Quick and Dirty Point-source Polarization Estimate from 500d SPTpol Data

Tom Crawford

January 26, 2016

Abstract

For the purposes of forecasting and design for future CMB experiments, we want to know the mean-squared polarization fraction of mJy-level sources at CMB frequencies (such as 150 GHz and 100 GHz), but we haven't done a dedicated polarized point-source analysis of SPTpol data yet. Here I take some 150 GHz maps not optimized for polarized point-source finding and do it anyway. After accounting for some biases, I find a mean-squared polarization fraction of $\langle \hat{p}^2 \rangle = (1.42 \pm 0.15) \times 10^{-3}$ for sources above 25 mJy and $\langle \hat{p}^2 \rangle = (1.42 \pm 0.67) \times 10^{-3}$ for sources between 6.7 mJy and 25 mJy. Both of these subsamples are expected to consist of $> 90\%$ radio sources.

1 Intro

For several purposes, the most pressing being the imminent creation of the CMB-S4 Science Book and the need to have a number about point-source polarization to put in it, we want to know the polarization fraction of point sources at CMB wavelengths. The current state of this knowledge in the literature is sketchy. In SPT publications, we often cite Battye et al. (2011) and Seiffert et al. (2007) as evidence for a low (few % level) polarization fraction for radio sources and dusty sources, respectively. The latter is based on a single source (Arp220) observed at 850 μm (350 GHz). The former uses a sample of over 100 sources, but at lower frequencies than we are interested in (8 - 43 GHz) and much brighter than the sample we care about (≥ 1 Jy as opposed to a few mJy). QUIET Collaboration et al. (2014) measure polarized flux for radio sources at higher frequencies (up to 90 GHz) and find fractions as large as 20%, but their polarization fraction calculations use non-simultaneous total intensity measurements, so source variability can significantly bias the results. Finally, our own EE power spectrum measurements (Crites et al., 2015) give a 95% upper limit to the rms polarization of radio sources of 20%, or 14% for the combined sample of radio and dusty sources. This is for exactly the frequency and flux range we care about, but it is a fairly blunt measurement.

For CMB-S4 forecasts, we would like to know more precise numbers for polarization fractions of mJy-level sources at observing frequencies around 150 GHz (and possibly 100 GHz)—specifically, we would like to know the mean-squared polarization fraction $\langle p^2 \rangle$ for sources below some flux cut that might be appropriate for Stage-3 or Stage-4 experiments. We have more information on this subject in SPTpol data than we have used so far. Eventually we want to do a dedicated polarized point-source analysis (including making maps appropriate for source extraction), but for now we can investigate source properties using maps made for other purposes. In this memo, I will use a set of 150 GHz maps made for power spectrum analysis of the main SPTpol 500-square-degree (500d) region.

2 Method

The basic analysis I'm doing here is to find sources in a total intensity or temperature (T) map, extract the source fluxes from that T map, extract the Q and U components of the polarized flux from the corresponding Q and U maps, calculate the polarization fraction of each source, and derive some statistics about the population, in particular $\langle p^2 \rangle$. There are two major complicating factors to this measurement:

1. Because p and $\langle p^2 \rangle$ are constructed from quadrature averages of Q and U:

   $$\langle p^2 \rangle = \left\langle \left(\frac{Q}{T}\right)^2 + \left(\frac{U}{T}\right)^2 \right\rangle, \tag{1}$$

   they will be biased by noise.

2. There is some instrumental polarization (or "$T \rightarrow P$ leakage"). Leakage that is uniform across all angular scales and map positions ("monopole leakage") will result in a mean value of Q/T or U/T, and this will also result in a bias on p or $\langle p^2 \rangle$ if not taken into account. Also, any leakage mechanism that has the right ℓ-space shape to affect flux measurements and fluctuates across the map will bias $\langle p^2 \rangle$ by adding variance.

For complication #1, we are lucky in that the SPTpol maps I am using have very uniform noise, so to a good approximation, the polarization fraction measurement of each source is biased by an amount equal to[1]

$$p^2_{\mathrm{noise},i} = \frac{\langle Q^2 \rangle}{T_i^2} + \frac{\langle U^2 \rangle}{T_i^2}, \tag{2}$$

where $\langle Q^2 \rangle$ is the variance in the Q map (and similar for U). Or, in the case in which $\langle U^2 \rangle = \langle Q^2 \rangle \equiv N^2_{\mathrm{pol}}$ (also a good approximation in SPTpol maps),

$$p^2_{\mathrm{noise},i} = \frac{2N^2_{\mathrm{pol}}}{T_i^2}. \tag{3}$$

We can then construct an estimator for $\langle p^2 \rangle$:

$$\begin{aligned} \langle \hat{p}^2 \rangle &= \langle p^2_{\mathrm{meas}} \rangle - \langle p^2_{\mathrm{noise}} \rangle \\ &= \frac{1}{N} \sum_{i=1}^{N} \frac{Q_i^2 + U_i^2}{T_i^2} - \frac{2N^2_{\mathrm{pol}}}{N} \sum_{i=1}^{N} \frac{1}{T_i^2} \\ &= \frac{1}{N} \sum_{i=1}^{N} \frac{Q_i^2 + U_i^2 - 2N^2_{\mathrm{pol}}}{T_i^2}. \end{aligned} \tag{4}$$

[1] There is also a bias due to the noise in T, but it is a *multiplicative* bias that goes as $\delta T^2/T^2$—i.e., $p^2 \rightarrow p^2(1 + \mathcal{O}(\delta T^2/T^2))$—so for sources with signal-to-noise of 10, this is a percent-level *fractional* bias in p^2.

Note that this is not the minimum-variance estimator for $\langle p^2 \rangle$; in fact, this estimator can get noisier as you add more sources. But the minimum-variance estimator would be very heavily weighted toward the brightest few sources, and the weighting would be a function of the true value of $\langle p^2 \rangle$. For ease of calculation and to ensure that the estimated value of $\langle p^2 \rangle$ includes a reasonable sample of sources, I will use the uniform-weight estimator in Equation 4.

To deal with complication #2, I will simply remove an estimate of the $T \rightarrow Q$ or $T \rightarrow U$ leakage from the Q or U estimate for every source. There are several ways to estimate these average leakage quantities, the easiest of which is to just take the mean value of Q/T or U/T for all sources. This will almost certainly bias the final estimate of $\langle p^2 \rangle$ from these same sources, so for the main analysis I will estimate the leakage from a different set of sources than I am using to measure $\langle p^2 \rangle$. In the limit that $T \rightarrow P$ leakage is not a function of angular scale ("monopole leakage"), I should also be able to use the estimates of monopole leakage from our power spectrum analyses. Thankfully, all of these methods (all sources, subset of sources, power spectrum) give consistent answers of roughly 2% $T \rightarrow U$ and 1% $T \rightarrow Q$. Making this subtraction explicit in the estimator for $\langle p^2 \rangle$,

$$\begin{aligned} \langle \hat{p}^2 \rangle &= \langle p^2_{\rm meas,corr} \rangle - \langle p^2_{\rm noise} \rangle \\ &= \frac{1}{N} \sum_{i=1}^{N} \frac{(Q_i - f_{TQ} T_i)^2 + (U_i - f_{TU} T_i)^2 - 2 N^2_{\rm pol}}{T_i^2}. \end{aligned} \tag{5}$$

This correction does not account for any $T \rightarrow P$ leakage that has the right ℓ-space shape to affect flux measurements but is not constant across the map (and thus will add variance and bias $\langle p^2 \rangle$). We do know of at least one such mechanism in SPTpol data; I have argued elsewhere that the effect on these $\langle p^2 \rangle$ measurements should be negligible. If I'm wrong, then the results in this memo should be taken as upper limits on $\langle p^2 \rangle$, not measurements.

3 Data products

As mentioned in the Intro, I am using SPTpol maps made for the purposes of estimating polarized power spectra. These maps contain approximately 2-1/2 seasons of SPTpol observations (2013 May through 2015 October) of a roughly 500-square-degree field centered at right ascension $0^{\rm h}$, declination -57.5°. The maps are made (as all SPTpol maps are) by filtering the detector timestreams and binning and averaging the observations of each map pixel using inverse-variance weighting and ignoring pixel-pixel noise correlations (see Section 4.5 of Keisler et al. 2015 for details). The map is pixelized using the Lambert equal-area azimuthal projection, with pixels 0.5 arcminutes on a side. The filtering used to make these maps is minimal: a low-order polynomial and a set of low-order Fourier modes are projected out of each detector's timestream on each scan across the field. The combined effect of these steps is a scan-direction spatial high-pass filter with a cutoff at approximately $\ell = 80$. This minimal filtering results in some loss of signal-to-noise on point-source scales in the total intensity map (even after a matched filter has been applied) compared to a map that has been timestream filtered more heavily, because some of the large-spatial-scale atmospheric noise leaks to smaller scales in the mapmaking process. It's not a big hit (tens of percent in S/N), and it simply limits the flux range of sources that can be explored (it doesn't result in any bias in the measurement of $\langle p^2 \rangle$, for example).

One effect that could result in a bias on $\langle p^2 \rangle$ is the fact that the brightest sources in this map ($S > 50$ mJy) have been masked in the timestream filtering step. That means that the filtering—and, hence, noise properties—will be slightly different at the locations of these sources than in the rest of the map. I will be using the rms at random locations in the map for my estimate of $N^2_{\rm pol}$, so if the noise at the source locations is significantly different, that will result in a bias on $\langle p^2 \rangle$. I have checked this by calculating the rms in matched-filtered difference maps (made from all left-going scans minus all right-going scans) at the location of bright sources vs. random locations and found no detectable difference (at the 5% level), so I think this is not an issue. Since the filtering only affects very low-ℓ modes, while the point-source signal is mostly high-ℓ, this is not surprising.

Before finding sources in the T map and extracting fluxes from T, Q, and U maps, I apply a point-source-matched filter to each map. This filter is constructed to be roughly optimal for detecting point-like things in the T map, assuming the T map contains only primary CMB, white noise, and the sources of interest, and that it has not been filtered, and that the SPT 150 GHz beam is a $1.2'$ FWHM Gaussian. All of these assumptions are wrong, of course, but they're not very wrong. As with using slightly non-optimal maps, this just slightly raises the flux threshold of the resulting catalog. (It also would result in a small error in the conversion from filtered map value to flux, but as I mention below, that factor divides out in the calculation of polarization fraction, so I can ignore it.) I find sources in the T map down to signal-to-noise of 10, using a simple SExtractor-like group finder. I filter the Q and U maps with the T-matched filter, and for every $> 10\sigma$ source I extract the map values in all three maps. Since the conversion from map value to source flux will be the same in all three maps (because they have been filtered identically), and I am only interested in ratios of polarized to unpolarized flux, I don't bother applying any conversions to the map values.

4 Results

Before giving results, I will quickly reiterate the steps in the calculation of $\langle \hat{p}^2 \rangle$:

- Grab T, Q, and U maps.
- Create a point-source-matched filter for the T map and filter all three maps with it.
- Find sources down to 10σ in the T map.
- Extract filtered map values from all three maps at the location of every $> 10\sigma$ source.
- Split the sources into a bright set and a dim set, and calculate f_{TQ} and f_{TU} from the mean value of Q/T and U/T in the dim set.
- Estimate $N^2_{\rm pol}$ from the variance in the filtered Q and U maps.
- Use Equation 5 to calculate the estimator for mean-squared polarization $\langle \hat{p}^2 \rangle$ in the bright set.

As a secondary result, and understanding that this result could be biased by the $T \rightarrow P$ projection step, I will also calculate $\langle \hat{p}^2 \rangle$ on the dim set.

The source-finding step in the T map results in a catalog of 387 sources above 10σ (roughly one per square degree). To decide how to split the sources into bright and dim, I've done toy simulations of the

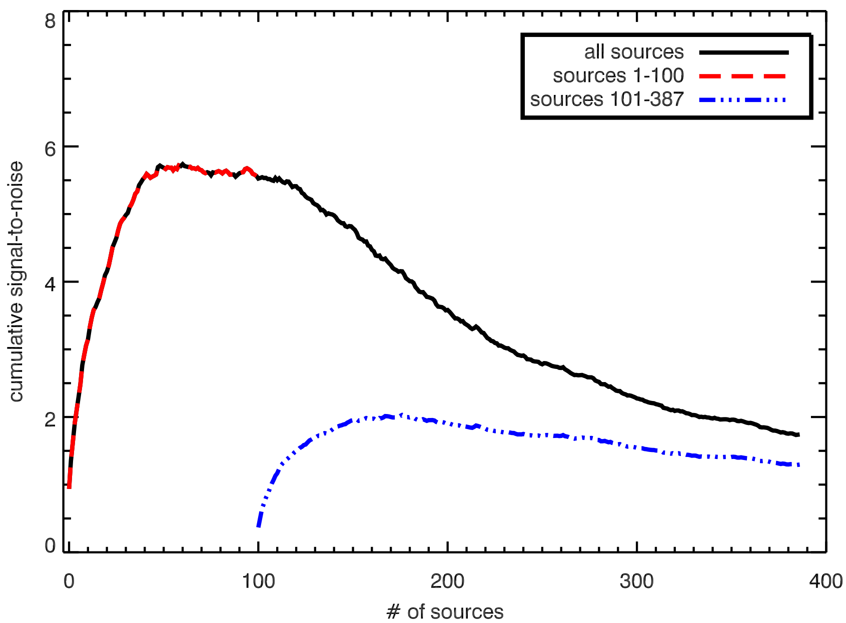

Figure 1: Expected signal-to-noise on $\langle p^2 \rangle$ as a function of sources included in the calculation. The solid black and dashed red lines show the cumulative S/N starting from the brightest source and going down in the catalog. (The S/N is not monotonic with number of sources included, because the estimator $\langle \hat{p}^2 \rangle$ is not the optimal estimator; see Section 2 for details.) The blue dot-dashed line shows the cumulative S/N starting with the 101st-brightest source and going farther down the catalog. These results are specific to the source catalog and map noise in this analysis.

$\langle \hat{p}^2 \rangle$ calculation using the actual total intensity values for the found sources and the actual noise level in the filtered Q and U maps. I have to guess at the true distribution of polarization fraction for this simulation, and I've used a Gaussian with mean zero and $\sigma_p = 0.03$ (so that $\langle p^2_{\mathrm{true}} \rangle = 0.0009$). With this guess, I find that the signal-to-noise on $\langle p^2 \rangle$ saturates (at about 6) after I include approximately the brightest 100 sources (see Figure 1). This seems like a nice round number to choose, so I use the top 100 sources as my bright set and the remaining 287 sources to estimate $T \rightarrow P$. Using a rough conversion from filtered map value to flux, the 100th-brightest source is roughly 25 mJy. (The full range of the catalog is 6.7 mJy to 533 mJy). For the secondary result, I expect signal-to-noise on $\langle p^2 \rangle$ for the dim set of approximately 1.5 (see Figure 1).

Source set	Approximate flux range	$\langle \hat{p}^2 \rangle \times 10^3$
Bright	25 - 533 mJy	1.42 ± 0.15
Dim	6.7 - 25 mJy	1.42 ± 0.67

Table 1: Summary of main results.

The main results of this memo are summarized in Table 1. The values of $\langle \hat{p}^2 \rangle$ for the bright and dim source sets are:

$$\begin{aligned} \text{Bright}: \langle \hat{p}^2 \rangle &= (1.42 \pm 0.15) \times 10^{-3} \\ \text{Dim}: \langle \hat{p}^2 \rangle &= (1.42 \pm 0.67) \times 10^{-3} \end{aligned}$$

The agreement to three significant digits between the bright and dim sets is a statistical fluke that goes away if I tweak the definition of either set. (For instance, if I use sources 101-300 for the dim set, I get a value of $\langle p^2 \rangle$ consistent with zero.) It is worth noting that the best-fit values of $\langle \hat{p}^2 \rangle$ I get are $\sim$50% higher than the value I used to estimate the signal-to-noise as a function of number of sources, and to estimate the uncertainty on $\langle \hat{p}^2 \rangle$. If I re-run the toy simulation with a true underlying $\langle p^2 \rangle$ value equal to 0.0014, the optimal way to split the sources shifts slightly to lower source number, and the uncertainty on $\langle \hat{p}^2 \rangle$ increases by roughly 25%, so you should only trust these error bars to about that level.

4.1 Radio sources vs. dusty galaxies

We expect the two main families of emissive sources at millimeter wavelengths—synchrotron-emitting "radio sources" and dusty, star-forming galaxies—to have different mean-squared polarization fractions; namely, we expect radio sources to be more polarized (e.g., De Zotti et al. 2015). An obvious question about the result presented here is: what fraction of the sources in this sample are radio vs. dusty? As it turns out, this sample used here is expected to be over 90% radio sources. In the catalog of Mocanu et al. (2013), there were 145 sources above a raw flux of 25 mJy at 150 GHz and 493 sources between raw fluxes of 6.7 and 25 mJy at 150 GHz. Of the 145 sources above 25 mJy, 142 (98%) were classified as radio sources. Of the 493 sources between 6.7 and 25 mJy, 449 (91%) were classified as radio sources. The results here can thus be considered a measurement of (or upper limit on) the mean-squared polarization of radio sources at 150 GHz. Given that the dusty source counts have a much steeper negative slope with flux than radio sources, it is expected that dusty sources will dominate the 150 GHz counts below approximately 1 mJy (see, e.g., Figure 4 of Toffolatti et al. 1998). Thus, the mean-squared polarization fraction of the total population contributing to the contamination of Stage-3 or Stage-4 polarized power spectrum measurements should be lower than the result found here.

5 Conclusions

Using a quick and dirty analysis of some existing SPTpol maps, I have measured the mean-squared polarization fraction $\langle p^2 \rangle$ in sources selected in total intensity at 150 GHz. I estimate $\langle \hat{p}^2 \rangle$ for two sets of sources, one from roughly 25 mJy to 500 mJy and one from roughly 6.7 mJy to 25 mJy. (Both of these subsets are dominated by radio sources.) The best value of $\langle \hat{p}^2 \rangle$ for both sets is 1.42×10^{-3}. The fractional uncertainty on $\langle \hat{p}^2 \rangle$ is about 10% for the bright set and about 50% for the dim set. While I have corrected for monopole, uniform-across-the-map $T \rightarrow P$ leakage in this measurement, I have not corrected for leakage that has the right ℓ-space shape to affect flux measurements but is not constant across the map (and thus will add variance and bias $\langle p^2 \rangle$). If there is substantial leakage of this nature (and I argue there is not), then the results in this memo should be taken as upper limits on $\langle p^2 \rangle$, not measurements.

References

Battye, R. A., Browne, I. W. A., Peel, M. W., Jackson, N. J., & Dickinson, C. 2011, MNRAS, 413, 132

Crites, A. T., et al. 2015, ApJ, 805, 36

De Zotti, G., et al. 2015, J. of Cosm. & Astropart. Phys., 6, 018

Keisler, R., et al. 2015, ApJ, 807, 151

Mocanu, L. M., et al. 2013, ApJ, 779, 61

QUIET Collaboration, et al. 2014, ArXiv e-prints, 1412.1111

Seiffert, M., Borys, C., Scott, D., & Halpern, M. 2007, MNRAS, 374, 409

Toffolatti, L., Argueso Gomez, F., de Zotti, G., Mazzei, P., Franceschini, A., Danese, L., & Burigana, C. 1998, MNRAS, 297, 117