Reconstructing Small Scale Lenses from the Cosmic Microwave Background

Benjamin Horowitz,[1] Simone Ferraro,[1,2,3] Blake D. Sherwin,[4,1]

[1]*Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720*

[2]*Department of Astronomy, University of California, Berkeley, CA, USA 94720*

[3]*Miller Institute for Basic Research in Science, University of California, Berkeley, CA, 94720 USA*

[4]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA*

October 30, 2017

ABSTRACT

Cosmic Microwave Background (CMB) lensing is a powerful probe of the matter distribution in the Universe. The standard quadratic estimator, which is typically used to measure the lensing signal, is known to be suboptimal for low-noise polarization data from next-generation experiments. In this paper we explain why the quadratic estimator will also be suboptimal for measuring lensing on very small scales, even for measurements in temperature where this estimator typically performs well. Though maximum likelihood methods could be implemented to improve performance, we explore a much simpler solution, revisiting a previously proposed method to measure lensing which involves a direct inversion of the background gradient. An important application of this simple formalism is the measurement of cluster masses with CMB lensing. We find that directly applying a gradient inversion matched filter to simulated lensed images of the CMB can tighten constraints on cluster masses compared to the quadratic estimator. While the difference is not relevant for existing surveys, for future surveys it can translate to significant improvements in mass calibration for distant clusters, where galaxy lensing calibration is ineffective due to the lack of enough resolved background galaxies. Improvements can be as large as $\sim 50\%$ for a cluster at $z = 2$ and a next-generation CMB experiment with 1μK-arcmin noise, and over an order of magnitude for lower noise levels. For future surveys, this simple matched-filter or gradient inversion method approaches the performance of maximum likelihood methods, at a fraction of the computational cost.

1 INTRODUCTION

Weak gravitational lensing of the Cosmic Microwave Background (CMB) by the large-scale structure of the Universe produces a signature in the CMB anisotropies in both temperature and polarization and provides a unique probe of the matter distribution out to high redshift (Seljak (1995)). Lensing induces statistical anisotropies in the observed CMB, through which the mass distribution can be probed. The most common method of lensing reconstruction relies on the quadratic estimator (QE) of Hu and Okamoto (Hu & Okamoto (2002)). Previous work (Hirata & Seljak (2003a); Carron & Lewis (2017)) has shown that, since the QE is only an approximation to the maximum likelihood solution, significant improvements are possible in polarization at low noise levels, usually at the cost of a much greater computational complexity.

In this paper, we revisit lensing reconstruction on very small scales, in a regime where the primary CMB fluctuations have been Silk damped and are negligible (i.e. $\ell \gtrsim 4000$). We will first show that the quadratic estimator can be very suboptimal in this regime as well, since it is limited by the cosmic variance of the long-wavelength back-

ground gradient mode, while a true maximum likelihood solution should be able to reconstruct small lenses arbitrarily well, given small enough instrumental noise and residual foreground levels. We then show that the statistical power of the QE is further reduced by the fact that the errors become highly correlated, so that the utility of measuring more modes is reduced when the proper covariance between them is taken into account. Finally, we show that a simple gradient inversion matched filtering approach, as proposed by Seljak & Zaldarriaga (2000), is close to optimal for reconstruction of very small scale lenses, and that it is not limited by cosmic variance, thus in principle allowing arbitrarily large improvements over the QE (for small enough noise).

One application of this formalism with cosmological significance is cluster mass determination though CMB lensing, or "cluster lensing". The number counts of galaxy clusters are a key prediction of the Λ-CDM model which is sensitive to the overall matter density, Ω_m, as well as the amplitude of density fluctuations, σ_8, and hence the neutrino mass ΣM_ν. A key limiting factor of using galaxy clusters for cosmological studies is their mass calibration, which is difficult to obtain (Rozo et al. (2013)). Since it isn't directly observable,

cosmologists are left to study different proxies of mass, such as X-ray luminosity or Sunyaev-Zel'dovich signal, and correlate them with gravitational effects of the cluster, such as galaxy-cluster lensing. The uncertainty in the calibration of these relations now limits the constraining power of cluster counts as a cosmological probe (Mantz et al. (2014)).

While low-redshift massive clusters can be accurately weighed by using lensing of background galaxies along the line of sight (von der Linden et al. (2014); Madhavacheril et al. (2017)), this approach fails for more distant clusters (at $z \gtrsim 1$) because of the lack of enough resolved galaxies on the background. In addition, uncertainties in the photometric redshift distribution of source galaxies, which is challenging to determine precisely at high redshift, can lead to errors in the mass determination.

Meanwhile, the CMB provides a source at known redshift ($z \approx 1100$) and with well studied statistical properties. By looking at the lensing of the CMB by these galaxy clusters, we can avoid some of the shortcomings of galaxy lensing surveys and probe more distant clusters (Holder & Kosowsky (2004); Yoo & Zaldarriaga (2008); Yoo et al. (2010); Melin & Bartlett (2015)). While development of the theory behind these measurements dates back almost two decades (Zaldarriaga & Seljak (1998); Seljak & Zaldarriaga (2000); Lewis & King (2006)), it is only fairly recently that this lensing has been detected at high statistical significance (Melin & Bartlett (2015); Madhavacheril et al. (2015); Baxter et al. (2015))

These detections have used a modified version of the QE (Hu et al. (2007)) which relies on reconstructing the lensing convergence κ, and then applying a matched filter (Melin & Bartlett (2015)) on the resulting convergence map. However, as discussed in (Yoo & Zaldarriaga (2008)), QE obtains unbiased and optimal results only in the limit of no gravitational lensing and becomes progressively more biased and sub-optimal as the gradient in lensing potential increases. Clusters found in SZ samples have particularly strong potential gradients due to their large mass and pose a particularly serious problem for quadratic estimator techniques. The work by (Baxter et al. (2015)) employs a maximum-likelihood approach and should be immune to the discussion above, but is more computationally expensive.

Revisiting the work of Seljak & Zaldarriaga (2000) in the context of upcoming CMB experiments such as CMB S4 (Abazajian et al. (2016)), we discuss how an unbiased estimate for the mass of a small spherically symmetric lens (such as an idealized cluster) can be obtained with a simple matched filter of the temperature and polarization maps. In the low-noise regime, we show that these estimates are close to the maximum likelihood solution, and considerably better than the QE. Reconstructing the lensing potential on small scales is particularly important as upcoming CMB experiments can potentially reach sensitivities of $\sim 1\ \mu$K-arcmin and 1 arcmin beams size. With this increase in instrumental sensitivity, there should be corresponding increases in cluster mass calibration and cosmological constraints (Louis & Alonso (2016); Madhavacheril et al. (2017)).

In Section 2, we briefly review weak lensing of the CMB, explore the small-scale limit of the quadratic estimator and explain the matched filter approach, as well as our assumptions about the cluster mass profile. In Section 3, we introduce a matched filter technique we use directly on the
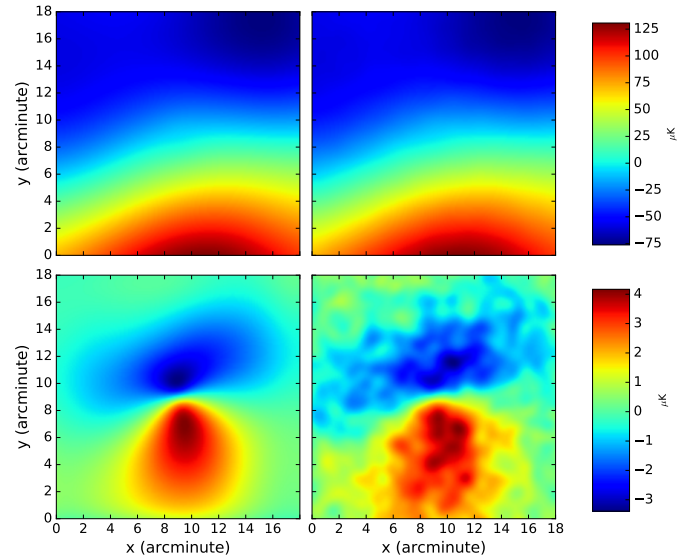
Figure 1. Top: Simulated images of CMB lensing caused by a $2 \times 10^{14}\ M_\odot$ NFW cluster at $z = 0.7$. Bottom: Unlensed image subtracted out. Left: Beam alone, Right: Beam and 1 μK-arcmin noise.

temperature map. In Section 4, we implement our technique on realistic lensed CMB maps and compare against the quadratic estimator, finding improved performance at low noise levels and high redshifts. In Section 5, we compare the results of the matched filter estimate to other lensing reconstruction techniques. Finally, in Section 6, we will summarize our results and look at prospects for applying it to future experiments.

For our analysis we will assume the flat ΛCDM Planck 2013 cosmology (Collaboration et al. (2013)), with $H_0 = 67.8$ km (Mpc s)$^{-1}$, and $\Omega_m = 0.307$. We will assume the Born approximation, and ignore field rotation/multiple-lens effects throughout this work.

2 FORMALISM

2.1 Effect of Cluster Lensing on the CMB

Here we briefly review the effects of gravitational lensing on the CMB; for a more complete review see (Lewis & Challinor (2006)). The observed CMB anisotropies can be described by their temperature fluctuations as a function of direction $\hat{\mathbf{n}}$, $\Theta(\hat{\mathbf{n}})$, as well as two Stokes parameters describing their polarization $Q(\hat{\mathbf{n}})$ and $U(\hat{\mathbf{n}})$. Since the effect of lensing is a simple coordinate remapping, the observed (i.e. lensed) field, $(\tilde{\Theta}, \tilde{Q}, \tilde{U})$, can be related to the primary fields, (Θ, Q, U), through a deflection angle field $\mathbf{d}(\hat{\mathbf{n}}) = \nabla\phi$, where ϕ is the lensing potential, by
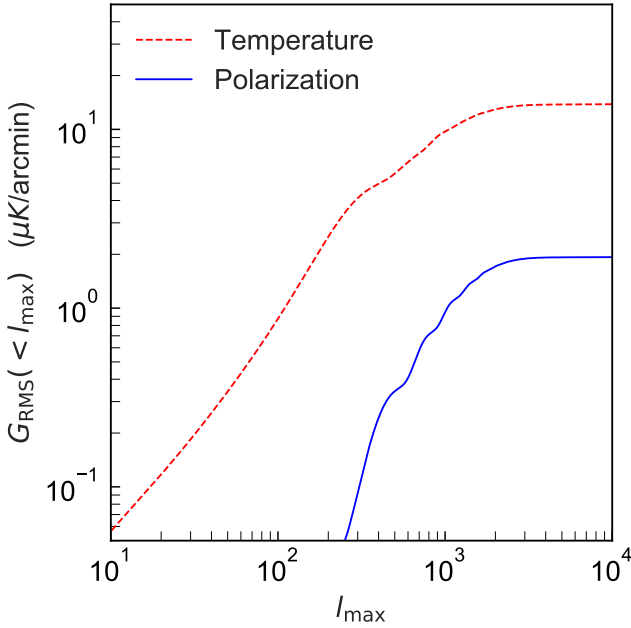
$$\tilde{\Theta}(\hat{\mathbf{n}}) = \Theta(\hat{\mathbf{n}} + \nabla\phi)) \approx \Theta(\hat{\mathbf{n}}) + \nabla\phi \cdot \nabla\Theta(\hat{\mathbf{n}}) + \dots \quad (1a)$$

$$\tilde{Q}(\hat{\mathbf{n}}) = Q(\hat{\mathbf{n}} + \nabla\phi)) \approx Q(\hat{\mathbf{n}}) + \nabla\phi \cdot \nabla Q(\hat{\mathbf{n}}) + \dots \quad (1b)$$

$$\tilde{U}(\hat{\mathbf{n}}) = U(\hat{\mathbf{n}} + \nabla\phi)) \approx U(\hat{\mathbf{n}}) + \nabla\phi \cdot \nabla U(\hat{\mathbf{n}}) + \dots \quad (1c)$$

Here we have truncated the expansion to first order in $\nabla\phi$, and which is referred to as the "gradient approximation".

Note that while the gradient approximation can be a poor approximation on large scales, it is very good on small

Figure 2. Gradient of the lensed temperature T and polarization E fields as a function of the filtering scale, $\ell_{\max}$, as defined in Eq 2. We note that the RMS gradient is almost a factor of 10 larger in temperature than in polarization, and this will make temperature lensing more sensitive to a fixed-mass lens than polarization on small scales (in absence of foregrounds).

enough scales, for the following reason: CMB fluctuations have most of the power on large scales, small scales being suppressed by Silk damping. To quantify which scales contribute to the gradient, we can calculate the variance of the gradient when including multipoles ℓ only up to $\ell_{\max}$. For temperature, this is

$$G_{\mathrm{rms}}^2(<\ell_{\max}) = \int_0^{\ell_{\max}} \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \, \ell^2 C_\ell^{TT}. \qquad (2)$$

and a similar definition holds for polarization. Figure 2 shows G_{rms} as a function of $\ell_{\max}$, and we can see that the gradient variance entirely originates at $\ell \lesssim 2000$, as pointed out in Hu et al. (2007). If we consider small lenses such that $\nabla\phi$ receives most of its contribution from $\ell > 2000$, the gradient approximation should be excellent, and in this paper we will study this regime, which is where analytic progress can be made. For larger lenses, where this assumption fails, more expensive numerical maximum-likelihood methods should be employed to ensure optimality (Hirata & Seljak (2003a); Carron & Lewis (2017); Millea et al. (2017)).

As we have discussed in the introduction, the advantages of CMB lensing over galaxy lensing for measuring cluster masses become large at $z \gtrsim 1$, and we will show that most of those clusters are small enough in the sky for the gradient approximation to hold. This makes cluster lensing an ideal application of our formalism.

For concreteness, in the rest of the paper we will work in terms of the temperature fluctuations Θ, but the same formalism also applies to Q and U, since they are deflected by the same vector. Sometimes it will also be useful to

work in terms of the lensing convergence, κ, defined as $\kappa(\hat{\mathbf{n}}) = -\frac{1}{2}\nabla \cdot \mathbf{d}(\hat{\mathbf{n}}) = -\frac{1}{2}\nabla^2\phi(\hat{\mathbf{n}})$. We also note, that since the rms gradient of temperature fluctuations is almost an order of magnitude larger than that for polarization, temperature reconstruction is expected to be the dominant source of information on about small scales. Polarization, while statistically less powerful, is less affected by some of the foreground contamination and can provide useful consistency checks.

2.2 Quadratic Estimator in the small-scale limit and Optimal Estimators

CMB lensing introduces coupling between long and short wavelength modes and it is possible to construct a minimum-variance quadratic estimator for the lensing potential ϕ. Note that the "minimum-variance" qualification here only applies to the class of estimators that are quadratic in the observed fluctuations, and is not a general statement of optimality. It has been previously shown that maximum likelihood or iterative methods can outperform the QE. In this section we show that the QE is very suboptimal in the small-scale and low noise regime since it is subject to the cosmic variance on the gradient mode, while (in this regime) the particular realization of the gradient on the background of the lens can be measured without cosmic variance. In this section we will take the limit in which the reconstructed mode $L \gg \ell_{\mathrm{Silk}}$ (in practice $L \gtrsim 4000$), low noise and no-foreground limit.

The standard QE of Hu and Okamoto (Hu & Okamoto (2002)) for temperature can be written as[1]:

$$\hat{\phi}_{\mathrm{QE}}(\mathbf{L}) = N(\mathbf{L}) \int_{\boldsymbol{\ell}} \tilde{\Theta}(\boldsymbol{\ell})\tilde{\Theta}(\mathbf{L}-\boldsymbol{\ell})f(\boldsymbol{\ell},\mathbf{L}) \qquad (3)$$

where the mode-coupling kernel is

$$f(\boldsymbol{\ell},\mathbf{L}) = \frac{(\mathbf{L}-\boldsymbol{\ell})\cdot\mathbf{L}C_{|\mathbf{L}-\boldsymbol{\ell}|}^{TT} + \boldsymbol{\ell}\cdot\mathbf{L}C_\ell^{TT}}{2C_\ell^{\mathrm{tot}}C_{|\mathbf{L}-\boldsymbol{\ell}|}^{\mathrm{tot}}} \quad . \qquad (4)$$

Here, C_ℓ^{tot} includes contributions from instrumental noise, foregrounds, and the (lensed) primary CMB. The reconstruction noise serves as the normalization in the estimator and represents the uncertainty in the reconstruction of $\phi(\mathbf{L})$ due to chance correlations between different modes in an unlensed, Gaussian realization,

$$N(\mathbf{L})^{-1} = \int_{\boldsymbol{\ell}} \frac{\left[(\mathbf{L}-\boldsymbol{\ell})\cdot\mathbf{L}C_{|\mathbf{L}-\boldsymbol{\ell}|}^{TT} + \boldsymbol{\ell}\cdot\mathbf{L}C_\ell^{TT}\right]^2}{2C_\ell^{\mathrm{tot}}C_{|\mathbf{L}-\boldsymbol{\ell}|}^{\mathrm{tot}}} \quad . \qquad (5)$$

Taking the expectation value of this estimator over many CMB realizations for fixed ϕ, this estimator recovers an unbiased mean deflection field (at first order in the lensing expansion). However, this reconstruction has significant noise for any given CMB realization. This can be seen by taking the limit of reconstruction of small scale modes, with L higher than where the primary CMB C_ℓ^{TT} has support, i.e. on scales much smaller than the Silk damping scale,

---

[1] For compactness, we use the notation $\int_{\boldsymbol{\ell}} \equiv \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2}$. Upper-case $\mathbf{L}$ denotes lensing multipole, while lower-case $\boldsymbol{\ell}$ denotes temperature map multipole.

$L \gg \ell_{\rm Silk}$. Since we will use L for the reconstructed multipole and ℓ for the multipoles used in the reconstruction, we'll denote this limit by $L \gg \ell$. We then have:

$$f(\boldsymbol{\ell}, \mathbf{L}) \xrightarrow{L \gg \ell} 2 \frac{\boldsymbol{\ell} \cdot \mathbf{L} C_\ell^{TT}}{2 C_\ell^{\rm tot} C_L^{\rm tot}}, \tag{6}$$

and

$$N(\mathbf{L})^{-1} \xrightarrow{L \gg \ell} 2 \int_{\boldsymbol{\ell}} \frac{\left[\boldsymbol{\ell} \cdot \mathbf{L} C_\ell^{TT} \right]^2}{2 C_\ell^{\rm tot} C_L^{\rm tot}}. \tag{7}$$

Therefore the quadratic estimator simplifies to[2]

$$\hat{\phi}_{\rm QE}(\mathbf{L}) \longrightarrow \frac{\int_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \mathbf{L} \, W_F(\ell) \tilde{\Theta}(\boldsymbol{\ell}) \tilde{\Theta}(\mathbf{L})}{\int_{\boldsymbol{\ell}} (\boldsymbol{\ell} \cdot \mathbf{L})^2 \, W_F(\ell) C_\ell^{TT}}, \quad \text{with } W_F(\ell) = \frac{C_\ell^{TT}}{C_\ell^{\rm tot}}. \tag{8}$$

In the low-noise, no foregrounds limit, we can take the Wiener filter $W_F(\ell) \to 1$ and note that

$$\int_{\boldsymbol{\ell}} (\boldsymbol{\ell} \cdot \mathbf{L})^2 \, W_F(\ell) C_\ell^{TT} \approx \int_{\boldsymbol{\ell}} (\boldsymbol{\ell} \cdot \mathbf{L})^2 \, C_\ell^{TT} = L^2 \frac{1}{2} (\nabla \Theta)^2_{\rm rms}. \tag{9}$$

The quadratic estimator takes a simpler form of

$$\hat{\phi}_{\rm QE}(\mathbf{L}) \longrightarrow \frac{\tilde{\Theta}(\mathbf{L})}{L} \frac{\int_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \hat{\mathbf{n}}_{\mathbf{L}} \, \tilde{\Theta}(\boldsymbol{\ell})}{\frac{1}{2} (\nabla \Theta)^2_{\rm rms}}, \tag{10}$$

where we have written $\mathbf{L} = L \, \hat{\mathbf{n}}_{\mathbf{L}}$. Here $\int_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \hat{\mathbf{n}}_{\mathbf{L}} \, \tilde{\Theta}(\boldsymbol{\ell}) \approx (\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}$ is the (realization dependent) measured background gradient projected in the direction parallel to $\mathbf{L}$. We can then rewrite

$$\hat{\phi}_{\rm QE}(\mathbf{L}) \longrightarrow \frac{\tilde{\Theta}(\mathbf{L})}{L} \frac{(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}{\frac{1}{2} (\nabla \Theta)^2_{\rm rms}}. \tag{11}$$

This is the small-scale limit of the QE, which we will interpret soon.

On the other hand, at $L > 4000$, the primary CMB is highly suppressed by diffusion damping and in our limit, all small scale fluctuation are created by lensing. In this regime, we can write the small-scale fluctuations as

$$\tilde{\Theta}(\hat{\mathbf{n}}) \approx \nabla \phi \cdot \nabla \Theta(\hat{\mathbf{n}}). \tag{12}$$

Note that we can treat the locally measured gradient as constant on scales smaller than $1/\ell_G$, where $\ell_G \approx 2000$ is the multipole where the gradient becomes saturated as shown in figure 2. Then, there is a one-to-one correspondence between small scale lenses $\nabla \phi$ and measured fluctuations $\tilde{\Theta}$. In Fourier space,

$$\tilde{\Theta}(\mathbf{L}) = \int_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot (\mathbf{L} - \boldsymbol{\ell}) \Theta(\boldsymbol{\ell}) \phi(\mathbf{L} - \boldsymbol{\ell})$$
$$\xrightarrow{L \gg \ell} L \phi(\mathbf{L}) \int_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \hat{\mathbf{n}}_{\mathbf{L}} \, \tilde{\Theta}(\boldsymbol{\ell}) = L \phi(\mathbf{L}) \, (\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}. \tag{13}$$

We can call the solution of the previous equation for $\phi(\mathbf{L})$ the "Gradient Inversion" (GI) solution:[3]

$$\hat{\phi}_{\rm GI}(\mathbf{L}) \longrightarrow \frac{\tilde{\Theta}(\mathbf{L})}{L} \frac{1}{(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}. \tag{14}$$

Note that modes with $\mathbf{L}$ perpendicular to the gradient direction have $(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}} = 0$ and cannot be reconstructed. This is because in this limit, lensing doesn't produce any effect perpendicular to the gradient direction and therefore no estimator can reconstruct such modes.

We can now compare the two estimators:

$$\begin{aligned} L \, \hat{\phi}_{\rm QE}(\mathbf{L}) &\longrightarrow \tilde{\Theta}(\mathbf{L}) \frac{(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}{\frac{1}{2} (\nabla \Theta)^2_{\rm rms}}, \\ L \, \hat{\phi}_{\rm GI}(\mathbf{L}) &\longrightarrow \tilde{\Theta}(\mathbf{L}) \frac{1}{(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}. \end{aligned} \tag{15}$$

Given a good enough experiment, $\tilde{\Theta}(\mathbf{L})$ and $(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}$ can be made measured to arbitrary accuracy, and therefore the error on $\hat{\phi}_{\rm GI}(\mathbf{L})$ can be made arbitrarily small. Note that in this limit neither $\tilde{\Theta}(\mathbf{L})$ nor $(\nabla \Theta)_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}$ are random variables, and should therefore not be marginalized over.

This is not the case for the quadratic estimator, which can be rewritten as (in the limit of small noise), and identifying $\hat{\phi}_{\rm GI}$ with the "true" lensing potential $\phi_{\rm true}$,

$$\hat{\phi}_{\rm QE}(\mathbf{L}) = \hat{\phi}_{\rm GI}(\mathbf{L}) \frac{(\nabla \Theta)^2_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}{\frac{1}{2} (\nabla \Theta)^2_{\rm rms}} \longrightarrow \phi_{\rm true}(\mathbf{L}) \frac{(\nabla \Theta)^2_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}{\frac{1}{2} (\nabla \Theta)^2_{\rm rms}} \tag{16}$$

Firstly, we notice that $\langle (\nabla \Theta)^2_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}} \rangle = \frac{1}{2} (\nabla \Theta)^2_{\rm rms}$, so that $\langle \hat{\phi}_{\rm QE} \rangle = \phi_{\rm true}$, where the expectation value is taken over realizations of the background gradient, for a fixed ϕ. We see that while unbiased, the "error" that the QE makes is proportional to the difference between the (square) true gradient and the rms gradient. Since this quantity is on average of order the rms gradient itself, the fractional error of the QE is always of order unity per mode.

We can formalize the above intuition, showing that even with arbitrarily small experimental noise, there is a lower limit to the statistical error on the quadratic estimator $\sigma(\hat{\phi}_{\rm QE})$. Defining

$$R_\nabla^2 = \frac{(\nabla \Theta)^2_{\rm true, \hat{\mathbf{n}}_{\mathbf{L}}}}{\frac{1}{2} (\nabla \Theta)^2_{\rm rms}}, \tag{17}$$

we find that

$$\left(\frac{\sigma^2(\hat{\phi}_{\rm QE})}{\phi^2_{\rm true}} \right)_{\rm min} (\mathbf{L}) = \langle (R_\nabla^2)^2 \rangle - \langle R_\nabla^2 \rangle^2 = 3 - 1 = 2 \tag{18}$$

In summary, the GI estimator can have arbitrarily low noise per mode for a good enough experiment, while the QE is limited to $S/N = 1/\sqrt{2}$ per mode (which is the same has having cosmic variance on the gradient mode).

---

[2] Technically, the convergence in the following equation would only be "in probability" and not at the field level, if $\Theta(\boldsymbol{\ell})$ and $\phi(\boldsymbol{\ell})$ were random variables with values that are uncorrelated for different $\boldsymbol{\ell}$. In practice, the finite (and small) window function used to define the local gradient and nonlinear evolution of the potential make both continuous functions and the limit well defined. This subtlety doesn't affect any of our results.

[3] The GI solution corresponds to the maximum likelihood solution in the limit considered here, where both the gradient and the short scale modes can be measured on this particular realization with $S/N \gg 1$. This can be seen by writing a likelihood based on Equation 12 and taking the noise to zero.
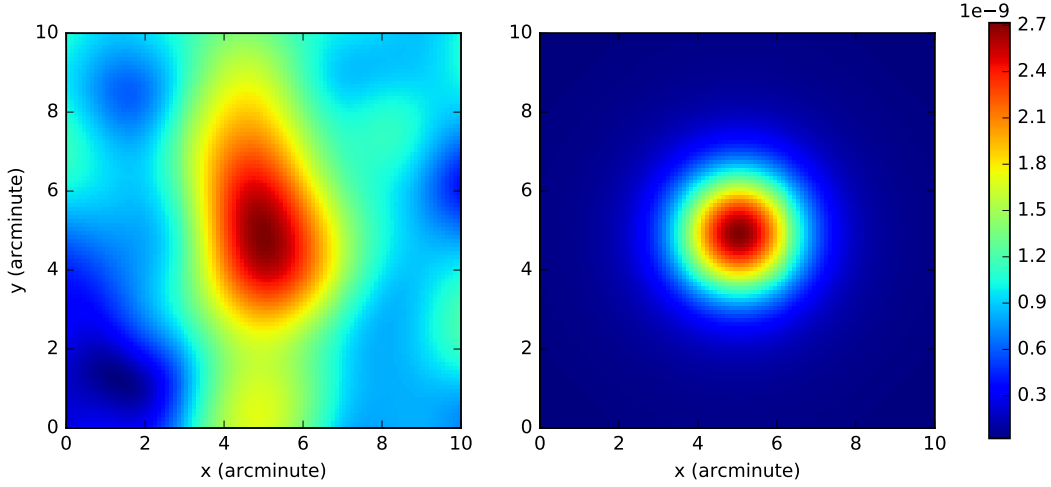
Figure 3. Left: Reconstruction of the convergence map of a single cluster in a constant gradient in the $\hat{y}$ direction, assuming 1 arcmin beam and noise level of 1 μK-arcmin. Cluster mass of $2 \times 10^{14} M_\odot$ at $z = 2.0$ with an NFW profile (see Section 3.1). Note that modes perpendicular to the gradient are not reconstructed in this limit, hence the residual vertical band in the direction of the gradient. Right: True convergence of the same cluster.

2.3 One gradient, many independent modes?

In this section we explore the covariance of the small scale lensing modes estimated using the quadratic estimator, in the very high-L, very low noise limit. In this regime, we have shown that

$$\hat{\phi}_{\mathrm{QE}}(\mathbf{L}) \longrightarrow \phi_{\mathrm{true}}(\mathbf{L}) \; \frac{(\nabla\Theta)^2_{\mathrm{true},\hat{\mathbf{n}}_{\mathbf{L}}}}{\frac{1}{2}(\nabla\Theta)^2_{\mathrm{rms}}} = \phi_{\mathrm{true}}(\mathbf{L}) \; R^2_\nabla \quad (19)$$

Then the noise covariance between different small scale modes sharing the same background gradient becomes:

$$N^{\phi\phi}_{\mathrm{QE}}(\mathbf{L},\mathbf{L}') \equiv \langle (\hat{\phi}_{\mathrm{QE}}(\mathbf{L}) - \phi_{\mathrm{true}}(\mathbf{L})) \, (\hat{\phi}_{\mathrm{QE}}(\mathbf{L}') - \phi_{\mathrm{true}}(\mathbf{L}')) \rangle_\Theta \tag{20}$$

where crucially the average is taken over realization of the long-wavelength CMB fluctuations, for a fixed large scale structure ϕ_{true}. we have

$$\begin{aligned} N^{\phi\phi}_{\mathrm{QE}}(\mathbf{L},\mathbf{L}') &\longrightarrow \phi_{\mathrm{true}}(\mathbf{L})\phi_{\mathrm{true}}(\mathbf{L}')\langle (R^2_\nabla - 1)^2 \rangle \\ &= 2 \, \phi_{\mathrm{true}}(\mathbf{L})\phi_{\mathrm{true}}(\mathbf{L}') \end{aligned} \tag{21}$$

where we have used $\langle (R^2_\nabla)^2 \rangle = 3$ and $\langle R^2_\nabla \rangle = 1$.

We conclude that for small scale modes not only the noise covariance matrix $N^{\phi\phi}_{\mathrm{QE}}(\mathbf{L},\mathbf{L}')$ is not diagonal, but the off-diagonal terms become as large as the on-diagonal ones, making the matrix singular. This means that when detector noise is negligible, modes sharing the same gradient become 100% percent correlated, and treating them as independent (for example when forecasting cluster lensing constraints for future surveys), may lead to underestimation of the overall statistical noise.

We also note that this is a feature of the quadratic estimator, and won't be a limiting factor in a maximum-likelihood approach that includes information about the large-scale gradient. In principle, a maximum likelihood approach has no limit on the statistical significance that it can achieve[4].

2.4 Example: Lensing Map Reconstruction

In the previous section, we have derived a ϕ estimator for the pure gradient, low noise limit, which is found in Eq 14. Equivalently, we can rephrase that in terms of the lensing convergence $\kappa(\mathbf{L}) = -\frac{L^2}{2}\phi(\mathbf{L})$,

$$\hat{\kappa}_{\mathrm{GI}}(\mathbf{L}) \longrightarrow -\frac{L\tilde{\Theta}(\mathbf{L})}{2} \; \frac{1}{(\nabla\Theta)_{\mathrm{true},\hat{\mathbf{n}}_{\mathbf{L}}}} \; . \tag{22}$$

In Figure 3, we show a comparison between the true and reconstructed convergence field, in the case of a massive lens with mass $2 \times 10^{14} M_\odot$ and following an NFW profile, as described in the next section. We note that as expected, the modes with variation perpendicular to the gradient direction are not correctly reconstructed, explaining the vertical band in the reconstructed map. Those modes can be filtered out or downweighted in a real analysis since they have infinite variance.

In presence of finite noise, the error on the estimate of a given $\kappa(\mathbf{L})$ noise, depends on the errors on both the large-scale gradient $(\nabla\Theta)_{\mathrm{true},\hat{\mathbf{n}}_{\mathbf{L}}}$ and the small scale temperature fluctuation $\tilde{\Theta}(\mathbf{L})$.

We expect the error on the gradient to be almost negligible in most cases, unless the gradient on the patch of interest happens to be much smaller than the rms, since most current or future CMB experiments should be close to

---

[4] Further neglecting post-Born corrections and field rotation, as well as foregrounds and residual primary CMB.

cosmic variance limited at $\ell \lesssim 2000$,

$$\left(\frac{\sigma(\hat{\kappa}_{\mathrm{GI}})}{\kappa}\right)^2 = \left(\frac{\sigma(\tilde{\Theta})}{\tilde{\Theta}}\right)^2 + \left(\frac{\sigma((\nabla\Theta)_{\mathrm{true},\hat{\boldsymbol{n}}_{\mathbf{L}}})}{(\nabla\Theta)_{\mathrm{true},\hat{\boldsymbol{n}}_{\mathbf{L}}}}\right)^2. \quad (23)$$

Assuming that the errors are on the gradient and the small scales are independent. Note that the noise properties of $\hat{\kappa}_{\mathrm{GI}}$ are in general not trivial, since the noise is anisotropic (it depends on the angle between the gradient and **L**), and the direction of the anisotropy depends on position on the sky. We defer a full treatment to future work, but we will discuss noise estimates in the specific case of cluster lensing in what follows.

3   MEASURING CLUSTER MASSES WITH MATCHED FILTERING

A straightforward application arises when the lens profile is known up to a normalization factor, such as when determining the mass M of a cluster, lying on a constant gradient $\boldsymbol{G} = (\nabla\Theta)_{\mathrm{true}}$. We have shown that the on small scales, the GI solution is linear in the measured temperature fluctuations, and therefore the minimum variance estimator for the overall amplitude of the lensing $A = M|G|$, can be obtained by a matched filter of the CMB map (Haehnelt & Tegmark (1996)). Once both the amplitude A and the background gradient $\boldsymbol{G}$ have been measured, the mass can be obtained by direct inversion $M \approx A/|G|$. In the context of cluster lensing, this was first investigated by (Seljak & Zaldarriaga (2000)), and here we revisit this point.

A mass of a given profile on a pure gradient background CMB will create a dipole-like structure $g(\boldsymbol{\theta})$, aligned with the gradient direction, and with a free amplitude A that we wish to measure. We then model the total small-scale CMB anisotropy as

$$\tilde{\Theta}(\boldsymbol{\theta}) = Ag(\boldsymbol{\theta}) + n(\boldsymbol{\theta}) \quad (24)$$

where $g(\boldsymbol{\theta})$ is the angular profile of the deflection angle of a known profile (i.e. NFW, Sersic, etc.) caused by the cluster, A is an amplitude depending on the mass of the cluster and the CMB gradient, and $n(\boldsymbol{\theta})$ is noise, either from the instrument, foregrounds, or residual primary CMB. It is useful to work with the noise in Fourier space and its power spectrum,

$$\langle n(\boldsymbol{\ell})\, n(\boldsymbol{\ell}')\rangle = (2\pi)^2 C_\ell^{\mathrm{tot}}\delta(\boldsymbol{\ell} + \boldsymbol{\ell}'). \quad (25)$$

The value of A can be found by applying a linear filter Ψ on the temperature map at the cluster position

$$\hat{A} = \int d^2\boldsymbol{\theta}\ \Psi(\boldsymbol{\theta})\tilde{\Theta}(\boldsymbol{\theta}), \quad (26)$$

where the optimal filtering function is given by (Haehnelt & Tegmark (1996)),

$$\Psi(\boldsymbol{\ell}) = \left[\int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \frac{|g(\boldsymbol{\ell})|^2}{C_\ell^{\mathrm{tot}}}\right]^{-1} \frac{g(\boldsymbol{\ell})}{C_\ell^{\mathrm{tot}}}, \quad (27)$$

can be found by minimizing the error on $\hat{A}$. We show components of our matched filter for a NFW profile in Figure 4.

The variance on A is given by

$$\sigma_A^2 = \left[\int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \frac{|g(\boldsymbol{\ell})|^2}{C_\ell^{\mathrm{tot}}}\right]^{-1} \quad (28)$$
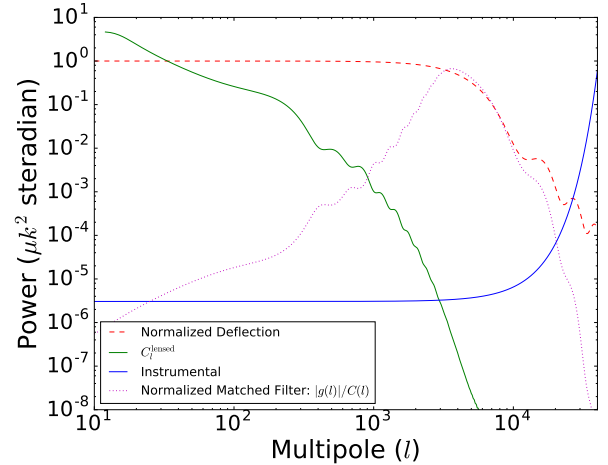
Figure 4. Components of matched filter in ℓ-space. At low ℓ the cluster power will be suppressed by the primary CMB fluctuations and at high ℓ it will be suppressed by the instrumental beam. Instrumental noise is for a 6 μK-arcmin experiment with a 1 arcmin beam.

Assuming an axi-symmetric cluster profile we can write

$$g(\boldsymbol{\ell}) = g(\ell)\cos\phi_\ell \quad (29)$$

where ϕ_ℓ is the angle between the vector $\boldsymbol{\ell}$ and the gradient direction.

Note that $\sigma_A \to 0$ as the noise power spectrum tends to zero, which is what we expect.

3.1   NFW Profile

For our work, we will assume that halos follow a NFW profile Navarro et al. (1997), with density given by

$$\rho(r) \propto \frac{1}{r/r_s(1 + r/r_s)^2}, \quad (30)$$

with scale radius r_s and normalization constant dependent on halo mass. The halo mass is related to the scale radius via the concentration parameter, c, which we will fix to 3.2 to allow direct comparison against Hu et al. (2007). In practice, variation of the concentration and scale radius can be corrected for by combining CMB lensing information with other proxies for density, such as tSZ and X-ray observations. The lensing deflection profile (with the speed of light equal to one) for an NFW halo is

$$|\nabla\phi(\theta)| \propto g(\theta) = 16\pi GM\frac{D_{LS}D_L}{D_S}\frac{\rho_s r_s \theta_s}{\theta}h\left(\frac{\theta}{\theta_s}\right) \quad (31)$$

where,

$$\rho_s = \frac{200}{3}\rho_{\mathrm{crit}}\frac{c^3}{\ln c(1+c) - c/(1+c)} \quad (32)$$

and D_L is the comoving distance to the lensing, D_S is the comoving distance to the source (i.e. the CMB), D_{LS} is the comoving distance between the lens and the source, θ_s is the angle subtended by the scale radius and M is the cluster

mass. The functional dependence of the profile is

$$
h(x) = \begin{cases} \ln(x/2) + \frac{2}{\sqrt{x^2-1}}\arctan\sqrt{\frac{x-1}{x+1}}, & (x > 1) \\ \ln(x/2) + \frac{2}{\sqrt{1-x^2}}\arctan\sqrt{\frac{1-x}{1+x}}, & (x < 1) \\ \ln(x/2) + 1\,, & (x = 1)\,. \end{cases} \quad (33)
$$

4 IDEALIZED EXAMPLE ON CMB MAP

To demonstrate this technique, we generated CMB realizations using HEALPix (Górski et al. (2005)[5]), with $N_{\text{side}} = 2048$, from which we extracted $20'$ square cutouts. The effect of lensing from an NFW profile with a given concentration and mass is then added, together with detector noise, as shown in Figure 1. We then reconstruct the mass using the matched filter technique described above, the first step being a measurement of the average the background gradient, as discussed below.

4.1 Reconstruction of the background gradient

Correct determination of the cluster mass depends on being able to accurately extract the mean background gradient Holder & Kosowsky (2004), without bias from the cluster lensing signal or deviation from the pure gradient approximation. Fortunately, as discussed in section 2.1, this gradient has little variation on cluster scales. We can define the average gradient on a small patch centered on the cluster i as $\boldsymbol{G}_i = [(\nabla\Theta)_{\text{true}}]_{\text{patch } i}$, so that

$$
\boldsymbol{G}_i = \frac{1}{C} \int_{\ell<2000} \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \left[\boldsymbol{\ell}\,\tilde{\Theta}(\boldsymbol{\ell})\right]_{\text{patch } i} \quad (34)
$$

where C is proportional to the area being integrated over. One should be careful to not average over too large of a patch such that variations in the primary CMB would become a concern. For our analysis we find that a $8'$ box around the center of the cluster works adequately.

This measure is robust even in the presence of noise, resulting in less than a 0.1% error in gradient extraction for an experiment with a 1 arcmin beam and 1 μK-arcmin sensitivity, compared to the ideal case. The error scales roughly linearly with instrumental sensitivity and has very small dependence on lensing halo mass and redshift.

In addition, the presence of clusters does not strongly bias the measurement of the gradient itself. This shown explicitly in Figure 5, where we simulated the lensing of the CMB by massive clusters, extracted the gradient from those lensed images, and applied the matched filter prescription to measure the mass. As shown, there is no significant bias in the measured mass as it relates to the real mass.

4.2 Measuring the lens mass

Here we briefly summarize the procedure used for estimating the mass of simulated clusters. For a given simulated lensed image, we perform the following procedure:

(i) Find the direction and magnitude of the gradient $\boldsymbol{G}_i$ at the center of the cluster, as in Equation 34.
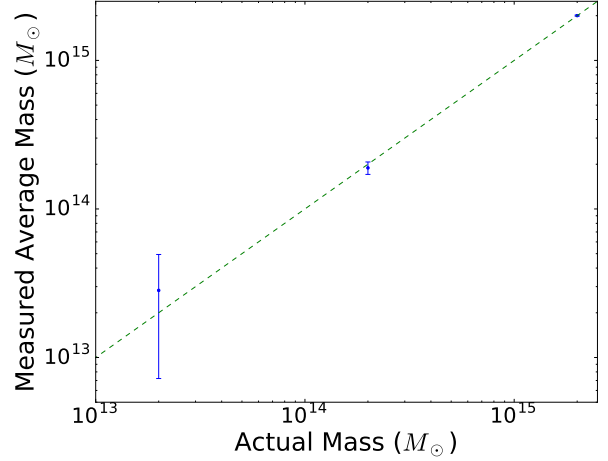
Figure 5. Cluster mass reconstruction as a function of mass for our 1000 realistic lensed CMB realizations. We have fixed $z = 0.7$, $c = 3.2$, $\Delta_T = 6\mu K$-arcmin, and beam FWHM = 1 arcmin. Our reconstructed mass has no measurable bias.

(ii) Choose the axis of the matched filter (i.e. the $\phi_\ell = 0$ direction), aligned with the direction of the gradient. This is important due to the $\cos\phi_\ell$ term in Equation 29 establishing the antisymmetry of the filter in real space.

(iii) Apply the matched filter defined in Equation 27 and perform the integral in Equation 26 to obtain an amplitude $\hat{A}_i$.

(iv) Since $A_i = |G_i|M_i$ when correctly normalized, the mass M_i can be estimated as $\hat{M}_i = \hat{A}_i/|G_i|$, with error dependent on the local size of the gradient as expected: $\sigma(M_i) = \sigma_A/|G_i|$. For a sample of many different clusters, the mean mass can be obtained by inverse noise weighting each measurement as explained in the next section.

4.3 Gradient-based weighting of samples

Assume that we have a collection of N clusters all of the same mass, M, but each in their own background gradient, G_i, which we assume is extracted with no noise from observations. Then the matched filter output for each cluster i, $\hat{A}_i$ (from Equation 26), and it's uncertainty, σ_A (which is approximately independent of the cluster mass in a uniform sample, and only determined by detector noise and foregrounds), are given by

$$
\hat{A}_i = |G_i|M_i, \quad (35)
$$
$$
\sigma_A = |G_i|\sigma(M_i). \quad (36)
$$

The minimum variance estimate of the mean mass is obtained by inverse noise weighing individual measurements. Noting that from Equation 36, $\sigma(M_i)$ is inversely proportional to $|G_i|$, we have

$$
\hat{M} = \frac{\sum_i \frac{|G_i|\hat{A}_i}{\sigma_A^2}}{\sum_i \frac{|G_i|^2}{\sigma_A^2}}. \quad (37)
$$

Where the uncertainty on the mean mass is

$$
\sigma^2(\hat{M}) = \left[\sum_i \frac{|G_i|^2}{\sigma_A^2}\right]^{-1} = \frac{1}{N}\frac{\sigma_A^2}{G_{\text{rms}}^2}\,, \quad (38)
$$

The latter equality being true in the limit of a large sample size. Equation 38, together with the matched filter error σ_A given by Equation 28, provides the basis for our results in Figure 6. Note that uniform weighting would lead to an increase of the mass uncertainty.

4.4   Results

In Figure 6 we show the analytical calculation of the mass uncertainty using the formalism of Equation 28, allowing the integral to be cut off by the beam (see Figure 4). Our fiducial case include the residual *lensed* CMB as a source of noise in C_ℓ^{tot}, while we also show the case in which the unlensed CMB is used instead. In practice, the fiducial case is conservative in the sense that no delensing is assumed, while any amount of delensing (either internal (Carron et al. (2017)) or with the use of other tracers (Smith et al. (2012))) will place the noise between our two extreme cases.

For comparison, we show the performance of the quadratic estimator (QE) of Hu et al. (2007), as well as the maximum likelihood (ML) result of Raghunathan et al. (2017). For the QE curve we integrate over ℓ well past the beam cutoff scale ($\ell \approx 10,000$) for a fair comparison between the methods. The matched filter outperforms the quadratic estimator in the low noise limit as well as the high redshift limit. For a cluster of a constant mass, its apparent size decreases with increasing redshift and the background gradient is closer to a pure gradient where the match filter technique is optimal.

As expected by the analytical calculations in Section 2.2, the quadratic estimator performance saturates at signal to noise ratio $S/N \sim 1$ per cluster. Meanwhile, the matched filter keeps improving for small values of Δ_T.

As a check, we validate our error estimate on simulations. To do this, we have taken 1,000 cutouts from full-sky realizations of the lensed primary CMB, as described in Section 4 and placed an NFW cluster with constant mass and $c = 3.2$ at the center of the image and lensed the background primary CMB. We then measured the gradient as described in Section 4.1, and found their mass using the procedure outlined in Section 4.2. We then weight the clusters according to their background gradient, as described in 4.3, to find the average measured mass; the results of which are shown in Figure 6.

4.5   Effect of Beam Size

We show the effect of beam size in Figure 7. As beam size decreases, less of the power from the deflection angle is cut off (see Figure 4), and there is a significant improvement when going to higher resolution, up to about 1 arcmin beam (FWHM) for a $2 \times 10^{14} M_\odot$ cluster at $z = 0.7$. However, when the cluster is well resolved, the improvements become marginal. Smaller beam size may be useful in extracting other information from the cluster, such as concentration or other details about the profile.

4.6   Effect of Foreground Emission and other Secondary Anisotropies

Here we briefly comment on the impact of foregrounds on the matched filter estimator and we leave a full exploration with simulation and correlated emission to future work. We note that we expect many of the issues outlined in the detailed study of Raghunathan et al. (2017) in the context of a maximum likelihood approach to be relevant here as well, with some caveats that we now explain.

Firstly, the effects of residual foreground emissions uncorrelated with the cluster in question and with known power spectra can be incorporated in this formalism as an additional source of noise contribution in Equation 25. In this case, the noise appearing in the plots would refer to "effective noise" after component separation, and we note that most of foregrounds can be in part mitigated by a multifrequency analysis.

An exception is the kinematic Sunyaev-Zel'dovich (kSZ) effect caused by the bulk motion of free electrons (Sunyaev & Zeldovich (1970)). Unlike other foregrounds, the kSZ effect cannot be subtracted out by using multifrequency observations as it preserves the black body spectrum of the CMB, and therefore it represents an additional source of noise that is hard to overcome. As the measurements of the kSZ effect improve, it may be possible to construct a template of the emission and reduce the residual kSZ noise by a factor of order unity.

Any emission from the cluster itself, or from matter correlated with it should also only amount to extra noise, and not a bias, unless this emission is correlated with our estimate of the background gradient of the unlensed CMB. This is because the matched filter only detects dipole contributions aligned with the background gradient, due to the $\cos \phi_\ell$ term in Equation 29.

While not technically a foreground, gravitational lensing from uncorrelated structures in the universe produce extra small-scale anisotropy which again acts as a source of noise, which is included in the fiducial analysis throughout the paper. In Figure 6, we show that the difference in error caused by using the lensed vs unlensed CMB power spectra is only relevant at small noise levels. The effect of matter correlated with the halo is to enhance the lensing signal by what is usually known as the "two halo term". In a real analysis, the whole matter profile would be obtained and the one and two halo contributions would be fit simultaneously, just as performed in most galaxy-galaxy lensing analyses.

5   RELATION TO PREVIOUS WORK

Here we briefly compare the match filtered technique with other work in literature. Direct matched filtering techniques were first discussed as applicable to CMB cluster lensing in (Seljak & Zaldarriaga (2000)), and extensions in (Holder & Kosowsky (2004); Vale et al. (2004); Dodelson (2004)). Most of the subsequent work has focused on the use of quadratic estimators, as modified by (Hu et al. (2007)) to ensure unbiased results (Melin & Bartlett (2015); Maturi et al. (2005)) This and other works based on the QE suffer from the statistical limitations discussed in our paper.

Iterative techniques based on the quadratic estimator
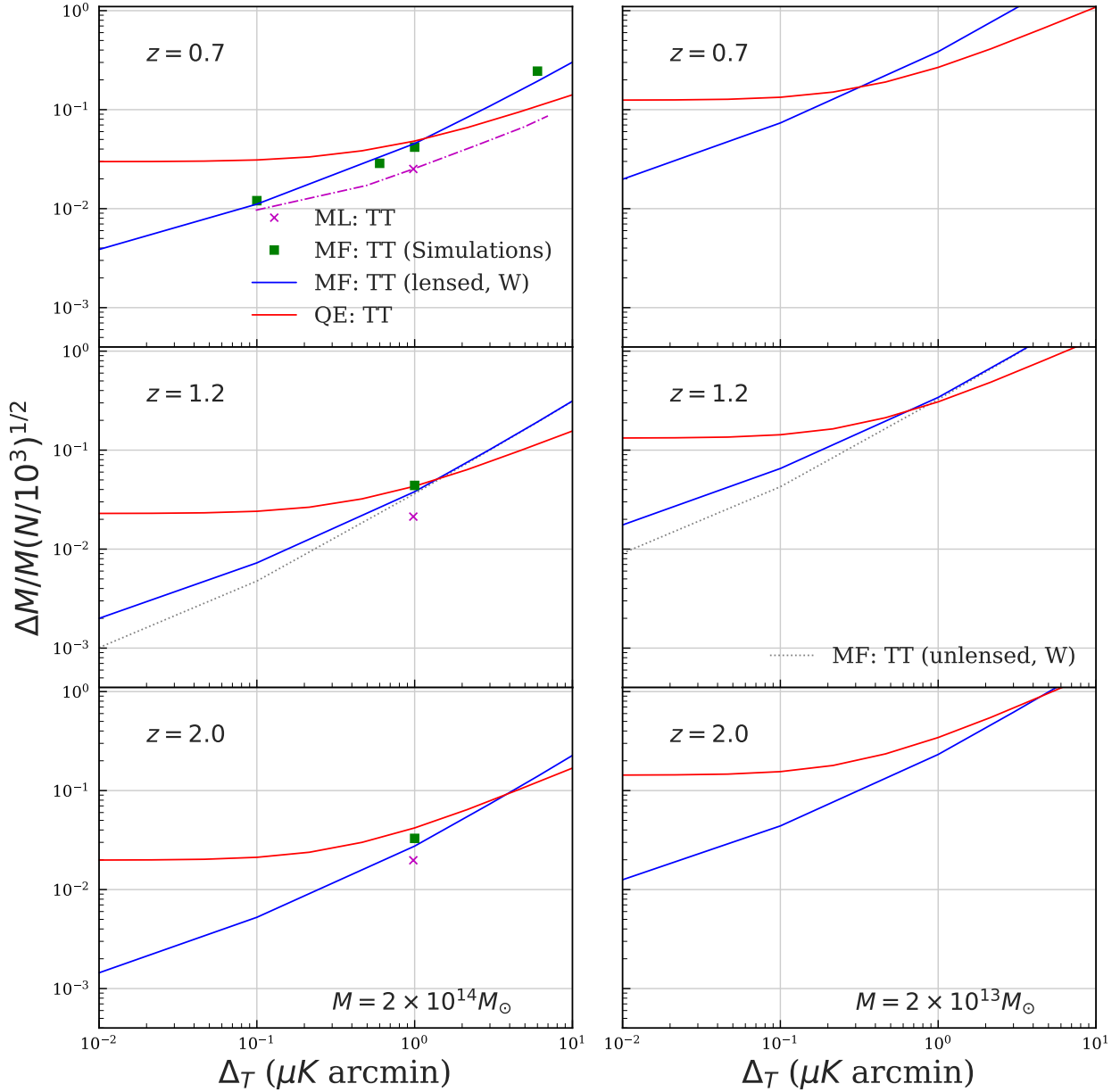
Figure 6. Mass sensitivity assuming 1 arcmin beam and no foregrounds using different techniques for redshifts $z = 0.3, 0.7$, and 1.2, as well as for two different masses, $2 \times 10^{14} M_\odot$ (left) and $2 \times 10^{13} M_\odot$ (right). In addition to the matched filter result from this work (MF) we show the quadratic estimator (QE) from Hu et al. (2007), which has an essentially flat redshift dependence at $z > 0.5$ for a 1 arcmin beam (see also Figure 3 in Melin & Bartlett (2015)). We have also shown the maximum likelihood (ML) TT result from Raghunathan et al. (2017) (see Figure 2). In the middle panel, we show the effect of using the lensed vs unlensed CMB as a source of noise. For all ℓ integrals we use an $\ell_{\min} = 10$ and $\ell_{\max} = 40,000$ (beyond the beam cutoff scale). The QE outperforms the MF for large values of the noise, since in this regime most of the contribution to the S/N will come from modes with $\ell < 4000$, where the MF is no longer optimal.

have been introduced in (Yoo & Zaldarriaga (2008); Yoo et al. (2010)). In the limit of large number of iterations the result should converge to the maximum likelihood solution in a unbiased way.

A maximum likelihood framework for cluster masses was developed in (Lewis & King (2006); Baxter et al. (2015)) and further studied and extended to polarization in (Raghunathan et al. (2017)), which also provides extensive discussion of possible contamination from foregrounds. While it is

not possible to outperform a full maximum likelihood approach, Figure 6, shows that for high redshift clusters or small noise limit, the matched filter presented in this paper approaches the maximum likelihood performance, but with only a small fraction of the computational cost.

On the experimental front, cluster lensing has first been detected using the quadratic estimator by (Madhavacheril et al. (2015)) using data from the ACTPol experiment and CMASS galaxies from the BOSS survey. Subsequently, it
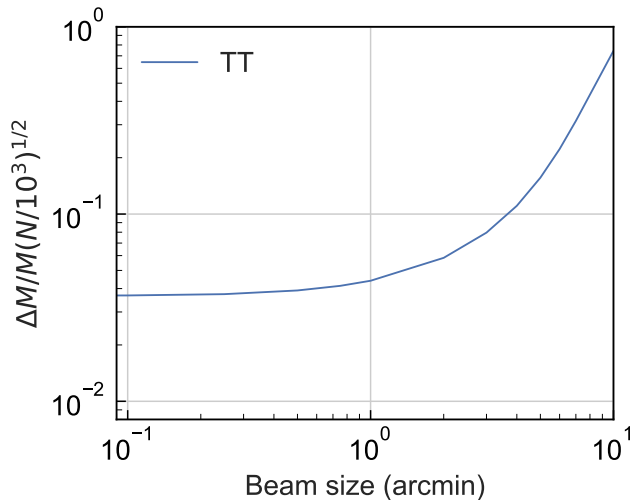
Figure 7. Effect of beam size on ability to extract mass, assuming 1 μK-arcmin instrumental noise, a $2 \times 10^{14} M_{\odot}$ cluster at $z = 0.7$, and a $N = 1000$ sample. At 0.3 arcmin effectively all the power of an NFW cluster is captured in the matched filter and no more improvements are possible by decreasing beam size.

has also been detected by the SPT collaboration using a Maximum-Likelihood technique (Baxter et al. (2015, 2017)), as well as by the Planck collaboration (Ade et al. (2016)).

Recently there has been work towards implementing a full maximum likelihood lensing map reconstruction without any assumptions about the shape of the lensing potential (Carron & Lewis (2017)), based on previous work (Hirata & Seljak (2003b)). Going forward, further developing these numerical techniques and applying them to small scale lenses would be very useful, but it currently appears challenging to obtain rapid convergence on small scales.

6 SUMMARY

In this work we have explored the limitations of quadratic estimator reconstruction of CMB lensing in the small scale, low noise limit. We have shown that the quadratic estimator is fundamentally limited by suboptimal weighting, while in no-foregrounds, small noise limit, small lenses should be measurable arbitrarily well. This problem can be overcome by using different techniques, such as direct gradient inversion presented in this paper, or using more computationally expensive maximum likelihood methods.

An obvious application of the direct gradient inversion is reconstruction of cluster masses from CMB lensing. Once a profile for the potential is assumed, optimal weighting of the reconstructed modes is equivalent to a matched filter that we have explored in this work and which out-performs the quadratic estimator over a wide range of instrumental sensitivities and redshifts; focused primarily on the low noise and small angular scale limit. The matched filtered approach relies on the antisymmetric nature of lensing as well as the separation of scales between cluster-lensing and the variation in the primary CMB. The antisymmetric nature of the matched filter is a useful feature which allows it to be less

sensitive to kSZ and other foregrounds coming from the cluster.

The matched filter studied here can be extended to fit additional parameters such as the cluster concentration (see for example (Dodelson (2004))) or the amplitude of the two-halo term. Alternatively, tight priors can be put on them by studying the halo clustering, or detailed tSZ/X-ray observations. The effect of miscentering between the observed and true halo center needs to be quantified since it will lead to a bias if not properly accounted for, as pointed out by (Raghunathan et al. (2017)). At the same time, correlation between the selection function and other observables can produce biases in the inferred mean mass for the sample. For example, if halos elongated along the line of sight are preferentially selected, the inferred mass is likely to be overestimated due to the alignment of the halo with the surrounding cosmic structure. Detailed simulations and modeling are required to estimate the size of these effects for the particular sample being analyzed.

The technique discussed in this paper approaches the maximum likelihood sensitivity for the noise level expected of future CMB experiments, and is computationally very efficient. Moreover, the noise calculation is analytically tractable, allowing direct forecasting and Fisher-type analysis without having to simulate a large number of clusters.

References

Abazajian K. N., et al., 2016, arXiv preprint arXiv:1610.02743
Ade P. A. R., et al., 2016, Astron. Astrophys., 594, A24
Baxter E., et al., 2015, The Astrophysical Journal, 806, 247
Baxter E. J., et al., 2017
Carron J., Lewis A., 2017, arXiv preprint arXiv:1704.08230
Carron J., Lewis A., Challinor A., 2017, Journal of Cosmology and Astroparticle Physics, 2017, 035
Collaboration P., et al., 2013, ASTRONOMY and ASTROPHYSICS, 571
Dodelson S., 2004, Phys. Rev., D70, 023009
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Haehnelt M. G., Tegmark M., 1996, Monthly Notices of the Royal Astronomical Society, 279, 545
Hirata C. M., Seljak U., 2003a, Physical Review D, 67, 043001
Hirata C. M., Seljak U., 2003b, Physical Review D, 68, 083002
Holder G., Kosowsky A., 2004, The Astrophysical Journal, 616, 8

Hu W., Okamoto T., 2002, The Astrophysical Journal, 574, 566

Hu W., DeDeo S., Vale C., 2007, New Journal of Physics, 9, 441

Lewis A., Challinor A., 2006, Physics Reports, 429, 1

Lewis A., King L., 2006, Physical Review D, 73, 063006

Louis T., Alonso D., 2016, arXiv preprint arXiv:1609.03997

Madhavacheril M., et al., 2015, Physical review letters, 114, 151302

Madhavacheril M., Battaglia N., Miyatake H., 2017

Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, Monthly Notices of the Royal Astronomical Society, 440, 2077

Maturi M., Bartelmann M., Meneghetti M., Moscardini L., 2005, Astron. Astrophys., 436, 37

Melin J.-B., Bartlett J. G., 2015, Astronomy & Astrophysics, 578, A21

Millea M., Anderes E., Wandelt B. D., 2017

Navarro J. F., Frenk C. S., White S. D., 1997, The Astrophysical Journal, 490, 493

Raghunathan S., et al., 2017, arXiv preprint arXiv:1705.00411

Rozo E., Rykoff E. S., Bartlett J. G., Evrard A. E., 2013, arXiv preprint arXiv:1302.5086

Seljak U., 1995, arXiv preprint astro-ph/9505109

Seljak U., Zaldarriaga M., 2000, The Astrophysical Journal, 538, 57

Smith K. M., Hanson D., LoVerde M., Hirata C. M., Zahn O., 2012, Journal of Cosmology and Astroparticle Physics, 2012, 014

Sunyaev R. A., Zeldovich Ya. B., 1970, Astrophys. Space Sci., 7, 3

Vale C., Amblard A., White M., 2004, New Astronomy, 10, 1

Yoo J., Zaldarriaga M., 2008, Physical Review D, 78, 083002

Yoo J., Zaldarriaga M., Hernquist L., 2010, Physical Review D, 81, 123006

Zaldarriaga M., Seljak U., 1998, Physical Review D, 58, 023003

von der Linden A., et al., 2014, Monthly Notices of the Royal Astronomical Society, 439, 2