EXTREME DIGITISATION FOR GROUND-BASED COSMIC MICROWAVE BACKGROUND EXPERIMENTS

L. Balkenhol1 and C. L. Reichardt1

ABSTRACT

The large size of the time ordered data of cosmic microwave background experiments presents challenges for mission planning and data analysis. These issues are particularly significant for Antarctica and space-based experiments, which depend on satellite links to transmit data. We explore the viability of reducing the time ordered data to few bit numbers to address these challenges. Unlike lossless compression, few bit digitisation introduces additional noise into the data. We present a set of one, two, and three bit digitisation schemes and measure increase in noise in the cosmic microwave background temperature and polarisation power spectra. The digitisation noise is independent of angular scale and is well-described as a constant percentage of the original detector noise. Three bit digitisation increases the map noise level by < 2%, while reducing the data volume by a factor of ten relative to 32-bit floats. Such digitisation is a promising strategy for upcoming experiments.

Subject headings: cosmic background radiation — polarization — techniques: miscellaneous

1. INTRODUCTION

Observations of the cosmic microwave background (CMB) have played a key role in cosmology (Penzias & Wilson 1965; Smoot et al. 1992; Bennett et al. 2013; Hanson et al. 2013; Planck Collaboration et al. 2018). Current ground-based CMB experiments, like SPT-3G (Benson et al. 2014) at the South Pole telescope (SPT) and Adv. ACTpol (Thornton et al. 2016), at the Atacama cosmology telescope (ACT), target science goals such as the discovery of inflationary gravitational waves, measuring the number of relativistic species, the neutrino mass sum, and mapping the large-scale distribution of matter through gravitational lensing and the Sunyaev-Zeldovich (SZ) effects. Through lensing and the SZ effects the CMB probes structure formation, reionisation, and is a powerful test for dark matter and dark energy models (Abazajian et al. 2016; Di Valentino et al. 2018; Ishino et al. 2016; Kogut et al. 2011).

Over the last two decades, CMB experiments have gone from single detectors to over 10,000 detectors. The CMB community has developed a variety of compression techniques and computational approaches to handle the increasing volume of data (Tristram & Ganga 2007). These include the compression of time-ordered data (TOD) into maps (Tegmark 1997b), bandpower estimation (Tegmark 1998), and the pseudo-$C_l$ method (Brown et al. 2005).

A potential hurdle for experiments at remote locations are the transmission limitations of satellite links. Space-based experiments have employed a combination of lossless and lossy compression techniques, including reduced bits in the TOD (Gaztanaga et al. 1998; Maris et al. 2003). Antarctica-based experiments that transmit a portion of their results via a satellite link downsample their data to meet telemetry allocations, but have not yet used few bit digitisation of the TOD. As we approach the next generation ground-based experiment, CMB-S4 (Abazajian et al. 2016), and the launch of the new space-based missions COrE+ (Delabrouille et al. 2018), Lite-BIRD (Ishino et al. 2016), and PIXIE (Kogut et al. 2011), we must consider potential transmission bottlenecks carefully.

In this work we present the method of extreme digitisation, which reduces a many bit (often 32 or 64 bit) signal to a few bits for ground-based experiments. We apply extreme digitisation to simulated TOD and detail the resulting effects on temperature and polarisation power spectra. We find that an optimal three bit digitisation scheme adds < 2% to the map noise level.

This work is structured as follows. In section §2 we detail the challenges that come with handling large data volumes, introduce the process of extreme digitisation and lay out the framework used to test its performance. We present results for white and 1/f detector noise in section §3. We summarise our findings in §4.

2. DIGITISATION

2.1. The Challenges Of Large Data Sets

The science goals of upcoming CMB experiments depend on achieving substantially faster mapping speeds. Given CMB detectors are generally photon noise limited, improving the mapping speed means adding more detectors. As a result the number of detectors (and data volume) of ground-based experiments has followed an exponential trend like Moore’s law, doubling approximately every 2 years (Abazajian et al. 2016; Abazajian et al. 2015).

The South Pole is one of the best sites for CMB observations on Earth (Chamberlin 2001; Ruhl et al. 2004). CMB-S4 plans to include several telescopes at the South Pole (Abazajian et al. 2016; Barron et al. 2018), which will generate a data influx of $\sim 10^6$Tb/d. Transferring this data volume via satellite would be expensive. For context, the transmission allocation for a current CMB experiment, SPT-3G, is 150Gh/d. The transmission bottleneck could be overcome by recovering the full data on hard drives every summer and transmitting a downsampled version of the data. Downsampling eliminates high frequency information, which makes it unsuitable for science on small angular scales, such as SZ galaxy clusters.
The potential delay (i.e. only getting the high frequency data once a year) also introduces risks by delaying when potential faults or issues at high frequency are noticed. One can reduce these issues by running substantial portions of the analysis at the South Pole, but this comes with its own costs and challenges.

Beyond transmission challenges, a larger data volume increases the size and cost of disk arrays and makes end-to-end simulations of experiments for the purpose of optimisation on systematics estimation more time consuming. Given the sheer size of upcoming data sets, full end-to-end simulations may prove impractical for CMB-S4 (Abazajian et al. 2016). The exponential growth of data once a year also introduces risks by delaying when the potential delay (i.e. only getting the high frequency data once a year) also informs us that the threshold levels should lie midway between their adjacent output levels. Setting the derivative of equation 2 with respect to \( y_i \) to zero gives the additional condition

\[
\frac{\partial D}{\partial y_i} = -2 \int_{x_i}^{x_{i+1}} (x - y_i) p(x) dx = 0.
\]

(5)

This implies that we should choose \( y_i \), such that it halves the area underneath \( p(x) \) in the interval from \( x_i \) to \( x_{i+1} \).

To progress further, we must consider the probability distribution \( p(x) \) of the input signal. For Gaussian detector noise, we have \( p(x) = (1/\sqrt{2\pi \sigma^2}) e^{-x^2/2\sigma^2} \). We also assume that we are in the low signal to noise regime, as is appropriate for ground-based CMB experiments. The solution for this case was given by Max (1960). For one bit digitisation, the result is

\[
\hat{s}_1(t) = \begin{cases} 
1 & s(t) > 0 \\
-1 & s(t) \leq 0 
\end{cases}
\]

(7)

For two bit digitisation we have

\[
\hat{s}_2(t) = \begin{cases} 
1.51\sigma, & s(t) \geq 0.98\sigma \\
-0.45\sigma, & 0 \leq s(t) < 0.98\sigma \\
-1.51\sigma, & -0.98\sigma \leq s(t) < 0 \\
-1.50\sigma, & s(t) \leq 0 
\end{cases}
\]

(8)

Finally the optimal three bit digitisation is described by the eight-level function

\[
\hat{s}_3(t) = \begin{cases} 
2.15\sigma, & s(t) \geq 1.75\sigma \\
-1.34\sigma, & 1.05\sigma \leq s(t) < 1.75\sigma \\
0.76\sigma, & 0.50\sigma \leq s(t) < 1.05\sigma \\
-0.25\sigma, & -0.50\sigma \leq s(t) < 0 \\
-0.76\sigma, & -1.05\sigma \leq s(t) < -0.50\sigma \\
-1.34\sigma, & -1.75\sigma \leq s(t) < -1.05\sigma \\
-2.15\sigma, & -1.75\sigma \leq s(t) < 0 
\end{cases}
\]

(9)

The digitisation schemes listed above do not conserve signal power and need to be re-calibrated to recover the correct power in the map. We demonstrate that the schemes laid out by Max (1960) can be rescaled to conserve power and derive an analytical expression for the normalisation constant in the appendix. We use this analytical rescaling for all results in this work.

2.3. From Time Ordered Data To Maps

To investigate the performance of the derived digitisation schemes, we simulate timelapse level scans over a single CMB realisation. We add detector noise and apply the digitisation schemes above to the simulated I, Q, and U TOD. We also retain the original 64 bit TOD to construct control maps. The different timelapses are binned into HEALPix (Górski et al. 2005) maps with resolution NSIDE = 4096 for a T, E, B power spectrum analysis.

We use the healpy Python package to generate a single CMB realisation for the best-fit Planck 2015 cosmology.

\[\text{Available at http://healpix.sourceforge.net/}^{2}\]
Specifically, the key cosmological parameters are \( H_0 = 67.8\text{ km s}^{-1}\text{Mpc}^{-1} \), \( \Omega_m = 0.308 \), \( n_s = 0.968 \), and \( \tau = 0.066 \) (Planck Collaboration et al. 2016).

We simulate observing a 600 deg\(^2\) patch of the sky with a single detector. The detector noise level for temperature is 500\(\mu\text{K}\sqrt{\text{s}}\) and for polarisation observations \( \sqrt{2} \times 500\mu\text{K}\sqrt{\text{s}} \), i.e. the corresponding photon noise limit. Constant elevation scans (CES) are performed beginning at right ascension (RA) and declination (DEC) \((0^\circ, 0^\circ)\). After covering \(24^\circ30'\) along RA the detector is reset to a RA of \(0^\circ\) and an offset in DEC corresponding to the pixel size. Through repetition of CES with constant steps in DEC the survey patch is covered. This imitates the scan strategy of the SPT (Schaffer et al. 2011). The entire scan strategy is repeated 100 times with offsets in the starting RA and DEC up to the size of one pixel, ensuring that pixels are sampled uniformly.

We create naive I,Q, and U maps from the simulated I, Q, and U TOD, i.e. the value of a map pixel is the average of all TOD samples lying within that pixel. We create four maps: one using the original 64 bit TOD and one for each of the digitisation schemes. We save the output maps at 800, 8,000, 80,000, 1,024,000, 10,240,000 and 102,400,000 hits per pixel.

2.4. Power Spectrum Estimation

To minimise boundary effects in the subsequent analysis we apodise the observed patch using a cosine mask. For the case of white detector noise we use PolSpice (Szapudi et al. 2001) to compute TT, EE and BB power spectra from output maps.

For the later considered case of anisotropic noise, we use healpy to calculate the spherical harmonic coefficients\( a_{\ell m} \). Since the scan strategy of the simulated observations is similar to the SPT’s, smaller values of \( m \) correspond to larger angular features along the scan direction (Chown et al. 2018). Because of the anisotropic nature of the noise, equally weighting all modes leads to inaccurate power spectra. As a first order approximation to optimal weighting we discount the lowest modes in scan direction.

We rebin the calculated power spectra as mentioned in section §3 and assign Fisher matrix error bars (Tegmark 1997a).

As seen in Figure 2 (appendix) all spectra of extremely digitised data recover the input to a satisfying degree. Differences between few bit and the input power spectra are due to observing a single CMB realisation and residual boundary effects.

3. RESULTS

3.1. White Noise

We analyse the results of the simulated observations to quantify the distortion digitisation causes in the power spectra and investigate whether the additional noise has any angular scale dependence.

We infer the fractional increase of the original map noise level, \( \Delta\sigma/\sigma \), to quantify the distortion caused by extreme digitisation. We calculate \( \Delta\sigma/\sigma \) from the noise dominated angular scales in the power spectra via

\[
\frac{\Delta\sigma}{\sigma} = \sqrt{1 + \frac{C^N}{C^L}} - 1, \tag{10}
\]

where \( C^N \) and \( C^L \) are the detector noise level and the additional digitisation noise respectively. Before calculating \( C^N/C^L \) we rebinned the power spectra to \( \Delta\ell \approx 100 \). This ensures that points in the noise tail are independent, allowing us to extract an uncertainty for \( \Delta\sigma/\sigma \).

The deduced additional noise for one, two, and three bit digitisation is shown in Table 1 in the appendix. We see that 3 bit digitisation performs best, followed by two bit and one bit. This is expected, since more bits allow a more faithful representation of the input signal. We observe that the fractional increase, \( \Delta\sigma/\sigma \), is independent of the hits per pixel in the maps and that the compression performs as well for polarisation as it does for temperature data. Finally, it is striking how little noise is added. On average one bit digitisation leads to a \((25.1 \pm 0.9)\%\) increase in the map noise level, two bit yields a \((6.4 \pm 0.4)\%\) increase and three bit adds only \((1.8 \pm 0.2)\%\). The uncertainties given are inferred from the standard deviation of rebinned points in the \( C^N/C^L \) noise tail. The performance of few bit digitisation is impressive, given that all schemes considered reduce TOD volume by at least an order of magnitude.

To investigate whether the noise added through digitisation has any angular scale dependence, we subtract the input map from simulated observations. The power in the difference map is plotted in Figure 1. There is no sign of any angular scale dependence.

3.2. 1/f Noise

While digitisation clearly works extremely well for white noise, real ground-based CMB experiments face several sources of low-frequency noise. High on the list is atmospheric noise, which Lay & Halverson (2000) describe as

\[
|N_{1/f}(\ell)|^2 = |N_{\text{white}}|^2 \left[ 1 + \left( \frac{\ell}{\ell_{\text{knee}}} \right)^{-3/2} \right]^2, \tag{11}
\]

where \( N_{\text{white}} \) denotes white detector noise. In this expression \( \ell \) describes the multipole moment of angular features of corresponding size along the scan direction. The scale at which such modes are amplified appreciably by the second term in equation 11 is set by \( \ell_{\text{knee}} \). We choose \( \ell_{\text{knee}} = 1000 \) for temperature and \( \ell_{\text{knee}} = 50 \) for polarisation. In practice CMB experiments suppress power in the lowest modes through a range of techniques, including fitting polynomials, sines, and cosines to the TOD. We simulate successful removal of these modes by applying a high pass filter to the noise, such that

\[
|N(\ell)|^2 = |N_{1/f}(\ell)|^2 \exp \left[ - \left( \frac{\ell_{\text{min}}}{\ell} \right)^{12} \right], \tag{12}
\]

with \( \ell_{\text{min}} = 336.3 \) for temperature and \( \ell_{\text{min}} = 16.8 \) for polarisation.

We use the same digitisation schemes as for white noise, but modify the parameter \( \sigma \) in equations 7, 8 and 9 by considering Parseval’s theorem on the above noise profile. The appropriate value is obtained by evaluating the integral

\[
\sigma^2 = \int_0^\infty |N(\ell)|^2 d\ell. \tag{13}
\]
After modifying the digitisation schemes derived in subsection §2.2 accordingly, we carry out the same simulations as described in subsection §2.3 with the noise profile given in equation 12. We save the output maps at 8,000, 102,400 and 1,024,000 hits per pixel. The analysis procedure of subsection §3.1 is used to characterise the digitisation noise.

We recover similar results to the white noise case, with three bit digitisation performing the best, followed by two bit, and one bit. Table 2 in the appendix displays the fractional increase to the detector noise level due to digitisation, \( \Delta \sigma / \sigma \). While polarisation observations seem to incur a slightly higher noise penalty compared to temperature, deviations do not exceed 3\( \sigma \). On average one bit digitisation leads to a (27.7 + 1.5)\% increase in the map noise level, two bit adds (6.2 + 0.5)\%, and three bit leads to an increase of only (1.7 + 0.2)\%. The digitisation noise shows no angular scale dependence. Thus the effectiveness of few bit digitisation is not limited to the white noise case, but extents to more realistic noise profiles.

4. CONCLUSION

In this work we have conducted an investigation of extreme digitisation as a technique in combating the challenges of large data volumes for ground-based CMB experiments. In particular the reduction of TOD volume by an order of magnitude decreases the transmission requirements from remote locations. The reduction may also streamline the analysis and simulation of TOD.

We present a set of one, two, and three bit digitisation schemes. For white and 1/f detector noise alike, we find that optimal three bit digitisation adds as little as < 2\% to the map noise level. This is true for temperature and polarisation observations. No change in the results is observed for maps of different hits per pixel and no angular scale dependence was observed in the added noise.

Given that the digitisation noise remains low at small angular scales, cluster finding algorithms may perform well on few bit TOD. Investigating this will also help determine the higher order statistical moments, i.e. skewness and kurtosis, of the digitisation noise.

We are grateful for insightful discussions about the prospects of extreme digitisation for CMB experiments with Andrew Melatos and Patrick Clearwater. We thank Srinivasan Raghunathan, Federico Bianchini, Nikhel Gupta, Andrew Melatos, and Patrick Clearwater for valuable feedback on the manuscript. We acknowledge support from an Australian Research Council Future Fellowship (FT150100074), and also from the University of Melbourne. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science.

This research made use of the NumPy (Travis E. Oliphant 2006), SciPy (Jones et al. 2001–), Matplotlib (Hunter 2007), and Astropy (Astropy Collaboration et al. 2013) packages.
We use the standard deviation of points in the \( C_\gamma \) digitisation scheme. It is relatively simple to impose power conservation, since it boils down to rescaling the digitised depth of observations and channel for a given digitisation scheme.

\[
\text{Therefore } D_p \text{ occurring. The probabilities } D_p \text{ can be expressed as integrals over probability distributions. Assuming Gaussian detector noise we introduce normal distributions } p(\xi) \text{ with mean zero and standard deviation } \sigma. \]

\[
(D(\mu + \xi_1)D(\mu + \xi_2)) = \sum_{i=0}^{N} \sum_{j=0}^{N} \int_{x_{i,j} - \mu}^{x_{i+1,j} - \mu} \int_{x_{j} - \mu}^{x_{j+1} - \mu} y_i y_j p(\xi_1)p(\xi_2) d\xi_1 d\xi_2. \tag{3}
\]
TABLE 2
ADDITIONAL NOISE - 1/f DETECTOR NOISE

| Hits Per Pixel | Channel | 1 Bit | 2 Bit | 3 Bit |
|---------------|---------|-------|-------|-------|
| 8000          | TT      | 0.231 ± 0.017 | 0.058 ± 0.005 | 0.016 ± 0.002 |
|               | EE      | 0.300 ± 0.012 | 0.064 ± 0.004 | 0.018 ± 0.002 |
|               | BB      | 0.294 ± 0.022 | 0.062 ± 0.007 | 0.017 ± 0.003 |
| 102400        | TT      | 0.235 ± 0.014 | 0.060 ± 0.004 | 0.017 ± 0.002 |
|               | EE      | 0.299 ± 0.009 | 0.064 ± 0.004 | 0.017 ± 0.002 |
|               | BB      | 0.298 ± 0.025 | 0.064 ± 0.007 | 0.018 ± 0.002 |
| 1024000       | TT      | 0.238 ± 0.013 | 0.06 ± 0.005  | 0.017 ± 0.002 |
|               | EE      | 0.299 ± 0.008 | 0.063 ± 0.004 | 0.017 ± 0.002 |
|               | BB      | 0.297 ± 0.019 | 0.063 ± 0.006 | 0.017 ± 0.003 |
| Average       | TT      | 0.235 ± 0.015 | 0.059 ± 0.005 | 0.016 ± 0.002 |
|               | EE      | 0.299 ± 0.010 | 0.064 ± 0.004 | 0.017 ± 0.002 |
|               | BB      | 0.297 ± 0.019 | 0.063 ± 0.006 | 0.017 ± 0.003 |

Note. — Percent addition to the map noise level, \( \Delta \sigma / \sigma \), due to one, two, and three bit digitisation for 1/f detector noise. Uncertainties are obtained in the same way as for white noise. As in the white noise case we observe no significant trend with the numbers of hits per map pixel or channel for a given digitisation scheme. No significant deviation in the values compared to simulations assuming white detector noise is found.

We separate the above sums and recognise the error function.

\[
\langle D(\mu + \xi_1)D(\mu + \xi_2) \rangle = \left\{ \sum_{i=0}^{N} \frac{y_i}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{1}{(\sqrt{2\sigma})^{2n+1}} \left((x_{i+1} - \mu)^{2n+1} - (x_i - \mu)^{2n+1}\right) \right\}\times\left\{ \sum_{j=0}^{N} \frac{y_j}{y} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{1}{(\sqrt{2\sigma})^{2m+1}} \left((x_{j+1} - \mu)^{2m+1} - (x_j - \mu)^{2m+1}\right) \right\},
\]

(4)

\[
\approx \left\{ \sum_{i=0}^{N} \frac{y_i}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{1}{(\sqrt{2\sigma})^{2n+1}} \left(\sum_{k=0}^{2n+1} \frac{2n+1}{k} \mu^k \left(x_{i+1}^{2n+1-k} - x_i^{2n+1-k}\right)\right) \right\}\times\left\{ \sum_{j=0}^{N} \frac{y_j}{\sqrt{\pi}} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{1}{(\sqrt{2\sigma})^{2m+1}} \left(\sum_{l=0}^{2m+1} \frac{2m+1}{l} \mu^l \left(x_{j+1}^{2m+1-l} - x_j^{2m+1-l}\right)\right) \right\},
\]

(5)

Where we have used the Maclaurin series expansion of the error function and the binomial expansion. Moving the \( i, j \) summation forwards yields

\[
\langle D(\mu + \xi_1)D(\mu + \xi_2) \rangle = \frac{1}{\pi} \sum_{n,m=0}^{\infty} \frac{1}{n!(2n+1)} \frac{1}{m!(2m+1)} \frac{(-1)^{n+m}}{(\sqrt{2\sigma})^{2n+2m+2}} \times \sum_{k,l=0}^{2n+1} \left(\sum_{i,j=0}^{N} y_i y_j \psi_{i,j}^k \psi_{i,j}^l\right),
\]

(7)

with

\[
\psi_{i,j}^k = x_{i+1}^{2n+1-k} - x_i^{2n+1-k}.
\]

(8)

We split the innermost term in equation 7 and shift indices such that
Fig. 2.— Input power spectra used to generate the CMB template map plus detector noise level (black) and digitisation noise (gray); power spectra using three (violet, right), two (orange, middle), and one (green, left) bit TOD. Main panels show the TT, EE, and BB power spectra; accompanying panels show the residual of the input plus noise curves and the power spectra recovered from few bit TOD. We assume white detector noise and show results for maps with 1,024,000 hits per pixel. Spectra are binned to $\Delta \ell = 100$. Variations are attributed to observing a single CMB realisation and residual boundary effects. Even though small angular scales have small errors in the temperature power spectrum, no significant trend in the residual is found. No qualitative differences are visible for other depths of observation.
\[
\sum_{i,j=0}^{N} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} = \left( \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} + \sum_{i=N/2}^{N} \sum_{j=0}^{N/2} + \sum_{i=0}^{N/2} \sum_{j=N/2}^{N} + \sum_{i=N/2}^{N} \sum_{j=N/2}^{N} \right) y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m}. \tag{9}
\]

\[
= \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} \left( y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} + y_{N-i}y_{j} \psi_{N-i}^{k,n} \psi_{j}^{l,m} + y_{i}y_{N-j} \psi_{i}^{k,n} \psi_{N-j}^{l,m} + y_{N-i}y_{N-j} \psi_{N-i}^{k,n} \psi_{N-j}^{l,m} \right). \tag{10}
\]

Equations 4 and 6 in subsection \S 2.2 demand that for a Gaussian input distribution \(y_{i} = -y_{N-i}\) and \(x_{i} = -x_{N+1-i}\). This simply states symmetry of the digitisation thresholds and output levels about zero (or after a global mean has been subtracted). Therefore

\[
\sum_{i,j=0}^{N} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} = \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} \left( \psi_{i}^{k,n} \psi_{j}^{l,m} + (-1)^{k+1} \psi_{i}^{k,n} \psi_{j}^{l,m} + (-1)^{l+1} \psi_{i}^{k,n} \psi_{j}^{l,m} + (-1)^{k+l} \psi_{N-i}^{k,n} \psi_{N-j}^{l,m} \right), \tag{11}
\]

since \(\psi_{N-i}^{k,n} = (-1)^{k} \psi_{i}^{k,n}\). For ground-based CMB observations \(\mu\) is small. Therefore we calculate the leading order terms in \(\mu\) in equation 7. Since this equation has powers of \(\mu^{k+l}\) we split the sum over \(k, l\) as follows

\[
\sum_{i,j=0}^{N} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} = \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} y_{i}y_{j} \left( \psi_{i}^{k,n} \psi_{j}^{l,m} \psi_{i}^{k,n} \psi_{j}^{l,m} \psi_{i}^{k,n} \psi_{j}^{l,m} \psi_{i}^{k,n} \psi_{j}^{l,m} \right) = 0, \tag{12}
\]

For odd \(k\) and \(l\) equation 11 may be simplified to

\[
\sum_{i,j=0}^{N} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} = \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} y_{i}y_{j} \left( \psi_{i}^{k,n} \psi_{j}^{l,m} + \psi_{i}^{k,n} \psi_{j}^{l,m} + \psi_{i}^{k,n} \psi_{j}^{l,m} + \psi_{i}^{k,n} \psi_{j}^{l,m} \right) = 4 \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m}. \tag{13}
\]

This term survives and leads to even powers of \(\mu\) in equation 7. Considering the case of odd \(k\) and even \(l\) we observe

\[
\sum_{i,j=0}^{N} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} = \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} y_{i}y_{j} \left( \psi_{i}^{k,n} \psi_{j}^{l,m} + \psi_{i}^{k,n} \psi_{j}^{l,m} \psi_{i}^{k,n} \psi_{j}^{l,m} \right) = 0, \tag{14}
\]

and similarly

\[
\sum_{i,j=0}^{N} y_{i}y_{j} \psi_{i}^{k,n} \psi_{j}^{l,m} = \sum_{i=0}^{N/2} \sum_{j=0}^{N/2} y_{i}y_{j} \left( \psi_{i}^{k,n} \psi_{j}^{l,m} + \psi_{i}^{k,n} \psi_{j}^{l,m} \psi_{i}^{k,n} \psi_{j}^{l,m} \right) = 0, \tag{15}
\]

for odd \(k\) and even \(l\). Therefore all odd powers of \(\mu\) in equation 7 vanish. Moreover there is no \(\mu\) independent term, since this could only be produced by \(k = l = 0\), but even \(k\) and \(l\) terms are shown to vanish. Therefore

\[
\langle D(\mu + \xi_{1})D(\mu + \xi_{2}) \rangle = c\mu^{2} + \mathcal{O}(\mu^{4}) \tag{16}
\]

where \(c\) is some constants that depend on the specific digitisation scheme chosen. Returning to equation 1 we see

\[
\gamma^{2} = \frac{\langle (\mu + \xi_{1})(\mu + \xi_{2}) \rangle}{\langle D(\mu + \xi_{1})D(\mu + \xi_{2}) \rangle} = \frac{1}{c + \mathcal{O}(\mu^{2})} = \frac{1}{c} \tag{17}
\]

where we have ignored terms beyond quadratic order. By investigating \(k = l = 1\) terms in equation 7 we obtain an expression for \(\gamma\).

\[
c = \sum_{i,j=0}^{N} y_{i}y_{j} \pi \left[ \sum_{m=0}^{\infty} \frac{n!(2n+1)}{m!(2m+1)} \frac{1}{(\sqrt{2})^{2m+2}} \left( \binom{2m+1}{1} \binom{2n+1}{1} (x_{i+1}^{2n+1} - x_{i}^{2n+1}) \right) \right] \tag{18}
\]

\[
= \sum_{i,j=0}^{N} \frac{y_{i}y_{j}}{2\pi \sigma^{2}} \left[ \sum_{m=0}^{\infty} \frac{n!}{m!} \frac{1}{(\sqrt{2})^{2m}} (x_{i+1}^{2m} - x_{i}^{2m}) \right] \tag{19}
\]

\[
= \sum_{i,j=0}^{N} \frac{y_{i}y_{j}}{2\pi \sigma^{2}} \left\{ \exp \left[ -\left( \frac{x_{i+1}}{\sqrt{2} \sigma} \right)^{2} - \exp \left[ -\left( \frac{x_{i}}{\sqrt{2} \sigma} \right)^{2} \right] \right] \right\} \exp \left[ -\left( \frac{x_{j+1}}{\sqrt{2} \sigma} \right)^{2} \right] \tag{20}
\]

\[
= \sum_{i,j=0}^{N} \frac{y_{i}y_{j}}{2\pi \sigma^{2}} \left\{ \exp \left[ -\left( \frac{x_{i+1}}{\sqrt{2} \sigma} \right)^{2} \right] - \exp \left[ -\left( \frac{x_{i}}{\sqrt{2} \sigma} \right)^{2} \right] \right\} \exp \left[ -\left( \frac{x_{j+1}}{\sqrt{2} \sigma} \right)^{2} \right].
\]
The normalisation constant therefore is given by

$$\gamma = \left( \sum_{i,j=0}^{N} \frac{y_i y_j}{2\pi \sigma^2} \left\{ \exp \left[ -\left( \frac{x_{i+1}}{\sqrt{2\sigma}} \right)^2 \right] - \exp \left[ -\left( \frac{x_{j+1}}{\sqrt{2\sigma}} \right)^2 \right] \right\} \right)^{-1/2} \tag{21}$$

This can be calculated for a given digitisation scheme $N, x_i, y_i$ and standard deviation of the noise $\sigma$. Notice that if the digitisation scheme is chosen such that $y_i$ and $x_i$ depend linearly on $\sigma$ the normalisation constant is independent of $\sigma$.

REFERENCES

Abazajian, K. N., et al. 2015, Astroparticle Physics, 63, 66
Abazajian, K. N., et al. 2016
Astropy Collaboration, et al. 2013, A&A, 558, A33
Barron, D., et al. 2018, Journal of Cosmology and Astroparticle Physics, 2018, 009
Bennett, C. L., et al. 2013, ApJS, 208, 20
Benson, B. A., et al. 2014, in Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, Vol. 9153, 91531P
Brown, M. L., Castro, P. G., & Taylor, A. N. 2005, MNRAS, 360, 1262
Chamberlin, R. A. 2001, Journal of Geophysical Research: Atmospheres, 106, 20101
Chown, R., et al. 2018, ArXiv e-prints, 1803.10682
Delabrouille, J., et al. 2018, J. Cosmology Astropart. Phys., 4, 014
Di Valentinio, E., et al. 2018, J. Cosmology Astropart. Phys., 4, 017
Gaztanaga, E., Barriga, J., Romeo, A., Fosalba, P., & Elizalde, E. 1998, Astrophysical letters & communications, 37
Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, ApJ, 622, 759
Hanson, D., et al. 2013, Physical Review Letters, 111, 141301
Hunter, J. D. 2007, Computing in Science Engineering, 9, 90
Ishino, H., et al. 2016, Proc. SPIE Int. Soc. Opt. Eng., 9904, 99040X
Jenet, F. A., & Anderson, S. B. 1998, PASP, 110, 1467
Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python
Kogut, A., et al. 2011, J. Cosmology Astropart. Phys., 7, 025
Lay, O. P., & Halverson, N. W. 2000, ApJ, 543, 787
Maris, M., Maino, D., Burigana, C., Mennella, A., Bersanelli, M., & Pasian, F. 2003, Mem. Soc. Astron. Italiana, 74, 488
Max, J. 1960, IRE Transactions on Information Theory, 6, 7
Max, J. 1960, IRE Transactions on Information Theory, 6, 7
Max, J. 1960, IRE Transactions on Information Theory, 6, 7
Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 1149
Planck Collaboration, et al. 2016, A&A, 594, A13
Planck Collaboration, et al. 2018, ArXiv e-prints, 1807.06205
Ruhl, J., et al. 2004, in Proc. SPIE, Vol. 5498, Z-Spec: a broadband millimeter-wave grating spectrometer: design, construction, and first cryogenic measurements, ed. C. M. Bradford, P. A. R. Ade, J. E. Aguirre, J. J. Bock, M. Dragovan, L. Duband, L. Earle, J. Glenn, H. Matsuhara, B. J. Naylor, H. T. Nguyen, M. Yun, & J. Zmuidzinas, 11–29
Schaffer, K. K., et al. 2011, ApJ, 743, 90
Smoot, G. F., et al. 1992, ApJ, 396, L1
Szapudi, I., Prunet, S., Pogosyan, D., Szalay, A. S., & Bond, J. R. 2001, ApJ, 548, L115
Tegmark, M. 1997, Phys. Rev. D, 56, 4514
Tegmark, M. 1998, in Eighteenth Texas Symposium on Relativistic Astrophysics, ed. A. V. Olinto, J. A. Frieman, & D. N. Schramm, 270
Tristram, M., & Ganga, K. 2007, Reports on Progress in Physics, 70, 899
Tristram, M., & Ganga, K. 2007, Reports on Progress in Physics, 70, 899
