

Implementation of Van Vleck Correction for the MWA

Pyxie Star

May 27, 2020

Quantization in the MWA digital signal pathway introduces non-linear artifacts into the data. Formulae for correcting these artifacts with a Van Vleck correction were presented in [4]. An implementation of the correction has been written into `pyuvdata` as an option when reading raw MWA correlator output files. This memo describes MWA quantization, re-derives the correction formulae, and discusses the implementation.

1 MWA Quantization

Quantization occurs in 3 stages of the MWA digital signal pathway, twice in the digital receivers and then in the correlator. It is this final quantization that the implemented Van Vleck correction addresses. All three quantization stages are described below in a summary of the digital signal pathway.

The analog signal from a tile is attenuated to ± 1 Volt, and a bandpass is applied to limit the frequency range to 80-300 MHz. This signal then goes to a digital receiver, described in [1], where it is sampled at 655.36 MHz and quantized to 8-bit values. The quantized data is then cast from real to complex, and also channelized into 256 1.28 MHz 'coarse' channels, by a polyphase filter bank (PFB). The PFB has an 8 tap subfilter followed by a 512 point FFT, and functions as follows. A Kaiser windowing function is applied to 4096 data samples. The 4096 windowed samples are split into 8 'phases' of 512 samples each, which are summed to result in 512 inputs to the FFT. The FFT outputs 256 complex values each consisting of a 16-bit real and 16-bit imaginary pair. Since each frequency channel is complex, that is, carrying two sampled values, the sampling rate is halved to 327.68 MHz. The incoming data is then shifted by 512 samples and the PFB procedure applied to this next grouping. After the FFT, a gain is applied to the data, and the second quantization occurs, taking the 16-bit real, 16-bit imaginary pair to a 5-bit real, 5-bit imaginary pair. For each observing session, some subset of 24 coarse channels are chosen and sent to the correlator.

The MWA correlator has two stages: a PFB which channelizes the data, and a cross-multiply and accumulate module to perform the correlations. The correlator PFB has 12-tap subfilter followed by a 128-point FFT, applying a Hanning window to the data

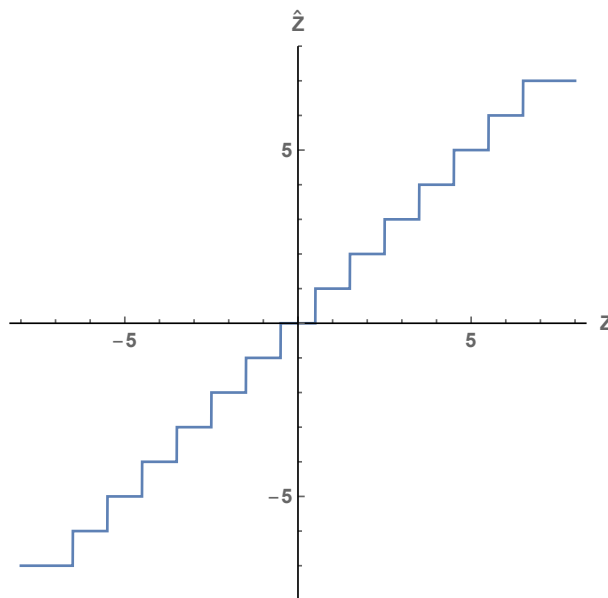


Figure 1: The 4-bit quantization pattern of the MWA.

stream. Outputs of this second PFB are quantized to 4-bit real, 4-bit imaginary pairs, and these integers are cast to floats for input into the cross-multiply. The correlator is described in greater detail in [2], and its PFB described in [3].

2 Van Vleck Correction

The final 4-bit quantization in the correlator is assumed to have a dominant impact, and so is addressed by the Van Vleck correction. That is, we effectively treat the values in the 4-bit quantization as being drawn from an analog zero-mean Gaussian distribution, ignoring the intermediate signal processing stages. Let Z be the analog signal from an antenna and \hat{Z} the quantized signal. The quantization pattern $\hat{Z}(Z)$ is shown in figure 1; essentially, all Z values within a certain range are placed into a certain quantization bin. After correlation, we no longer have access to \hat{Z} , and so the Van Vleck correction instead treats the statistics of \hat{Z} , assuming that we are in a regime in which we can approximate the auto and cross correlations as measuring variances and covariances as described in section 3.

First, consider the relation between the variance of \hat{Z} and the variance of Z . Following [4], the variance $E[\hat{Z}^2]$ is related to the probability of Z falling into each of the

quantization bins as follows

$$\begin{aligned}
E[\hat{Z}^2] &= 0^2 \cdot P(-0.5 < Z < 0.5) + (-1)^2 \cdot P(-1.5 < Z < -0.5) + 1^2 \cdot P(0.5 < Z < 1.5) + \dots \\
&\quad + (-7)^2 \cdot P(Z < -6.5) + 7^2 \cdot P(Z > 6.5) \\
&= 0^2 \cdot P(-0.5 < Z < 0.5) + 1^2 \cdot [P(-1.5 < Z < 1.5) - P(-0.5 < Z < 0.5)] + \dots \\
&\quad + 7^2 \cdot [1 - P(-6.5 < Z < 6.5)]
\end{aligned} \tag{1}$$

where we have taken advantage of the symmetry of the distribution of Z . With Z drawn from a Gaussian distribution, the probability that $Z \in [-a, a]$

$$P(Z \in [-a, a]) = \text{erf}\left(\frac{a}{\sigma_Z \sqrt{2}}\right) \tag{2}$$

where σ_Z is the standard deviation of Z . Combining terms gives

$$\begin{aligned}
E[\hat{Z}^2] &= 0^2 \cdot \text{erf}\left(\frac{0.5}{\sigma_Z \sqrt{2}}\right) + 1^2 \cdot \left[\text{erf}\left(\frac{1.5}{\sigma_Z \sqrt{2}}\right) - \text{erf}\left(\frac{0.5}{\sigma_Z \sqrt{2}}\right)\right] + \dots + 7^2 \cdot \left[1 - \text{erf}\left(\frac{6.5}{\sigma_Z \sqrt{2}}\right)\right] \\
&= (-1)\text{erf}\left(\frac{0.5}{\sigma_Z \sqrt{2}}\right) + (-3)\text{erf}\left(\frac{1.5}{\sigma_Z \sqrt{2}}\right) + \dots + (-13)\text{erf}\left(\frac{6.5}{\sigma_Z \sqrt{2}}\right) + 7^2 \\
&= 7^2 - \sum_{k=0}^6 (2k+1)\text{erf}\left(\frac{k+0.5}{\sigma_Z \sqrt{2}}\right).
\end{aligned} \tag{3}$$

This result can be written in terms of the standard deviation of \hat{Z} :

$$\hat{\sigma}_Z = \left[7^2 - \sum_{k=0}^6 (2k+1)\text{erf}\left(\frac{k+0.5}{\sigma_Z \sqrt{2}}\right)\right]^{1/2} \tag{4}$$

Now consider the impact of quantization on the covariance and correlation between two quantized \hat{Z}_1 and \hat{Z}_2 , where, since the distributions are zero-mean, the covariances

$$\kappa = E[Z_1 Z_2], \tag{5}$$

$$\hat{\kappa} = E[\hat{Z}_1 \hat{Z}_2]; \tag{6}$$

and the correlations

$$\rho = \frac{\kappa}{\sigma_1 \sigma_2}, \tag{7}$$

$$\hat{\rho} = \frac{\hat{\kappa}}{\hat{\sigma}_1 \hat{\sigma}_2}. \tag{8}$$

Again following [4], we use Prices's theorem, which, for two random variables X and Y , states

$$\frac{\partial \langle f(X, Y) \rangle}{\partial \langle XY \rangle} = \left\langle \frac{\partial f}{\partial X} \frac{\partial f}{\partial Y} \right\rangle \tag{9}$$

where brackets indicate expectation values. Using the function $f(Z_1, Z_2) = \hat{Z}_1 \hat{Z}_2$, we obtain

$$\frac{\partial \hat{\kappa}}{\partial \kappa} = \left\langle \frac{\partial \hat{Z}_1}{\partial Z_1} \frac{\partial \hat{Z}_2}{\partial Z_2} \right\rangle \quad (10)$$

For our quantization function, the derivative

$$\frac{\partial \hat{Z}}{\partial Z} = \delta(Z - (-6.5)) + \delta(Z - (-5.5)) + \dots + \delta(Z - 5.5) + \delta(Z - 6.5) \quad (11)$$

so

$$\frac{\partial \hat{Z}_1}{\partial Z_1} \frac{\partial \hat{Z}_2}{\partial Z_2} = \sum_{i=-7}^6 \sum_{j=-7}^6 \delta(Z_1 - (i + 0.5)) \delta(Z_2 - (j + 0.5)) \quad (12)$$

To find the expectation value, we integrate with the joint normal probability density function

$$\begin{aligned} \left\langle \frac{\partial \hat{Z}_1}{\partial Z_1} \frac{\partial \hat{Z}_2}{\partial Z_2} \right\rangle &= \sum_{i=-7}^6 \sum_{j=-7}^6 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dZ_1 dZ_2 \delta(Z_1 - (i + 0.5)) \delta(Z_2 - (j + 0.5)) \\ &\quad \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{Z_1^2}{\sigma_1^2} + \frac{Z_2^2}{\sigma_2^2} - \frac{2\rho Z_1 Z_2}{\sigma_1\sigma_2} \right) \right] \end{aligned} \quad (13)$$

to get

$$\begin{aligned} \left\langle \frac{\partial \hat{Z}_1}{\partial Z_1} \frac{\partial \hat{Z}_2}{\partial Z_2} \right\rangle &= \sum_{i=-7}^6 \sum_{j=-7}^6 \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{(i+0.5)^2}{\sigma_1^2} + \frac{(j+0.5)^2}{\sigma_2^2} \right. \right. \\ &\quad \left. \left. - \frac{2\rho(i+0.5)(j+0.5)}{\sigma_1\sigma_2} \right) \right] \end{aligned} \quad (14)$$

Writing $\rho = \kappa/\sigma_1\sigma_2$, equation 10 becomes

$$\begin{aligned} \hat{\kappa} &= \sum_{i=-7}^6 \sum_{j=-7}^6 \int_0^\rho d\rho' \frac{1}{2\pi\sqrt{1-\rho'^2}} \exp \left[-\frac{1}{2(1-\rho'^2)} \left(\frac{(i+0.5)^2}{\sigma_1^2} + \frac{(j+0.5)^2}{\sigma_2^2} \right. \right. \\ &\quad \left. \left. - \frac{2\rho'(i+0.5)(j+0.5)}{\sigma_1\sigma_2} \right) \right] \end{aligned} \quad (15)$$

3 Implementation

Let the correlator inputs from antennas 1 and 2, with arbitrary polarization, be \hat{Z}_1 and \hat{Z}_2 , where \hat{Z} is a sum of a sky signal and receiver noise: $\hat{Z} = S + N$. S and N are both complex circular Gaussian random variables. That is, the real and imaginary parts are each Gaussian distributed with zero mean, and are uncorrelated with each

other. Alternatively, we can think of this as there being no preferred phase in either the sky signal or in the receiver noise. We additionally assume that \hat{Z} itself is a complex circular Gaussian random variable, which might require assumptions about the lack of correlation between sky signal and instrument noise (see Appendix A).

In the limit of large time integration t , we can approximate the autocorrelation of \hat{Z} as the variance:

$$\frac{1}{t} \sum_i^t \hat{Z}_i \hat{Z}_i^* \rightarrow E[\hat{Z} \hat{Z}^*]. \quad (16)$$

Since Z is circular random, in this limit the autocorrelation is related to the variance of the real and imaginary parts of Z :

$$\frac{1}{t} \sum_i^t \hat{Z}_i \hat{Z}_i^* = 2E[\text{Re}(\hat{Z})^2] = 2E[\text{Im}(\hat{Z})^2]. \quad (17)$$

Also in the limit of large t , the cross-correlation can be approximated as the covariance:

$$\frac{1}{t} \sum_i^t \hat{Z}_{1_i} \hat{Z}_{2_i}^* \rightarrow E[\hat{Z}_1 \hat{Z}_2^*]. \quad (18)$$

The covariance can be expanded and simplified:

$$\begin{aligned} E[\hat{Z}_1 \hat{Z}_2^*] &= E[\text{Re}(\hat{Z}_1) \text{Re}(\hat{Z}_2)] + E[\text{Im}(\hat{Z}_1) \text{Im}(\hat{Z}_2)] \\ &\quad + iE[\text{Im}(\hat{Z}_1) \text{Re}(\hat{Z}_2)] - iE[\text{Re}(\hat{Z}_1) \text{Im}(\hat{Z}_2)] \\ &= 2E[\text{Re}(\hat{Z}_1) \text{Re}(\hat{Z}_2)] + 2iE[\text{Im}(\hat{Z}_1) \text{Re}(\hat{Z}_2)] \end{aligned} \quad (19)$$

where the following relations for circular complex random variables have been used, which may require the assumption that $\hat{Z}_1 = S + N_1$ and $\hat{Z}_2 = S + N_2$ are drawn from the same distribution.

$$E[\text{Re}(\hat{Z}_1) \text{Re}(\hat{Z}_2)] = E[\text{Im}(\hat{Z}_1) \text{Im}(\hat{Z}_2)] \quad (20)$$

$$E[\text{Im}(\hat{Z}_1) \text{Re}(\hat{Z}_2)] = -E[\text{Re}(\hat{Z}_1) \text{Im}(\hat{Z}_2)] \quad (21)$$

To correct a cross-correlation between antennas 1 and 2, the first step is to find σ_1 and σ_2 . For antenna 1, the quantized standard deviation $\hat{\sigma}_1$ of the real and imaginary parts of Z_1 can effectively be found by taking the antenna's autocorrelation, dividing it by 2, and taking the square root. Equation 4 can be used to find σ_1 from $\hat{\sigma}_1$. This σ_1 , along with σ_2 calculated in the same way, are then used in equation 15 to correct the real and imaginary parts of the cross correlation. The real and imaginary parts of the cross correlation are also divided by 2 before correcting.

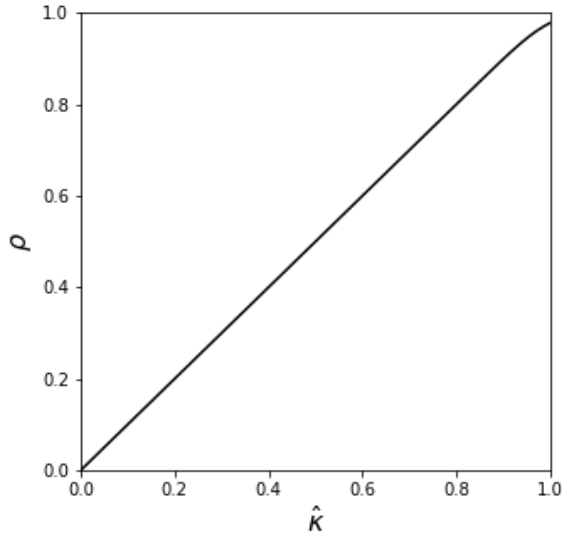


Figure 2: The cross-correlation Van Vleck correction function (equation 15) plotted with $\sigma_1 = 1.0, \sigma_2 = 1.0$.

The Van Vleck correction algorithm was implemented in `pyuvdata`. To verify the algorithm, equations 4 and 15 were plotted, as shown in figures 2 and 3. Comparison plots were then generated using the implementations of these functions in the code, verifying the functions' accuracy. Also, data corrected by the algorithm was taken back through equations 4 and 15, showing that the calculated quantized data matched the initial uncorrected input.

The correction algorithm has a high computational cost associated with equation 15, which has 14^2 exponential terms in the summation. By combining redundant terms, this number can be reduced by a factor of 2. Evaluation of the integral using `scipy.integrate.quad` requires 750 computations of the summation. Using a `scipy` root-finding function proved useful for finding the input of equation 4 given some output. However, this same function is too costly when correcting the cross-correlations, even though only 5 evaluations of equation 15 were required for the root-finding function to converge. An additional cost is memory usage, as running the root-finding function requires generation of a Jacobian matrix the size of which is dictated by the number of cross-correlations being corrected. This places a limit on the number of values which can be corrected simultaneously.

There are several possible solution for mitigating the cost of correcting the cross-correlations. One is to implement a `scipy` integration function which allows for vectorized integral evaluation, though this has potential for loss in precision. Additionally,

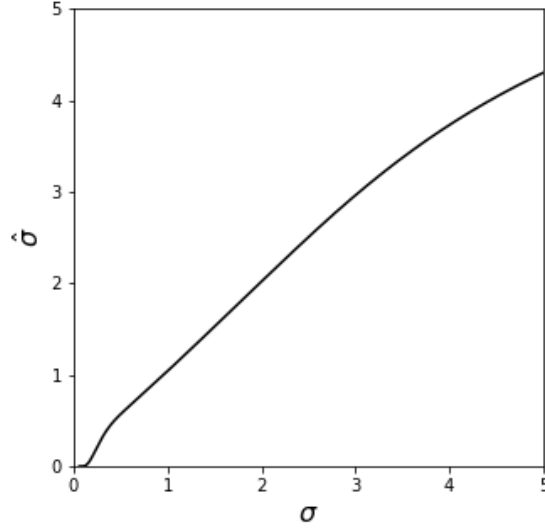


Figure 3: The auto-correlation Van Vleck correction function (equation 4).

moving from a root-finding algorithm to a lookup table might be more computationally efficient. Both root-finding and a lookup table may ultimately require parallelization to be feasible. Another, perhaps more complicated but also potentially interesting, alternative is to exploit the role of the corrected auto-correlations in equation 15 and use the shapes of the correction on the autos to determine the shape of the correction on the respective cross-correlation.

A Circular Complex Gaussian Random Variables

A circular complex random variable $Z = \text{Re}(Z) + i \text{Im}(Z)$ is defined as having a distribution that is invariant under a random phase shift:

$$E[Z] = E[e^{i\phi} Z] = 0, \quad (22)$$

$$E[ZZ] = E[e^{i\phi} Z e^{i\phi} Z], \quad (23)$$

$$E[ZZ^*] = E[e^{i\phi} Z e^{-i\phi} Z^*]. \quad (24)$$

The last of these conditions is trivial. The first and second are only satisfied if $E[Z] = E[ZZ] = 0$. Thus

$$E[ZZ] = 0 \quad (25)$$

$$= E[(\text{Re}(Z) + i \text{Im}(Z))(\text{Re}(Z) + i \text{Im}(Z))] \quad (26)$$

$$= E[\text{Re}(Z)^2] - E[\text{Im}(Z)^2] + iE[\text{Re}(Z) \text{Im}(Z)] + iE[\text{Im}(Z) \text{Re}(Z)] \quad (27)$$

and

$$E[\text{Re}(Z)^2] = E[\text{Im}(Z)^2], \quad (28)$$

$$E[\text{Re}(Z) \text{Im}(Z)] = -E[\text{Im}(Z) \text{Re}(Z)] = 0. \quad (29)$$

For a complex random variable, the variance $\sigma_Z^2 = E[ZZ^*] = E[\text{Re}(Z)^2] + E[\text{Im}(Z)^2]$. So in the case of a circularly complex random variable

$$\sigma_Z^2 = 2E[\text{Re}(Z)^2] = 2E[\text{Im}(Z)^2]. \quad (30)$$

Two circular complex random variables sampled from the same distribution have the following relations

$$E[\text{Re}(Z_1) \text{Re}(Z_2)] = E[\text{Im}(Z_1) \text{Im}(Z_2)] \quad (31)$$

$$E[\text{Re}(Z_1) \text{Im}(Z_2)] = -E[\text{Im}(Z_1) \text{Re}(Z_2)] \quad (32)$$

If this relation does not hold for circular complex random variables sampled from different distributions, then some assumptions are required for the correlator input model.

For example, consider the case of a correlator input $Z = S + N$, where S and N are both circular complex Gaussian random variables, that is, in addition to the conditions above, the real and imaginary parts of S and N are sampled from zero-mean Gaussian distributions. To treat Z as circularly complex requires $E[\text{Re}(Z) \text{Im}(Z)] = 0$. Thus

$$E[\text{Re}(S) \text{Im}(S)] + E[\text{Re}(N) \text{Im}(N)] + E[\text{Re}(S) \text{Im}(N)] + E[\text{Re}(N) \text{Im}(S)] = 0. \quad (33)$$

The first two terms vanish according to (29). If (32) applies to circular complex random variables sampled from different distributions, the last two terms vanish. If the relation does not apply, then it is required that $E[\text{Re}(S) \text{Im}(N)] = E[\text{Re}(N) \text{Im}(S)] = 0$, that is, that the sky signal and receiver noise are uncorrelated.

References

- ¹T. Prabu et al., “A digital-receiver for the murchison widefield array”, *Experimental Astronomy* **39**(1), 73–93 (2015).
- ²S. M. Ord et al., “The murchison widefield array correlator”, *Publications of the Astronomical Society of Australia* **32** (2015).
- ³S. J. McSweeney et al., *Mwa tied-array processing iii: microsecond time resolution via a polyphase synthesis filter*, 2020 PREPRINT.
- ⁴L. V. Benkevitch et al., *Van vleck correction generalization for complex correlators with multilevel quantization*, 2015.
- ⁵A. M. Levine, *An analysis of two cascaded fourier transforms and application to the coarse and fine pfb for the mwa*, 2012.

⁶A. M. Levine, *The effects of the two-stage pfb architecture on the response of the mwa*, 2012.

⁷A. Beardsley, “The murchison widefield array 21cm epoch of reionization experiment: design, construction, and first season results”, PhD thesis (University of Washington, Seattle, WA, 2015).