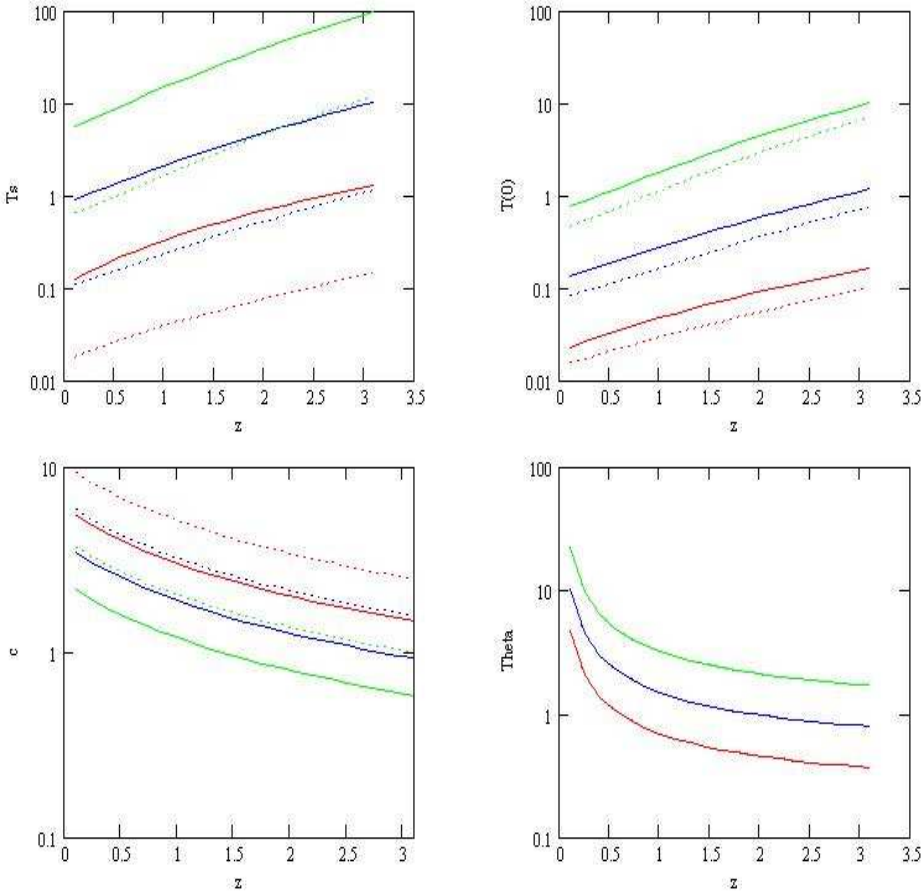

Figure 8: The parameters T_s/mK *(top left)*, $T(0)$/mK *(top right)*, c *(bottom left)*, and Θ *(bottom right)* for various cluster masses and α as a function of redshift. Each curve has only one parameter changing, the other fixed parameters correspond to those listed in the first six lines of table 3. The solid lines trace $\alpha = 1.5$, the dotted lines show $\alpha = 1$. The red curves have $10^{13} M_\odot$, the blue curves have $10^{14} M_\odot$, and the green curves has $10^{15} M_\odot$ virial mass. The Θ angular radii are independent of α, so the dotted and solid curves completely overlap.

This cutoff is not depicted on the graphs, it is anticipated to dominate when $\Theta \approx \Delta\phi$. For the ΛCDM cosmological model the angular radii of clusters does not decrease to zero, it approaches finite values: $\lim_{z\to\infty} \Theta(z) = 10.4$", 22.4", or 48.4" for clusters with low, medium, or large mass. Therefore, the $\Delta\phi = 2$" ALMA resolution is fine enough to see cluster edges at arbitrary redshifts.

In conclusion, medium to large mass clusters' edges should be clearly

observable with ALMA, as long as the edge is located near the virial radius. If the density cutoff is farther in the outer region, the edge will only be observed for the large mass clusters with $M > 10^{14} M_{\rm vir}$. The detection will be independent of α, z, and D.

6.2.2 Model II

For Model II, the T_s temperature is obtained from $x_{\rm max}$, which is unequal for the two hypotheses. From eq. (63) we get

$$\frac{S}{N} = \frac{\Theta(M_{\rm vir}, z)}{c(M_{\rm vir}, z)\Delta\phi} \sqrt{\int {\rm d}x^2 \left[\frac{T_{s_2} Y_2(x) - T_{s1} Y_1(x)}{\sigma_N}\right]^2} \tag{65}$$

Eqs. (64) and (65) corresponding to Model I and II can be compared with figs. 3 and 4, which plot the integrand of the two cases.

Table 3. compares the signal to noise ratio results for Model I and II. The T_{s2} value for $x_{\rm max2}$ was recalculated to substitute in eq. (65). The two models lead to approximately equal S/N detection ratios. This is not surprising since the parameter sets adopted in table 3, correspond to the fiducial model in accord with the relationships $c(M_{\rm vir}, z, \alpha)$ and $\Theta(M_{\rm vir}, z)$ (eqs. (12) and (20)). Model I and II will only be different, if there is a significant variance in the parameters around their calculated values.

Assuming that the noise is uncorrelated and Gaussian for different frequencies, measuring at many different frequencies the signal to noise ratio can be substantially higher.

6.3 The random distribution of S/N

One difficulty with the decision rule and signal to noise ratio discussed above is that the parameters are unknown prior to the measurement. Assume that the original signal (i.e. without noise) is $\boldsymbol{h_1}$ with parameter p. The measurement

$$\boldsymbol{y} = \boldsymbol{h_1}(p) + \boldsymbol{n} \tag{66}$$

can be used to give an estimate of p. Denote the estimated parameter[18] of $\boldsymbol{h_1}$ by p_1 and the estimated parameter of $\boldsymbol{h_2}$ by p_2. Once the parameters have been obtained, $\boldsymbol{h_1}(p_1)$ and $\boldsymbol{h_2}(p_2)$ can be fixed at the corresponding values.

[18] The least squares fit method for obtaining the parameters is relevant if the apriori distribution in the parameter space is assumed to be uniform.

Then the decision rule for fixed parameters can be used.

$$\ln L(\boldsymbol{y}) = -\frac{1}{\sigma_N^2}\left(\boldsymbol{y} - \frac{\boldsymbol{h_1}(p_1) + \boldsymbol{h_2}(p_2)}{2}\right) \cdot (\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1)) \tag{67}$$

If the apriori probabilities of the parameters are uniform, then the likelihood margin can be chosen independent of $\boldsymbol{y}$. If additionally, the apriori probabilities of the $\boldsymbol{h_1}(p_1)$ and $\boldsymbol{h_2}(p_2)$ are equal, then the likelihood margin can be chosen at $L(\boldsymbol{y}) = 0$. [19]

Since p_1 is obtained by minimizing $(\boldsymbol{y}-\boldsymbol{h_1}(p_1))^2$, the distribution of $\boldsymbol{h_1}(p_1)$ from (66) is

$$\boldsymbol{h_1}(p_1) = \boldsymbol{h_1}(p) + \boldsymbol{\Delta_1} \tag{68}$$

where p is the real parameter of the signal without the noise and $\boldsymbol{\Delta_1}$ is approximately a Gaussian random variable on the manifold $H_1 = \{\boldsymbol{h_1}(p_1) \,|\, p_1 \text{ arbitrary}\}$. $\boldsymbol{\Delta_1}$ has variance $\sqrt{\dim H_1}\sigma_N$.

Similarly, p_2 is obtained by minimizing $(\boldsymbol{y} - \boldsymbol{h_2}(p_2))^2$. The distribution of $\boldsymbol{h_2}(p_2)$ is

$$\boldsymbol{h_2}(p_2) = \boldsymbol{h_2}(p') + \boldsymbol{\Delta_2} \tag{69}$$

where p' is the parameter value at which $(\boldsymbol{h_1}(p) - \boldsymbol{h_2}(p'))^2$ is minimal. $\boldsymbol{\Delta_2}$ is defined by eq. (69). It can be shown that it is approximately a Gaussian random variable[20] on the manifold $H_2 = \{\boldsymbol{h_2}(p_2) \,|\, p_2 \text{ arbitrary}\}$, with variance $\sqrt{\dim H_2}\sigma_N$. Note that $\boldsymbol{\Delta_1}$ and $\boldsymbol{\Delta_2}$ is strongly correlated since both values are derived from the measurement $\boldsymbol{y}$.

The signal power

$$S^2 = (\boldsymbol{h_2} - \boldsymbol{h_1})^2 \tag{70}$$

is therefore a random variable. Let us define the empirical, most probable[21] and expected signal powers by

$$S^2_{emp} = (\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1))^2 \tag{71}$$

$$S^2_0 = (\boldsymbol{h_2}(p') - \boldsymbol{h_1}(p))^2 = \min_{p_2}(\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p))^2 \tag{72}$$

$$S^2_{exp} = \langle S^2_{emp} \rangle \tag{73}$$

[19] An equivalent realization of the decision process from a different perspective is the following. Take the sets $H_1 = \{\boldsymbol{h_1}(p_1) \in R^K \,|\, p_1 \text{ arbitrary}\}$ and $H_2 = \{\boldsymbol{h_2}(p_2) \in R^K \,|\, p_2 \text{ arbitrary}\}$ which are manifolds in the K-dimensional manifolds. Take the set $M = \{\boldsymbol{z} \in R^K \,|\, L(\boldsymbol{y}, p_1, p_2) = \kappa; \text{ with } p_1 \text{ and } p_2 \text{ arbitrary}\}$. M is a $K-1$=dimensional manifold. This separates the K-dimensional vector space into two disjoint regions R_1 and R_2. Now iff the measurement $\boldsymbol{y}$ is in R_1, the decision is made in favor of $\boldsymbol{h_1}$.

[20] The approximation is valid if H_2 is nearly a linear subspace in R^K, i.e. its curvature is much less than the uncertainty $\sqrt{\dim H_2}\sigma_N$.

[21] The estimated parameters have a Gaussian distribution around p and p'. The term "most probable" refers to the parameter distribution not the signal power distribution.

The expected signal power can be written in terms of S_0 using eq. (68) and (69)

$$S_{exp}^2 = \langle(\boldsymbol{h_2}(p) - \boldsymbol{h_1}(p') + \boldsymbol{\Delta_2} - \boldsymbol{\Delta_1})^2\rangle = S_0^2 + \langle(\boldsymbol{\Delta_2} - \boldsymbol{\Delta_1})^2\rangle \tag{74}$$

$$= S_0^2 + \langle\boldsymbol{\Delta_2^2}\rangle + \langle\boldsymbol{\Delta_1^2}\rangle - 2\langle\boldsymbol{\Delta_1}\boldsymbol{\Delta_2}\rangle \tag{75}$$

$$= S_0^2 + (\dim H_1 + \dim H_2)\sigma_N^2 - 2\langle\boldsymbol{\Delta_1}\boldsymbol{\Delta_2}\rangle \tag{76}$$

$$= S_0^2 + 2\sigma_N^2 - 2cos(\phi)^2\sigma_N^2 \tag{77}$$

where in the last line we have assumed that H_1 and H_2 are one dimensional and the scalar product of the corresponding normal basis vectors is $cos(\phi)$. This term can be obtained, for given $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$ hypothesis functions by

$$cos(\phi)^2 = \left(\frac{\mathrm{d}\boldsymbol{h_1}(p)}{\mathrm{d}p} \cdot \frac{\mathrm{d}\boldsymbol{h_2}(p')}{\mathrm{d}p'}\right)^2 / \left(\frac{\mathrm{d}\boldsymbol{h_1}(p)}{\mathrm{d}p}\right)^2 \left(\frac{\mathrm{d}\boldsymbol{h_2}(p')}{\mathrm{d}p'}\right)^2 \tag{78}$$

In general for any $(\boldsymbol{h_1}, \boldsymbol{h_2})$ we have

$$S_0^2 \leq S_{exp}^2 \leq S_0^2 + (\dim H_1 + \dim H_2)\sigma_N^2 \tag{79}$$

The noise power for choosing between $\boldsymbol{h_1}$ and $\boldsymbol{h_2}$ (with an arbitrary p_1 or p_2) is the variance of the measurement along $\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1)$, since the likelihood ratio (67) depends on only this component. Therefore

$$N^2 = \langle\left[\frac{\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1)}{||\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1)||} \cdot (\boldsymbol{y} - \boldsymbol{h_1}(p_1))\right]^2\rangle \tag{80}$$

$$= \sigma_N^2 + \langle\left[\frac{\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1)}{||\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p_1)||} \cdot \boldsymbol{\Delta_1}\right]^2\rangle \tag{81}$$

The first term is the statistical error, whereas the second term describes the systematic error resulting from the uncertain estimation of the parameters. For large signal to noise ratios this can be approximated to lowest order in Δ_1.

$$N^2 = \sigma_N^2 + \langle\left[\frac{\boldsymbol{h_2}(p') - \boldsymbol{h_1}(p)}{||\boldsymbol{h_2}(p') - \boldsymbol{h_1}(p)||} \cdot \boldsymbol{\Delta_1}\right]^2\rangle \tag{82}$$

$$= \sigma_N^2 + \langle(\cos\theta)^2\rangle\sigma_N^2 \tag{83}$$

where θ is the angle between the vector $\boldsymbol{h_2}(p') - \boldsymbol{h_1}(p)$ and $\boldsymbol{h_1}(p_1) - \boldsymbol{h_1}(p)$. In general for any $(\boldsymbol{h_1}, \boldsymbol{h_2})$

$$\sigma_N^2 \leq N^2 \leq 2\sigma_N^2 \tag{84}$$

In particular for the SZE brightness hypotheses, $H_2 \subset H_1$, with H_1 having two additional parameters. In this case, it can be shown that $\langle(\cos\theta)^2\rangle \approx 1/3$ independent of p.[22] Therefore

$$N^2 = \frac{4}{3}\sigma_N^2 \tag{85}$$

Although the expected signal to noise ratio can be calculated explicitly for a given parameter choice using eq. (77) and (85), it is useful to define its the theoretical bounds independent of the given form of the hypotheses. Comparing eq. (79) and (84), we get

$$\frac{S_0^2}{2\sigma_N^2} \leq \frac{S_{exp}^2}{N^2} \leq \frac{S_0^2}{\sigma_N^2} + \dim H_1 + \dim H_2 \tag{86}$$

where S_0 is given by eq. (72). The high bound is approached when H_2 tends to be parallel to H_1, the low bound is approached when H_2 is orthogonal to H_1 near $H_1(p)$ and $H_2(p')$.

For the most general case considered in this paper, $\boldsymbol{h_1}$ depends on parameters $(M_{\rm vir}, z, \alpha, T_s, c, x_{\rm max}, D)$ and $\boldsymbol{h_2}$ depends on $(M_{\rm vir}, z, \alpha, T_s, c)$. Since $M_{\rm vir}$, z, and α is assumed to be fixed, only 4 and 2 parameters are free in the two cases, respectively. Therefore the additive factors of eq. (86) are simply $\dim H_1 + \dim H_2 = 6$. Using (84) we get

$$\frac{3}{4}\frac{S_0^2}{\sigma_N^2} \leq \frac{S_{exp}^2}{N^2} \leq \frac{3}{4}\frac{S_0^2}{\sigma_N^2} + 4.5 \tag{87}$$

For a conservative estimate we shall use the low bound. Substituting S_0 from its definition, eq. (72)

$$\frac{S_{exp}^2}{N^2} = \frac{3}{4\sigma_N^2}\min_{p_2}[\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p)]^2 \tag{88}$$

This should be compared with eq. (62) describing the signal to noise ratio for fixed parameter values. The difference is the 3/4 factor and the $\min_{p_2}$ function over the p_2 parameter space. If the parameters describing the cluster are $p = (M_{\rm vir}, z, \alpha, T_s, c, x_{\rm max}, D)$ then the best fitting $\boldsymbol{h_2}$ hypothesis will have parameters $p' = (M_{\rm vir}, z, \alpha', T_s', c')$ generally different from the true values.

The result (88) can be written in terms of the original continuous functions $h_1(x)$ and $h_2(x)$. For a given detector with angular resolution $\Delta\phi$ and

[22] The reason this is not an exact equality stems from the fact that H_2 is not a complete linear subspace. It is curved and is constrained ($0 < D < x_{\rm max}$).

a cluster of apparent virial radius Θ the smallest detectable length scale is $\Delta x = c\Delta\phi/\Theta$. Converting the sum to an integral,

$$\frac{S_{exp}^2}{N^2} = \frac{3}{4\Delta x^2 \sigma_N^2} \min_{p_2} \int \mathrm{d}x^2 [h_2(x, p_2) - h_1(x, p)]^2 \tag{89}$$

$$= \frac{3\Theta^2}{4\Delta\phi^2 c^2 \sigma_N^2} \min_{p_2} \int \mathrm{d}x^2 [h_2(x, p_2) - h_1(x, p)]^2 \tag{90}$$

Taking the square root of the equation

$$\frac{S_{exp}}{N} = \frac{\sqrt{3}}{2} \frac{\Theta}{\Delta\phi c \sigma_N} \min_{p_2} \sqrt{\int \mathrm{d}x^2 [h_2(x, p_2) - h_1(x, p)]^2} \tag{91}$$

Eq. (91) measures the probability that the edge of the cluster is observable. Again, the angular radius Θ and the concentration parameter c can be obtained theoretically for given M_{vir} and z in a particular cosmology by eq. (12) and (20). Eq. (91) can be calculated for a given parameter set, substituting eq. (31) brightness profiles for $h_1(x, p)$ and $h_2(x, p_2)$.[23]

[23] Eq. (31) describes the profile without an edge, (i.e. $h_2(x, p_2)$) when taking $x_{\max} \gg c$ and $D = 0$.

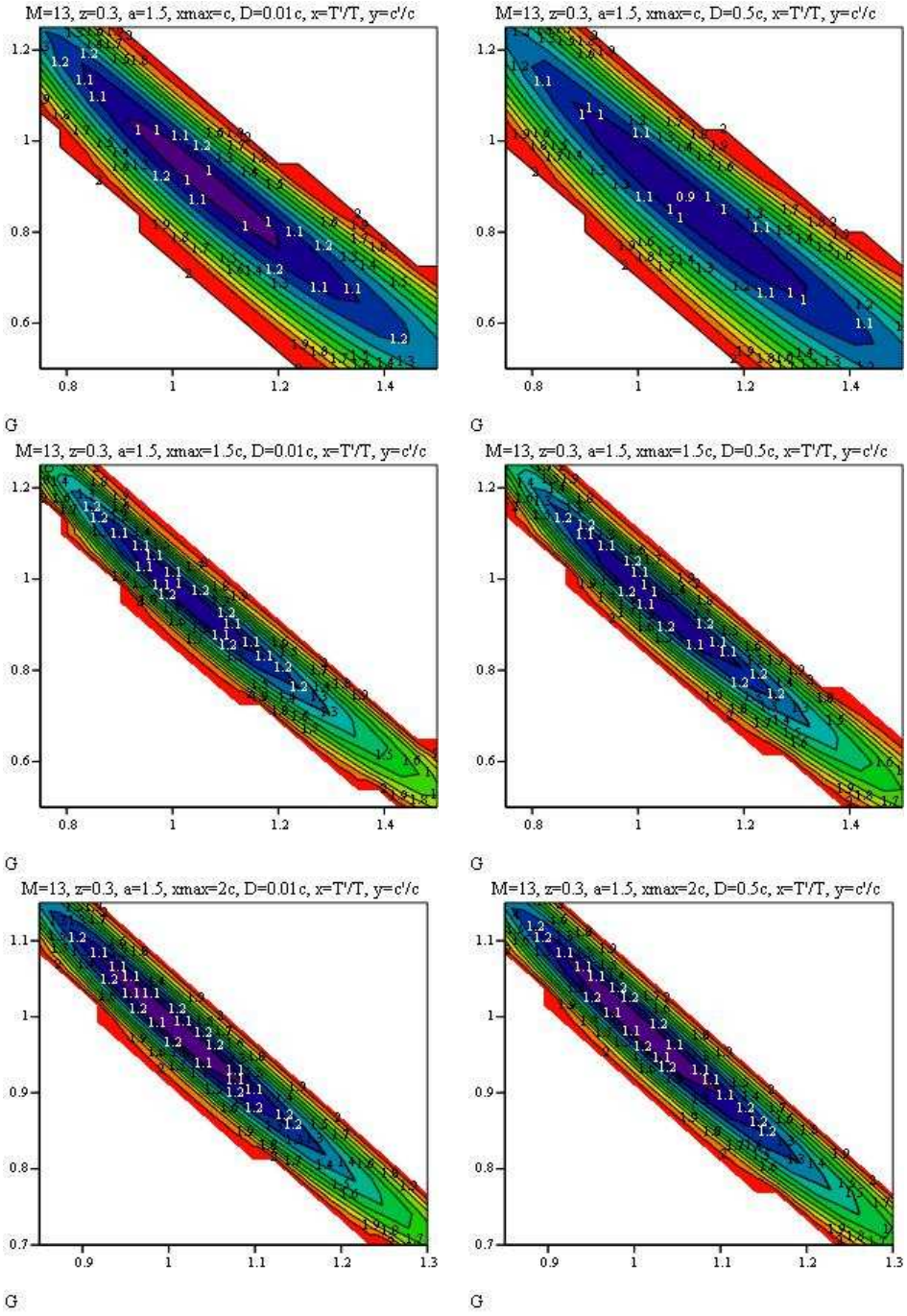

Figure 9: The $p_2 = (T_{s2}, c_2)$ parameter dependence of the S/N ratio for detecting the edge for $(M, z, \alpha) = (10^{13} M_\odot, 0.3, 1.5)$ and 6 choices for $(x_{\max}, D)$.

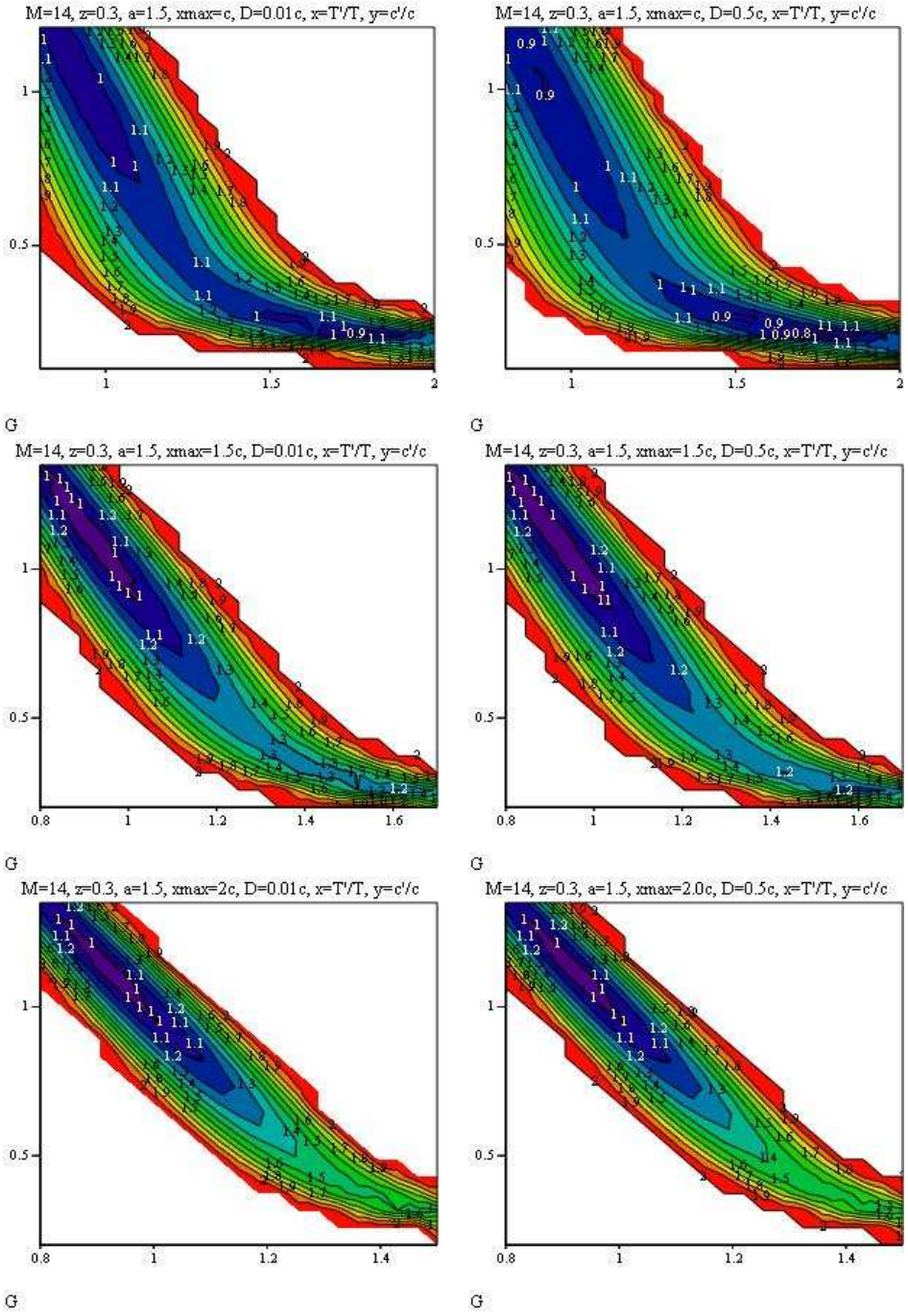

Figure 10: The $p_2 = (T_{s2}, c_2)$ parameter dependence of the S/N ratio for detecting the edge for $(M, z, \alpha) = (10^{14} M_\odot, 0.3, 1.5)$ and 6 choices for $(x_{\max}, D)$. Notice the multiple minima on the top panels.

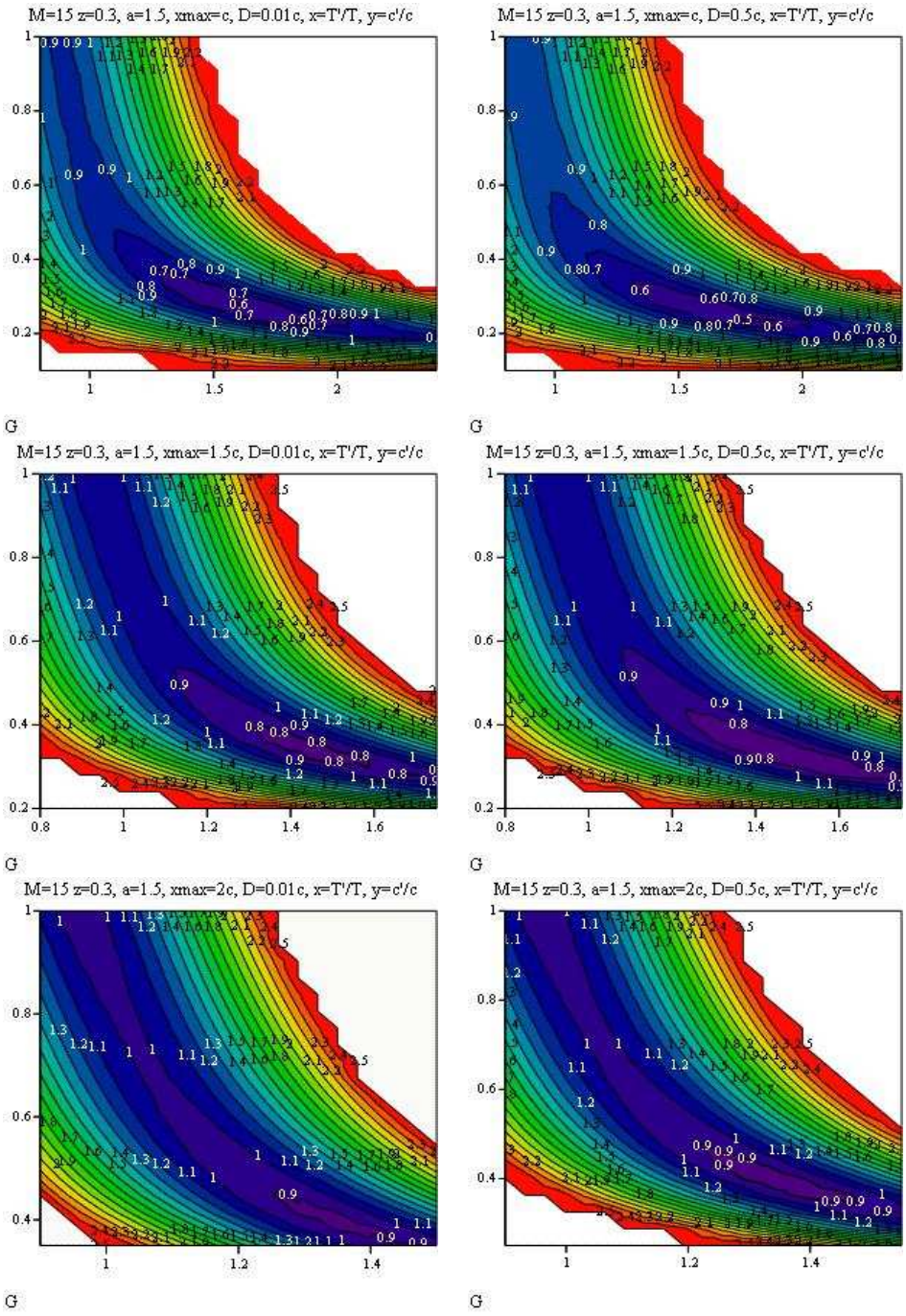

Figure 11: The $p_2 = (T_{s2}, c_2)$ parameter dependence of the S/N ratio for detecting the edge for $(M, z, \alpha) = (10^{15} M_\odot, 0.3, 1.5)$ and 6 choices for $(x_{\max}, D)$.

Figs. 9, 10, and 11 show the fractional deviation of the empirical signal

to noise ratio as compared to the fiducial value in Model I for various p_2 parameters. The best fitting p_2 parameter is at the minimum points. The $\sqrt{3}/2$ factor decrease because of the increase in the average noise power is not plotted for clearity. Thus, the plots assume the value 1 at the fiducial parameters. The numerical result presented in the last section on table 3 and fig. 7 considered $\boldsymbol{h_2}$ parameter values that were equal to the corresponding $\boldsymbol{h_1}$ parameter values. Figs. 9, 10, and 11 use the fiducial model with $p = (M_{\rm vir}, z, \alpha, c, T_s, x_{\rm max}, D$ with $(z = 0.3, \alpha = 1.5)$ and 18 different choices for $(M_{\rm vir}, x_{\rm max}, D)$. The (c, T_s) fiducial values were calculated with eqs. (12) and (41). The p_2 parameters had the same $(M_{\rm vir}, z, \alpha)$ but (T_s, c) were varied relative to the fiducial values. The signal to noise results listed in the previous section (table 3 and fig. 7) are true within a factor of $0.95 \times \sqrt{3}/2, 0.9 \times \sqrt{3}/2, 0.5 \times \sqrt{3}/2$ for $M_{\rm vir} = 10^{13} M_\odot, 10^{14} M_\odot, 10^{15} M_\odot$ respectively. The the edge detection probabilities remain significant for $M = 10^{15} M_{\rm vir}$ but become marginal for mid size clusters. The most interesting feature shown on figs. 9-10 is that the best fitting parameter values can deviate greatly off the fiducial parameters. In addition, the S/N of particular clusters have saddle points and multiple local minima in terms of the p_2 parameters. This is the case for $M_{\rm vir} = 10^{14} M_\odot$, where the local minimum near $p_2 = p$ is not global, as the global minimum (T_s', c') is at a much larger SZ temperature and a much smaller concentration.

6.3.1 Locating the minimum – obtaining p'

Finding the minimum value in H_2 parameter space in general can be computationally tedious. However, as 10 panels of figs. 9 and 10 display, the numerical value obtained naively by

$$p'^{(0)} = (\alpha', T_s', c') = (\alpha, T_s, c) \tag{92}$$

is a fair zeroth approximation in most cases.[24] A better approximation can be obtained by expanding $[\boldsymbol{h_2}(p_2) - \boldsymbol{h_1}(p)]^2$ to second order around $p'^{(0)}$ and finding its minimum in terms of p_2.

$$\begin{aligned}(\boldsymbol{h_2}(p') - \boldsymbol{h_1}(p))^2 \approx & (\boldsymbol{h_2} - \boldsymbol{h_1})^2 + 2(\partial_k \boldsymbol{h_2}) \cdot (\boldsymbol{h_2} - \boldsymbol{h_1}) x_k + \\ & + [(\partial_j \boldsymbol{h_2}) \cdot (\partial_k \boldsymbol{h_2}) + \partial_j \partial_k \boldsymbol{h_2} \cdot (\boldsymbol{h_2} - \boldsymbol{h_1})] x_j x_k \end{aligned} \tag{93}$$

where $x_j = p'^{(1)} - p'^{(0)}$ and

$$\partial_j \boldsymbol{h_2} = \frac{\partial \, \boldsymbol{h_2}(p_2)}{\partial p_{2_j}} \tag{94}$$

[24] The α parameter will not be varied, $\alpha' = \alpha$ is assumed throughout this study.

On the RHS it is assumed that $\boldsymbol{h_2}$ and its derivatives are evaluated at the naive value $p'^{(0)}$. Let us denote the quadratic coefficient by

$$M_{kj} = (\partial_k \boldsymbol{h_2}) \cdot (\partial_j \boldsymbol{h_2}) + (\boldsymbol{h_2} - \boldsymbol{h_1}(p)) \cdot \partial_k \partial_j \boldsymbol{h_2} \tag{95}$$

Since p' denotes the minimum value, the x_k derivative of eq. (93) must vanish

$$\frac{\partial}{\partial x_k}(\boldsymbol{h_2}(p') - \boldsymbol{h_1}(p))^2 \approx 2(\partial_k \boldsymbol{h_2}) \cdot (\boldsymbol{h_2} - \boldsymbol{h_1}) + 2M_{kj}x_j = 0 \tag{96}$$

Since $x_j = p'^{(1)} - p'^{(0)}$, the solution of this equation gives the next approximation of p'. Eq. (96) is solved by inverting the coefficient matrix. Thus

$$p_k'^{(1)} = p_k'^{(0)} - M_{kj}^{-1}(\partial_j \boldsymbol{h_2}) \cdot (\boldsymbol{h_2} - \boldsymbol{h_1}(p)) \tag{97}$$

$p_k'^{(1)}$ is the improved approximation of the $k^{\rm th}$ parameter of p'. Note that the I_s-dependence of the hypotheses is linear, implying $\partial_{T_s} \boldsymbol{h_2}(p_2) = \boldsymbol{h_2}(p_2, T_s = 1)$ and $\partial^2_{T_s} \boldsymbol{h_2}(p_2) = 0$ for all p_2. The α and c parameter derivatives are to be calculated numerically using eq. (31). This way, finding the minimum in eq. (88) is simplified to evaluating parameter derivatives of the $\boldsymbol{h_2}$ function at only $p'^{(0)}$.

Figs. 9, 10, and 11 show how the best fitting p' parameters are related to the naive choice p. The S/N function has a minimum at p'. Finding the critical point with eq. (97) yields the correct global minimum for $(M_{\rm vir}, z, \alpha) = (10^{13} M_\odot, 0.3, 1.5)$ for any $(x_{\rm max}, D)$ and $(10^{14} M_\odot, 0.3, 1.5)$ for any $(x_{\rm max}, D)$ unless $x_{\rm max} \approx c$. However, for $(M_{\rm vir}, z, \alpha, x_{\rm max}) \approx (10^{14} M_\odot, 0.3, 1.5, c)$ there are multiple local minima and saddle points. The $p' = p'^{(1)}$ approximation of eq. (97) yields the local minima in the vicinity, which is *not* the global minimum. For $(M_{\rm vir}, z, \alpha) \approx (10^{15} M_\odot, 0.3, 1.5)$ and arbitrary $(x_{\rm max}, D)$, the $p'^{(1)}$ value corresponds to the saddle point. Again, the $p' = p'^{(1)}$ approximation of eq. (97) breaks down.

The saddle points can be identified by calculating $\det M_{i,j}$

$$\det M_{ij}(p_2) \begin{array}{ll} > 0 & \text{iff } p_2 \text{ is at a local minimum or maximum} \\ < 0 & \text{iff } p_2 \text{ is at a saddle point} \end{array} \tag{98}$$

since M_{ij} is a 2×2 matrix and it is well-defined everywhere in the parameter space. The saddle point can be eluded if $\det M$ is calculated and a step is made towards $p'^{(1)}$ only if $\det M > 0$. If it is negative, then a step is made "downhill" along the negative gradient $-2(\partial_k \boldsymbol{h_2}) \cdot (\boldsymbol{h_2} - \boldsymbol{h_1})$.

Repeating eq. (97) or the downhill steps locates the local minimum near the fiducial parameter value p. Figs. 9, 10, and 11 indicate that this algorithm leads to the correct p' value for the 16 of the 18 cases considered on

the plots. Only the fiducial parameters $p = (M_{\rm vir}, z, \alpha) \approx (10^{14} M_\odot, 0.3, 1.5)$, with $(x_{\rm max}, D) = (c, 0.01c)$ or $(c, 0.5c)$ have non-global minimum in the close neighborhood of p. In this case, the minimum can be obtained with a manual input of $(T_s'^{(0)}, c'^{(0)}) = (2T_s, 0.2c)$, thereby eluding the local minimum.

7 Parameter estimation

In the previous section, we have examined the signal to noise ratio for the hypothesis test. The primary aim of the test was to make a distinction between the possibilities, that the cluster does or does not have an edge. Another problem is to examine the precision of the parameter estimation of $\boldsymbol{h_1}$ itself. We calculate the uncertainty of the parameter estimation with the Fisher matrix.

In this section we assume that the real signal is $\boldsymbol{h_1}$, with associated p_i parameters. Now let us suppose that a signal $\boldsymbol{y}$ arrives, and with no prior information, a parameter p_1 is to be chosen that best describes the data. Due to random noise, the measurement yields a parameter estimator with some uncertainty. The parameter estimator can be obtained with the maximum likelihood test. The likelihood function is

$$L(\boldsymbol{y}, p) = P(\boldsymbol{y} \,|\, \boldsymbol{h_1}(p)) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{[\boldsymbol{y} - \boldsymbol{h_1}(p)]^2}{2\sigma_N^2} \right) \tag{99}$$

and the log likelihood

$$\ln L(\boldsymbol{y}, p) = -\frac{[\boldsymbol{y} - \boldsymbol{h_1}(p)]^2}{2\sigma_N^2} \tag{100}$$

The parameters are chosen to maximize $L(\boldsymbol{y}, p)$, or equivalently to apply the least squares fit. The quantity

$$S^2 = [\boldsymbol{h_1}(p_1) - \boldsymbol{h_1}(p)]^2 \tag{101}$$

has a χ^2 distribution in terms of $\boldsymbol{h_1}(p_1)$. Taking the noise power to be the noise in a single bin, $N^2 = \sigma_N^2$, the false signal to noise

$$\frac{S}{N} = \frac{\sqrt{[\boldsymbol{h_1}(p_1) - \boldsymbol{h_1}(p)]^2}}{\sigma_N} \tag{102}$$

Eq. (102) is closely related to the χ-statistic that the parameter estimator is p_1 instead of p, the true value. Given a p true parameter set, the region within 2σ confidence (95%) for example is the set $\{p_1 \in H_1 \,|\, S/N < 2\}$. Thereby evaluating eq. (102), the precision of the parameter estimation can be readily read off. Figures 12, 13, and 14 depict this precision for various p true values. The contour plots have $p_1 = (M_{\rm vir}, z, \alpha, T_1, c_1, x_{\rm max1}, D_1)$ with $(x_(\rm max1), D_1)$ varied, while all other three parameters are fixed at the true value.

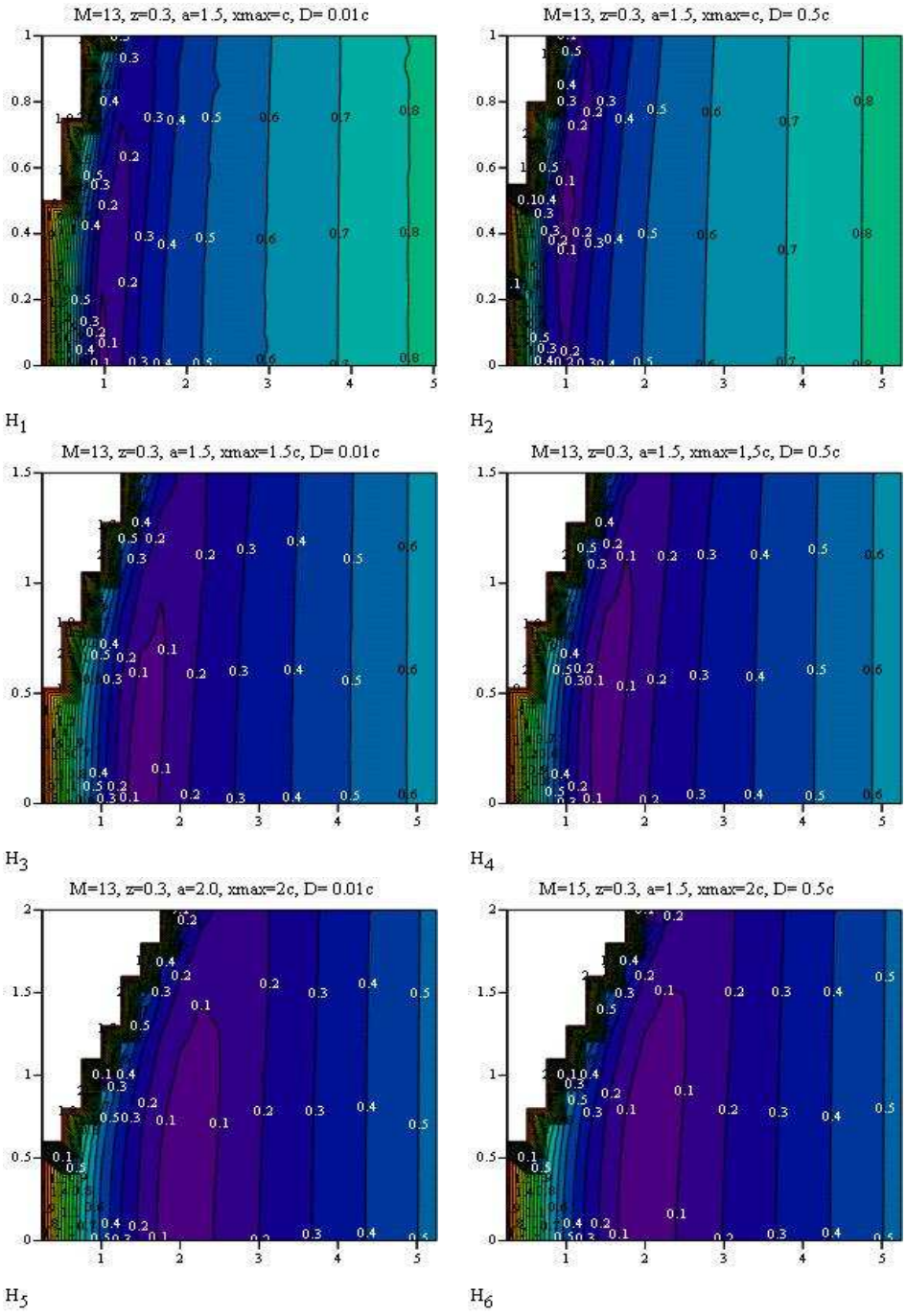

Figure 12: The parameter dependence χ-contours for $(M, z, \alpha) = (10^{13} M_{\odot}, 0.3, 1.5)$ and 6 choices for $(x_{\max}, D)$. The top left corners, where $D > x_{\max}$, were not calculated.

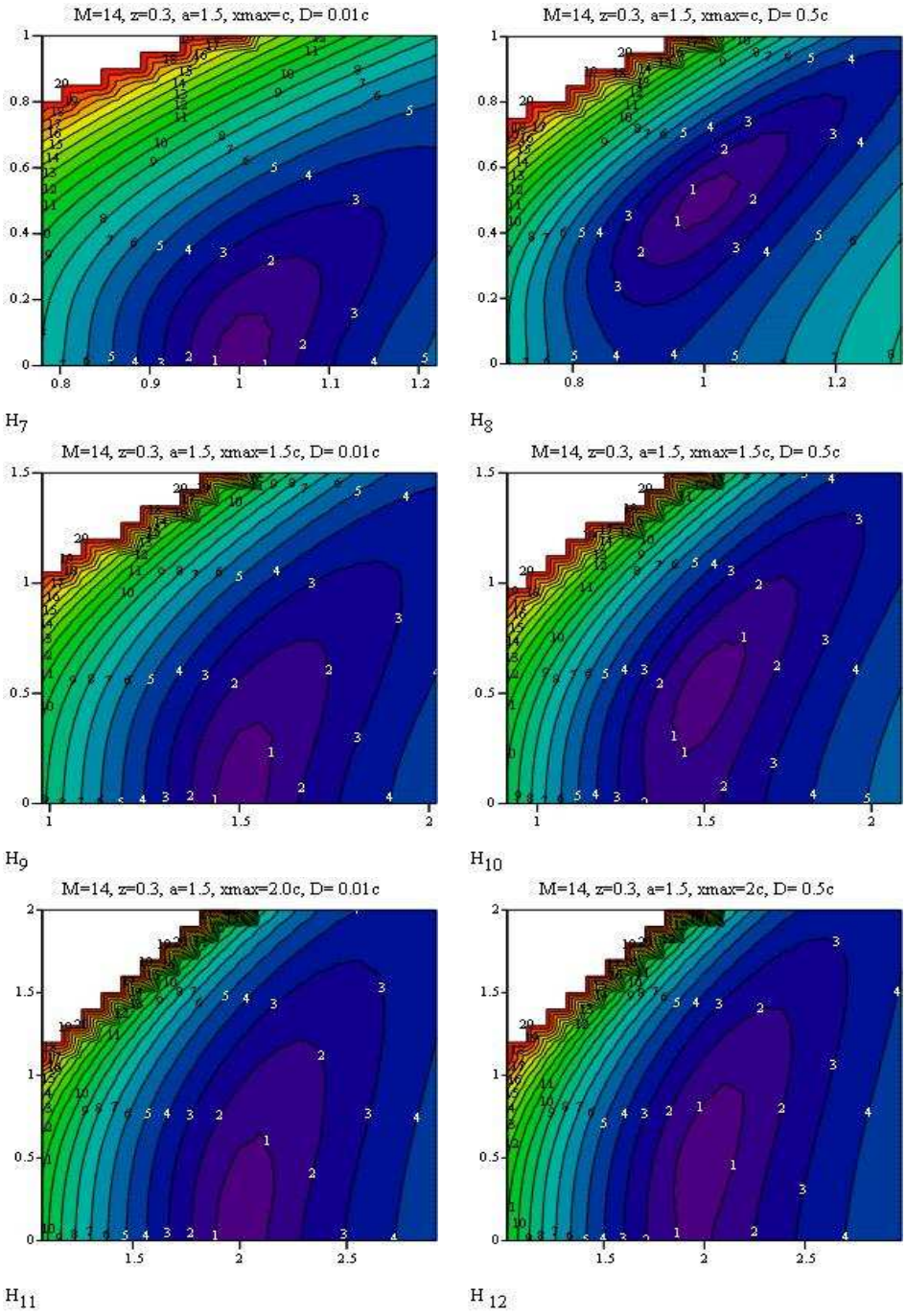

Figure 13: The parameter dependence χ-contours for $(M, z, \alpha) = (10^{14} M_{\odot}, 0.3, 1.5)$ and 6 choices for $(x_{\max}, D)$. The top left corners, where $D > x_{\max}$, were not calculated.

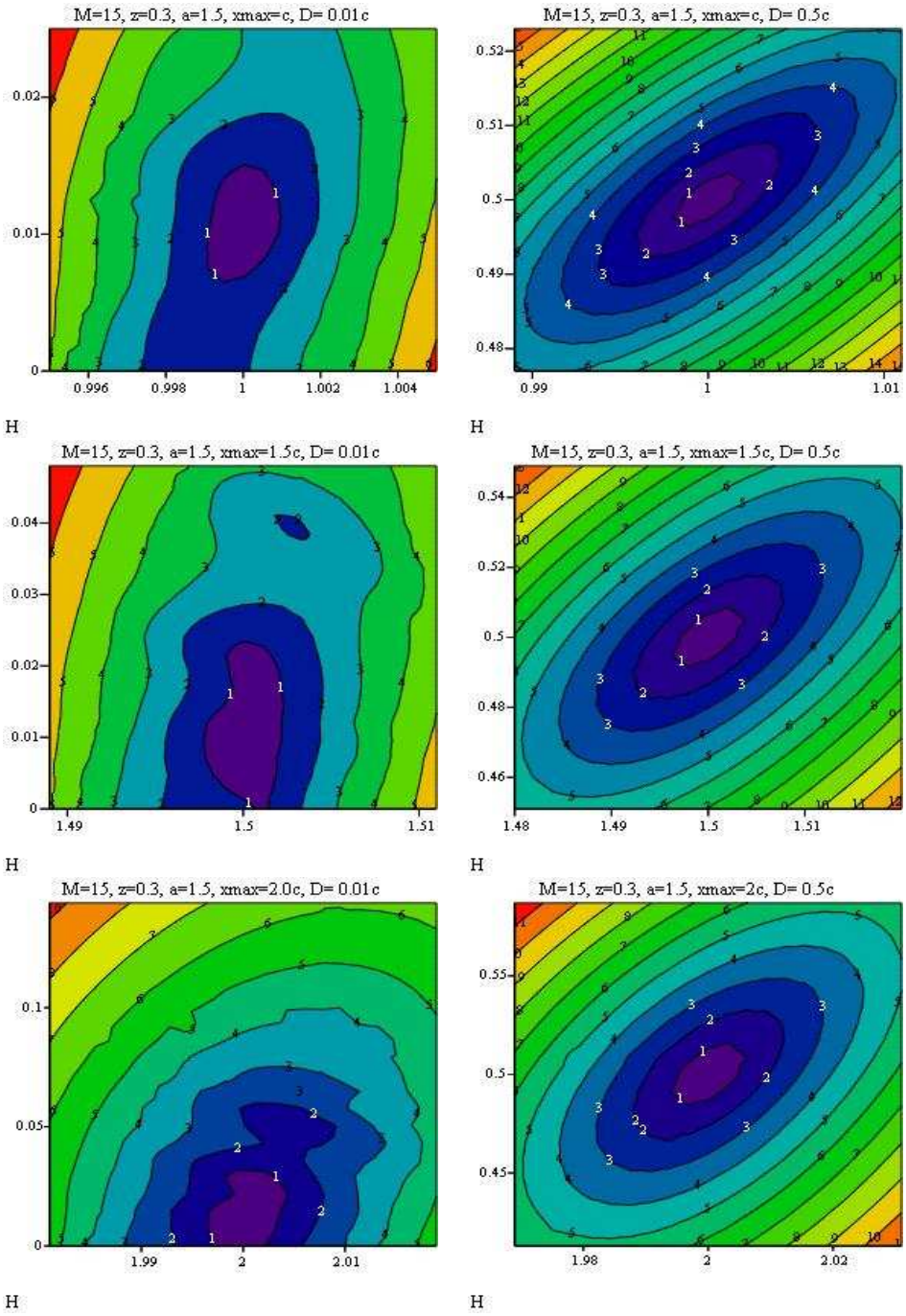

Figure 14: The parameter dependence χ-contours for $(M, z, \alpha) = (10^{15} M_{\odot}, 0.3, 1.5)$ and 6 choices for $(x_{\max}, D)$.

The Fisher information matrix is

$$I_F(p) = -\left\langle \frac{\partial^2 \ln L(\boldsymbol{y}, p)}{\partial p_i \partial p_j} \right\rangle_{\boldsymbol{y}} = \frac{1}{\sigma_N^2} \frac{\partial \boldsymbol{h_1}}{\partial p_i} \cdot \frac{\partial \boldsymbol{h_1}}{\partial p_j} \tag{103}$$

The finite sum can be approximated with the integral formula for a sensitive observation.

$$I_F(p) = \left(\frac{\Theta}{c\Delta\phi}\right)^2 \int_0^{x_{\max}+D} \frac{\partial h_1(x,p)}{\partial p_j} \frac{\partial h_1(x,p)}{\partial p_j} x \mathrm{d}x \tag{104}$$

Assuming that the likelihood distribution is Gaussian near the peak likelihood, we can use confidence limits for Gaussian statistics (i.e., χ^2) to obtain 68% and 95% confidence regions.

The minimum expected variance is related to the Fisher matrix with the Cramer-Rao bound

$$(\Delta p_j)^2 \geq (I_F^{-1})_{jj} \tag{105}$$

where equality holds if the distribution is well approximated with a Gaussian distribution.

Tables 5 and 4 show the 68% and 95% significance uncertainties for Model I and II as calculated from the Fisher matrix. The calculations indicate that the uncertainties are nearly the same for Model I and II regarding $x_{\max}$ and D. The Fisher matrix itself is decoupled in the corresponding subspaces, as the $(x_{\max}, T_s)$, $(x_{\max}, c)$, (D, T_s), (D, c) components are negligible compared to the diagonal elements. The variance in the parameter estimators are $(\Delta c/c, \Delta T_s/T_s, \Delta x_{\max}/x_{\max}, \Delta D/c) = (3\%, 7\%, 3\%, 13\%)$ for $(M_{\mathrm{vir}}, z, \alpha, c, T, x_{\max}, D) = (10^{14} M_\odot, 0.3, 1.5, 5, 0.1\mathrm{mK}, c, 0.01c)$. Therefore c, T, and $x_{\max}$ will be precisely obtained with ALMA for the majority of the clusters, while similar precision for D is possible for only massive clusters. Let us point out, that it is possible to deduce the α parameter from the $c(\alpha)$ relation. Table 5 shows, that the decision significance will be well over 95% even for clusters as small as $M_{\mathrm{vir}} = 10^{13} M_\odot$. Therefore, the observation with ALMA will surely end the controversy between the two possibilities for α.

Figures 12, 13, and 14 depict the likelihood contours for Model II with the direct evaluation of eq. (102), without the Fisher matrix approximation. The 1σ contours are well-approximated with ellipses, if $D - \Delta D > 0$ for $M_{\mathrm{vir}} \geq 10^{14} M_\odot$, e.g. $D > 0.1$, $x_{\max} < 1.7c$, and $M_{\mathrm{vir}} \geq 10^{14} M_\odot$. Therefore, in these cases, the inequality (105) assumes the equality for the parameter uncertainty. The parameter distribution around $D \approx 0$ is distorted, and higher confidence level contours for arbitrary D are banana shaped in the $x_{\max} - D$ plane.

$M_{\rm vir}$ $[M_\odot]$	z	α	Θ $[']$	c	T_s $[-\mu\mathrm{K}]$	$T(0)$ $[-\mu\mathrm{K}]$	$x_{\rm max}$ $[c]$	D $[c]$	S %
10^{13}	0.3	1	1.8	7.9	18	22	1	0.01	-
-	-	-	-	-	-	-	0.73	3.3	68
-	-	-	-	-	-	-	1.5	6.6	95
10^{13}	0.3	1.5	1.7	4.6	16	27	1	0.01	-
-	-	-	-	-	-	-	0.73	3.3	68
-	-	-	-	-	-	-	1.5	6.5	95
10^{14}	0.3	1	3.8	5.0	97	130	1	0.01	-
-	-	-	-	-	-	-	0.028	0.13	68
-	-	-	-	-	-	-	0.055	0.26	95
10^{14}	0.3	1.5	3.8	2.9	1100	160	1	0.01	-
-	-	-	-	-	-	-	0.027	0.13	68
-	-	-	-	-	-	-	0.054	0.26	95
10^{15}	0.3	1	8.1	3.1	560	780	1	0.01	-
-	-	-	-	-	-	-	0.001	0.005	68
-	-	-	-	-	-	-	0.002	0.010	95
10^{15}	0.3	1.5	8.1	1.8	7000	910	1	0.01	-
-	-	-	-	-	-	-	0.001	0.004	68
-	-	-	-	-	-	-	0.002	0.008	95

Table 4: Parameter variance, Model II

$M_{\rm vir}$ $[M_\odot]$	z	α	Θ $[']$	c	T_s $[-\mu\mathrm{K}]$	$T(0)$ $[-\mu\mathrm{K}]$	$x_{\rm max}$ $[c]$	D $[c]$	S %
10^{13}	0.3	1	1.8	7.9	18	22	1	0.01	-
-	-	-	-	1.0	15	12	0.76	3.3	68
-	-	-	-	2.0	30	24	1.5	6.6	95
10^{13}	0.3	1.5	1.7	4.6	16	27	1	0.01	-
-	-	-	-	0.80	132	22	0.76	3.3	68
-	-	-	-	1.6	264	45	1.5	6.6	95
10^{14}	0.3	1	3.8	5.0	97	130	1	0.01	-
-	-	-	-	0.14	6.4	4.8	0.029	0.13	68
-	-	-	-	0.29	13	9.6	0.058	0.27	95
10^{14}	0.3	1.5	3.8	2.9	1100	160	1	0.01	-
-	-	-	-	0.11	67	9.6	0.028	0.12	68
-	-	-	-	0.21	133	19	0.056	0.26	95
10^{15}	0.3	1	8.1	3.1	560	780	1	0.01	-
-	-	-	-	0.014	2.3	1.6	0.001	0.005	68
-	-	-	-	0.027	4.6	3.2	0.002	0.010	95
10^{15}	0.3	1.5	8.1	1.8	7000	910	1	0.01	-
-	-	-	-	0.011	28	3	0.001	0.004	68
-	-	-	-	0.023	56	7	0.002	0.008	95

Table 5: Parameter variance, Model I

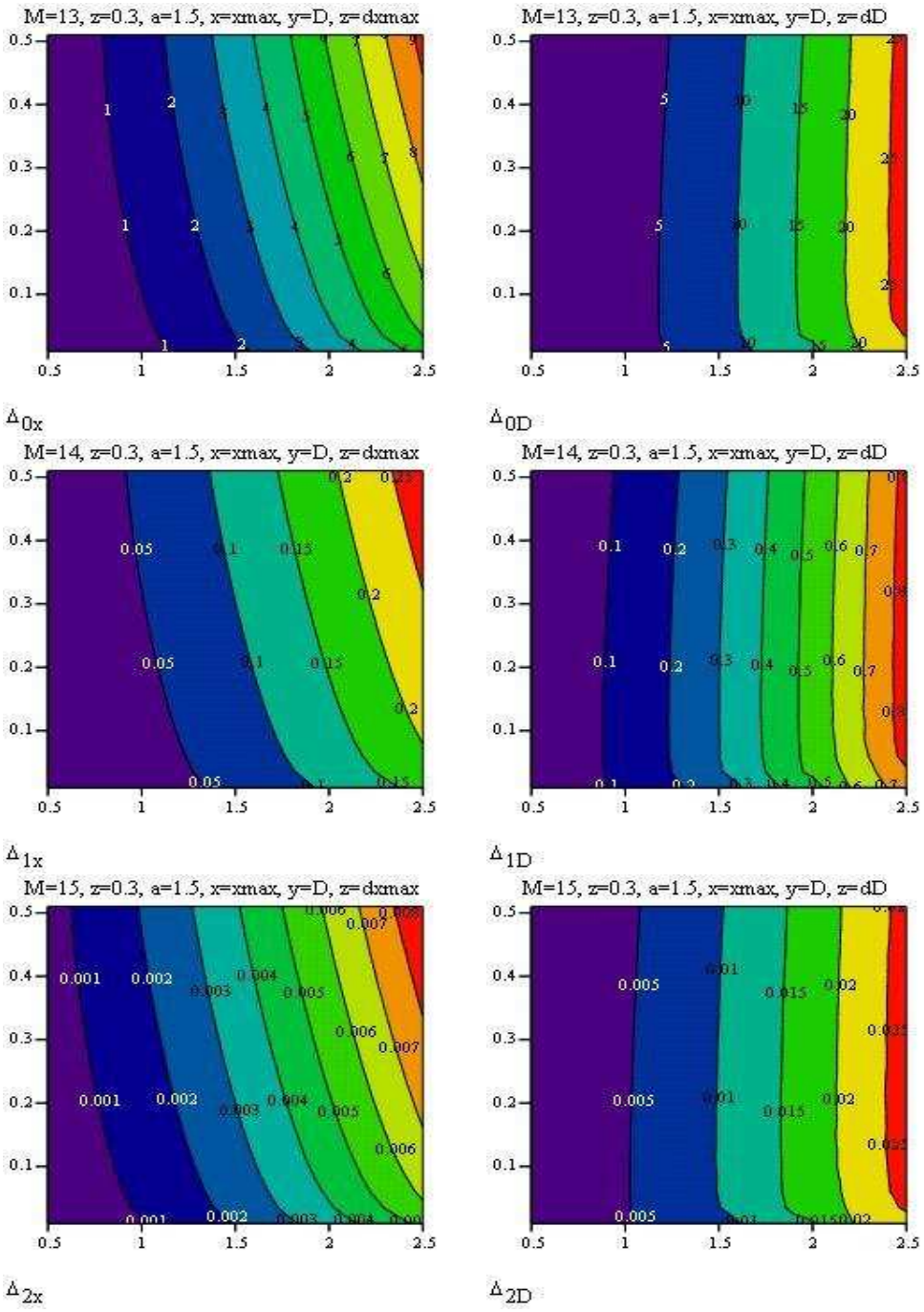

Figure 15: The $\Delta x_{\rm max}$ *(left panels)* and ΔD *(right panels)* variances are shown for various $x_{\rm max}$ *(x-axis)* and D values *(y-axis)*. Model II Fisher matrix was used, with $(z, \alpha) = (0.3, 1.5)$ and $M_{\rm vir} = 10^{13}, 10^{14}, 10^{15} M_\odot$. The units are c (with the appropriate Model II value).

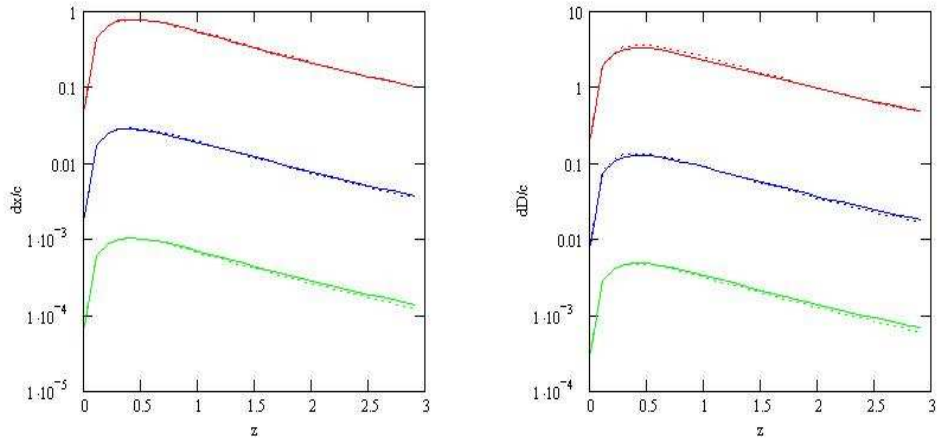

Figure 16: The $\Delta x_{\max}$ *(left)* and ΔD *(right)* uncertainty for various z. Model II was used, and the virial mass was taken 10^{13} *(red)*, 10^{14} *(blue)*, and $10^{15}{\cdot}M_{\odot}$ *(green)*. The units on are c (with the particular Model II value).

Figures 15 and 16 give an overview of the $\Delta x_{\max}$ and ΔD variances for various parameter choice. These plots were calculated using the Fisher matrix. The plots show that increasing $x_{\max}$ increases the uncertainties $\Delta x_{\max}$ and ΔD, while increasing D only slightly increases $\Delta x_{\max}$ and leaves ΔD unchanged. These plots correspond to various M_{vir} choices, but use fixed values for $z = 0.3$ and $\alpha = 1.5$. The z and α dependence is plotted on 16. It is clear, that the α parameter does not alter the $\Delta x_{\max}$ and ΔD uncertainties. Changing z yields a maximum in $\Delta x_{\max}$ and ΔD around $z = 0.4$. The parameter uncertainty decreases by a factor of 10 for $z \approx 0$ or $z \approx 3$ compared to the $z = 0.4$ value.

8 Conclusions

We have shown that the near-future Atacama Large Millimeter Array (ALMA) is well-suited for studies of the intra cluster medium density distribution using the Sunyaev-Zel'dovich (SZ) effect. The angular beam diameter and sensitivities are predicted to reach 2" and 10μK for this system, which exceed present detector resolutions with more than two orders of magnitude. A rich galaxy cluster observed at 100 GHz for 30 hours yields a SZ signal-to-noise ratio that is high enough to make inferences about the barionic and dark matter distributions. We showed that the ALMA detector will be capable of judging whether the current assumptions on the cluster physics need to be revised.

The hydrodynamic equations were solved for the self-similar gas density in the dark matter background. The dark matter profile was parameterized with α, and every calculation was evaluated for both values common in the literature, i.e. $\alpha = 1$ and 1.5. Within this framework, the resultant density profile as a function of radial distance is small but nonzero even in the faraway regions. As a modification, we introduced a hypothetical linear cutoff in the density profile, where the density falls to zero within a finite radial distance. This assumption is consistent with the theory of cluster evolution (see [14] and [15]).

We have calculated the change in the SZ image of the cluster with the ALMA detector. We assumed 10μK flat-power Gaussian noise. We constructed the optimal filter matched to the self-similar cluster density model. The corresponding likelihood function was used to see whether our fiducial model with an edge can be distinguished from the original model without an edge. It is important to emphasize that this test is practically model independent, since it is only weakly sensitive to the central region, but this sensitivity is significant in the outer region, where the various cluster models in the literature are identical. We calculated the signal to noise ratio in two ways, with 4 (Model I) and 2 free parameters (Model II), respectively. These parameters are the radial distance of the edge, $x_{\max}$, the length of the cutoff, D, the concentration parameter, c, and the central temperature decrement, $T(0)$, for Model I. For Model II, c and $T(0)$ were calculated from theory assuming no scatter around the correct values. Other parameters, such as the virial mass, $M_{\rm vir}$, the redshift, z, the dark matter exponent, α, and the virial angular radius, Θ, can be measured to a higher precision and were not varied. The results are within 10% for Model I and Model II. The signal to noise ratio is 470 for a rich cluster of virial mass $10^{15} \times M_{\odot}$ and it is 16 for a regular cluster with $10^{14} \times M_{\odot}$. The parameters such as c and $T(0)$ were assumed to be the correct values for both models with and without a cutoff.

Next, a more realistic case was considered, when the parameters of the cluster with and without an edge are fitted independently for a given measurement. As it turned out, there is a large systematical error in parameters c és $T(0)$ for the model without an edge, provided that the cluster does have an edge within approximately one virial radius. Nevertheless, the signal to noise ratio decreases only by 56% at most for detecting the edge. Therefore, the edges remain visible for normal and rich clusters.

Finally, we examined the precision of the parameter estimation of the model with an edge. As a result, typical values for normal and rich clusters are $(\Delta c/c, \Delta T(0)/T(0), \Delta x_{\max}/c, \Delta D/c) = (2.8\%, 3.7\%, 2.9\%, 13\%)$ and $(0.45\%, 0.2\%, 0.1\%, 0.5\%)$, respectively. Since there is a one-to-one correspondence between c and α, with $c(\alpha = 1)/c(\alpha = 1.5) = 1.7$, the SZ effect is capable to decide which α model best describes reality. The controversy between the two dark matter profiles will be solved. Measuring the SZ effect with ALMA will not only detect the cluster's edge, but it will be able to locate it with a high precision. This information will surely improve our understanding of the structure of matter within the clusters.

References

[1] Birkinshaw, M., "The Sunyaev-Zel'dovich effect," 1999, Physics Reports, 310, 97.

[2] Carlstrom, J., Holder, G., Reese, E. D., "Cosmology with the Sunyaev-Zel'dovich effect," 2002, ARA&A, 2002, 40, 643.

[3] Bartlett, J.G. "Sunyaev-Zel'dovich surveys: Analytic treatment of cluster detection", 2001, A&A, preprint astro-ph/0001267

[4] Holder, G.; Mohr, J.; Carlstrom, J.; Evrard, A.; Leitch, E. M., "Expectations for an Interferometric Sunyaev-Zeldovich Effect Survey for Galaxy Clusters," 2000, ApJ, 544, 629

[5] Barbosa, D., Bartlett, J., Blanchard, A., and Oukbir, J. 1996, A&A, 314, 13.

[6] Kneissl, R., Jones, M. E., Saunders, R., Eke, V. R., Lasenby, A. N., Grainge, K., Cotter, G., "Surveying the sky with the Arcminute MicroKelvin Imager: expected constraints on galaxy cluster evolution and cosmology," 2001, MNRAS, 328, 783

[7] Springel V., Martin W., Hernquist L., "Hydrodynamic simulations of the Sunyaev-Zeldovich effects," 2001, ApJ, 549, 681.

[8] Dawson, K. S., Holzapfel, W. L., Carlstrom, J. E., Joy, M., LaRoque, S. J., Reese, E. D., "A preliminary detection of arcminute-scale cosmic microwave background anisotropy with the BIMA array," 2001, ApJ, 553, 1L.

[9] Desert F, Benoit A, Gaertner S, Bernard J, Coron N, et al, "Observations of the Sunyaev-Zel'dovich effect at high angular resolution towards the galaxy clusters A665, A2163 and CL0016+16," 1998, New Astron., 3, 655.

[10] Holzapfel, W. L., Wilbanks, T. M., Ade, P. A., Church, S. E., Fischer, M. L., Mauskopf, P. D., Osgood, D. E., Lange, A. E. "The Sunyaev-Zeldovich infrared experiment: A millimeter-wave receiver for cluster cosmology," 1997, ApJ, 479, 17H.

[11] Mauskopf, P. D., Rownd, B. K., Edgington, S. F., Hristov, V. V., Mainzer, A. K., et al. 2000b. In Imaging at Radio through Submillimeter Wavelengths Astron. Soc. Pac., Conf. Ser., ed. J Mangum, p. E46. San Francisco: Publ. Astron. Soc. Pac.

[12] Jones, M., Saunders, R., Alexander, P., Birkinshaw, M., Dilon, N., et al., "An image of the. Sunyaev-Zel'dovich effect," 1993, Nature, 365, 320

[13] Mason, B. S., Myers S. T., Readhead A. C., "A measurement of H_0 from the Sunyaev-Zeldovich effect," 2001. ApJ, 555, L11.

[14] Bertschinger E., "Self-similar secondary infall and accretion in an Einstein-de Sitter universe," 1985, ApJS, 58, 39.

[15] Bagchi, J., Enßlin, T. A., Miniati, F., Stalin, C. S., Singh, M., Raychaudhury, S., Humeshkar, N. B., "Evidence for shock acceleration and intergalactic magnetic fields in a large-scale filament of galaxies ZwCl 2341.1+0000," 2002, NewA, 7, 249B.

[16] Shapiro, P. R., Iliev I. T., Raga, A. C., "The post-collapse equilibrium structure of cosmological haloes in a low-density universe," 2001, MNRAS, 325, 468.

[17] Cavaliere, A., Fusco-Femiano, R., "The Distribution of Hot Gas in Clusters of Galaxies," 1978, A&A, 70, 677.

[18] Mohr, J., Mathiesen, B., Evrard, A., "Properties of the intracluster medium in an ensemble of nearby galaxy clusters," 1999, ApJ, 517, 627.

[19] Holder, G., Carlstrom, J. E., "Understanding cluster gas evolution and fine-scale CMB anisotropy with deep Sunyaev-Zel'dovich effect surveys," 2001, ApJ, 558,515.

[20] E. Komatsu, U. Seljak, "Universal gas density and temperature profiles," 2001, MNRAS, **327**, 1353-1366.

[21] Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493.

[22] Moore, B., Quinn, T., Governato, F.,Stadel, J., Lake, G., "Cold collapse and core catastrophy," 1999, MNRAS, 310, 1147.

[23] Jing Y. P., Suto Y., 2000, ApJ, 529, L69.

[24] Suto Y., Sasaki S., Makino N., 1998, ApJ, 509, 544.

[25] Holder G., "Measuring Cluster Peculiar Velocities and Temperatures at cm and mm Wavelengths", ApJ perprint doi:10.1086/380916.

[26] Butler, B., Blitz, L., Wlech, J., Carlstrom, J., Woody, D., Churchwell, E., "Report on the Antenna Size Committee Meeting," 1999, MMA Memo. No. 243.

[27] Holdaway, M. A., Juan, R. P, "Modeling of the submillimeter opacity n Chanantor" 1999, MMA Memo. No. 187.

[28]

[29] Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 469, 494.

[30] Pen U., astro-ph/9904172.

[31] Barkana R., Loeb A., astro-ph/0010468.

[32] Evrard, A., "The intracluster gas fraction in X-ray clusters - Constraints on the clustered mass density," 1997, MNRAS, 292, 289.