A figure of merit for constraining instrumental polarization systematics in the delay spectrum

Zachary Martinot

August 15, 2017

Abstract

A figure of merit for comparing the polarization leakage of different antenna designs is derived. It is based on the smallest upper bound on the coupling of the polarization state of the sky to Vokes-I. Additionally, a symmetry of Mueller matrices is derived. This allows us to see how this figure of merit might be measured in practice. Further, it allows us to see explicitly how this figure relates to the IXR.

1 Introduction

The set of measured visibilities $\{\mathcal{V}_{ee}, \mathcal{V}_{en}, \mathcal{V}_{ne}, \mathcal{V}_{nn}\}$ obtained from the correlations of antenna feed polarizations e and n is described by

$$\boldsymbol{\mathcal{V}} = \begin{pmatrix} \mathcal{V}_{ee} & \mathcal{V}_{en} \\ \mathcal{V}_{ne} & \mathcal{V}_{ne} \end{pmatrix} = \int_{S^2} \boldsymbol{J}\boldsymbol{\mathcal{C}}\boldsymbol{J}^\dagger e^{-2\pi i \frac{\nu}{c} \vec{\mathbf{b}}\cdot\hat{\mathbf{s}}} \tag{1}$$

where $\boldsymbol{J} = \boldsymbol{J}(\nu, \hat{\mathbf{s}})$ is the frequency ν and position $\hat{\mathbf{s}}$ dependent instrumental Jones matrix, and $\boldsymbol{\mathcal{C}}$ is the polarized brightness tensor on the sky:

$$\boldsymbol{\mathcal{C}}(\nu, \hat{\mathbf{s}}) = I_c \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} I_f + Q & U - iV \\ U + iV & I_f - Q \end{bmatrix}. \tag{2}$$

Since the cosmological signal is (effectively) unpolarized, we can think of the sky as the sum of the scalar cosmological intensity I_c and the the polarized foregrounds with intensity I_f and polarization state described by Q, U, V. The total sky brightness is

$$I = I_c + I_f. \tag{3}$$

We then define a set of parameters analogous to the Stokes parameters; the Vokes parameters:

$$\mathcal{V}_I = \mathcal{V}_{ee} + \mathcal{V}_{nn} \tag{4}$$
$$\mathcal{V}_Q = \mathcal{V}_{ee} - \mathcal{V}_{nn} \tag{5}$$
$$\mathcal{V}_U = \mathcal{V}_{en} + \mathcal{V}_{ne} \tag{6}$$
$$\mathcal{V}_V = -i(\mathcal{V}_{en} - \mathcal{V}_{ne}). \tag{7}$$

These quantities may also be described as

$$\mathcal{V}_k = \frac{1}{2} \mathrm{Tr}(\boldsymbol{\sigma}_k \boldsymbol{\mathcal{V}}) \tag{8}$$
$$= \int_{S^2} (M_{k0} I + M_{k1} Q + M_{k2} U + M_{k3} V) e^{-2\pi i \frac{\nu}{c} \vec{b}\cdot\hat{\mathbf{s}}} \tag{9}$$

where k indexes the set of Pauli matrices with their indices permuted from the common Quantum Mechanics convention:

$$\boldsymbol{\sigma}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \boldsymbol{\sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \boldsymbol{\sigma}_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{10}$$

Here the (real-valued) functions M_{jk} are the the elements of the Mueller matrix $\boldsymbol{M}$:

$$M_{jk} = \frac{1}{2} \mathrm{Tr}\big(\boldsymbol{\sigma}_j \boldsymbol{J} \boldsymbol{\sigma}_k \boldsymbol{J}^\dagger\big) \tag{11}$$

$$\boldsymbol{M} = \begin{bmatrix} M_{00} & M_{01} & M_{02} & M_{03} \\ M_{10} & M_{11} & M_{12} & M_{13} \\ M_{20} & M_{21} & M_{22} & M_{23} \\ M_{30} & M_{31} & M_{32} & M_{33} \end{bmatrix} \tag{12}$$

To estimate the power spectrum of the unpolarized cosmological signal we will use the delay transform of $\mathcal{V}_I$. In general $\boldsymbol{J}$ will couple the polarization state of the sky – and thus any frequency structure it has – to Vokes-I. Additionally, this instrumental coupling is generally a function of frequency which may be distinct from that of the power pattern, and even a smooth spectrum polarization state on the sky could couple any such structure to Vokes-I. Thus it is desirable to minimize this instrumental polarization as much as possible.

2 A bound on instrumental polarization leakage

Here we will derive a figure of merit that constrains the polarization of an instrument, given its Jones matrix $\boldsymbol{J}$. Ideally, we would build an instrument that is insensitive to the polarization state of incident radiation. But at the meter-wavelengths of our observations this may be impractical. Never the less, we can consider how close we are to this idealization. The difference between a polarization insensitive Vokes-I

$$\mathcal{V}_u = \int_{S^2} M_{00} I e^{-2\pi i \frac{\nu}{c} \vec{b} \cdot \hat{\mathbf{s}}} \tag{13}$$

and the actual Vokes-I $\mathcal{V}_I$ is

$$\mathcal{V}_I - \mathcal{V}_u = \int_{S^2} (M_{01} Q + M_{02} U + M_{03} V) e^{-2\pi i \frac{\nu}{c} \vec{b} \cdot \hat{\mathbf{s}}} \tag{14}$$

which depends sensitively on the details of the polarization state of the sky. However, we can estimate the integral as follows:

$$\begin{aligned}
|\mathcal{V}_I - \mathcal{V}_u| &= \left| \int_{S^2} (M_{01} Q + M_{02} U + M_{03} V) e^{-2\pi i \frac{\nu}{c} \vec{b} \cdot \hat{\mathbf{s}}} \right| && (15) \\
&\leq \int_{S^2} |M_{01} Q + M_{02} U + M_{03} U| \quad \text{by the triangle inequality for integrals,} && (16) \\
&\leq \int_{S^2} \left(\sqrt{M_{01}^2 + M_{02}^2 + M_{03}^2} \right) \left(\sqrt{Q^2 + U^2 + V^2} \right) && (17) \\
&\leq \max_{S^2} \{I_p\} \int_{S^2} \sqrt{M_{01}^2 + M_{02}^2 + M_{03}^2} && (18)
\end{aligned}$$

where $I_p = \sqrt{Q^2 + U^2 + V^2}$ is the flux of polarized radiation on the sky. In going from the second to the third line we have used the Cauchy-Schwarz inequality: $|\vec{x} \cdot \vec{y}| \leq |\vec{x}||\vec{y}|$ with $\vec{x} = (M_{01}, M_{02}, M_{03})$ and $\vec{y} = (Q, U, V)$, and the fact that $\int f(x) \leq \int g(x)$ if $f(x) \leq g(x)$ for all x.

Thus, in the limit $\int_{S^2} \sqrt{M_{01}^2 + M_{02}^2 + M_{03}^2} \to 0$, we have $|\mathcal{V}_I - \mathcal{V}_u| \to 0$ in which case Vokes-I, and thus any power spectrum estimate derived from it, is independent of polarization on the sky:

$$\mathcal{V}_I = \int_{S^2} M_{00} I e^{-2\pi i \frac{\nu}{c} \vec{b} \cdot \hat{\mathbf{s}}} \tag{19}$$

Since the function $\sqrt{M_{01}^2 + M_{02}^2 + M_{03}^2}$ is real and positive this condition corresponds to it being identically zero everywhere. This would be the case of an instrument that is a perfect polarimeter at every point on the sky simultaneously, and is not realistic. But it does allow us to impose an ordering on different antenna designs by a guaranteed limit on polarization leakage that is independent of the polarization state of the sky. Defining the functional

$$\Upsilon[\boldsymbol{J}] = \int_{S^2} \sqrt{M_{01}^2 + M_{02}^2 + M_{03}^2} \tag{20}$$

as the "leakage bound", we can then compare the quality of two different antenna designs with different Jones matrices $\boldsymbol{J}_a$ and $\boldsymbol{J}_b$. If

$$\Upsilon[\boldsymbol{J}_a] < \Upsilon[\boldsymbol{J}_b], \tag{21}$$

then we say that the polarization quality of the antenna described by $\boldsymbol{J}_a$ is better than the polarization quality of the antenna described by $\boldsymbol{J}_b$.

Additionally, given an upper bound on the level of polarized flux on the sky, this allows one to set a crude specification for the leakage bound as a function of frequency using equation 18.

A basic code to compute equation 20 from a CST simulation of an instrument can be found here: https://github.com/zacharymartinot/PolLeakageMetric/

2.1 Delay spectrum bound

The leakage bound is a function of frequency, $\Upsilon(\nu) = \Upsilon[\boldsymbol{J}(\nu)]$, and the delay spectrum is taken over a fixed bandwidth. To obtain a bound on the polarization leakage to the delay spectrum computed over a bandwidth $\mathcal{B}$, we could take the maximum of Υ over that bandwidth,

$$\widehat{\Upsilon}[\mathcal{B}] = \max_{\mathcal{B}} \{\Upsilon(\nu)\} \tag{22}$$

since the delay spectrum leakage is bounded by

$$\left| \widetilde{\mathcal{V}}_I(\tau) - \widetilde{\mathcal{V}}_u(\tau) \right| \leq \int_{\mathcal{B}} \left(\Upsilon(\nu) \max_{S^2} \{I_p(\nu)\} \right) \tag{23}$$

$$\leq \widehat{\Upsilon}[\mathcal{B}] \int_{\mathcal{B}} \max_{S^2} \{I_p(\nu)\} \tag{24}$$

For a bound on the power spectrum, square both sides:

$$\left| \widetilde{\mathcal{V}}_I(\tau) - \widetilde{\mathcal{V}}_u(\tau) \right|^2 \leq \widehat{\Upsilon}[\mathcal{B}]^2 \left(\int_{\mathcal{B}} \max_{S^2} \{I_p(\nu)\} \right)^2 \tag{25}$$

3 A symmetry of Mueller matrices

Because the Mueller matrix is derived from a Jones matrix which is in turn derived from EM simulations by a reciprocity argument there is an expectation that the Mueller matrix should express some symmetry that reflects this. The hope that the matrix is symmetric is not true in general (although it might be nearly symmetric in particular instances). However, we will see that there is a symmetry of the matrix which might be useful for making a measurement of the leakage bound.

The integrand of the leakage bound (or "point-leakage bound") is the square-root of the sum of the squares of the off-diagonal components of the first **row** in equation 12:

$$M_p = \sqrt{M_{01}^2 + M_{02}^2 + M_{03}^2}. \tag{26}$$

We can form a similar quantity from the components of the first **column** that describe the spurious polarization that is apparently measured even when observing an unpolarized source:

$$M_u = \sqrt{M_{10}^2 + M_{20}^2 + M_{30}^2} \tag{27}$$

It turns out that the two quantities M_p and M_u are equal in general, which we prove now:

Proof. First, we will be using the cyclic permutation property of the trace of a product of matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$:

$$\mathrm{Tr}(\boldsymbol{ABC}) = \mathrm{Tr}(\boldsymbol{CAB}) = \mathrm{Tr}(\boldsymbol{BCA}) \tag{28}$$

Using this property and equation 11 the point-leakage bound can be written in terms of the Jones matrix as

$$\begin{aligned}
M_p^2 &= M_{01}^2 + M_{02}^2 + M_{03}^2 && (29)\\
&= \sum_{k=1}^{3} \frac{1}{4}\,\mathrm{Tr}\big(\boldsymbol{\sigma}_0 \boldsymbol{J} \boldsymbol{\sigma}_k \boldsymbol{J}^\dagger\big)^2 && (30)\\
&= \sum_{k=1}^{3} \frac{1}{4}\,\mathrm{Tr}\big(\boldsymbol{J} \boldsymbol{\sigma}_k \boldsymbol{J}^\dagger\big)^2 && (31)\\
&= \sum_{k=1}^{3} \frac{1}{4}\,\mathrm{Tr}\big(\boldsymbol{\sigma}_k \boldsymbol{J}^\dagger \boldsymbol{J}\big)^2 && (32)
\end{aligned}$$

Similarly, the bound on the spurious response to unpolarized radiation is

$$\begin{aligned}
M_u^2 &= M_{10}^2 + M_{20}^2 + M_{30}^2 && (33)\\
&= \sum_{k=1}^{3} \frac{1}{4}\,\mathrm{Tr}\big(\boldsymbol{\sigma}_k \boldsymbol{J} \boldsymbol{\sigma}_0 \boldsymbol{J}^\dagger\big)^2 && (34)\\
&= \sum_{k=1}^{3} \frac{1}{4}\,\mathrm{Tr}\big(\boldsymbol{\sigma}_k \boldsymbol{J} \boldsymbol{J}^\dagger\big)^2. && (35)
\end{aligned}$$

Now we observe that both $\boldsymbol{J}\boldsymbol{J}^\dagger$ and $\boldsymbol{J}^\dagger\boldsymbol{J}$ are Hermitian matrices which when considered as a vector space with the Hilbert-Schmidt norm as the inner product are spanned by the Pauli matrices in equation 10. Thus we can write them as

$$\begin{aligned}
\boldsymbol{J}\boldsymbol{J}^\dagger &= \sum_k \mathrm{Tr}\big(\boldsymbol{\sigma}_k \boldsymbol{J}\boldsymbol{J}^\dagger\big)\boldsymbol{\sigma}_k && (36)\\
&= \begin{bmatrix} M_{00} + M_{10} & M_{20} - iM_{30} \\ M_{20} + iM_{30} & M_{00} - M_{10} \end{bmatrix} && (37)
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{J}^\dagger\boldsymbol{J} &= \sum_k \mathrm{Tr}\big(\boldsymbol{\sigma}_k \boldsymbol{J}^\dagger\boldsymbol{J}\big)\boldsymbol{\sigma}_k && (38)\\
&= \begin{bmatrix} M_{00} + M_{01} & M_{02} - iM_{03} \\ M_{02} + iM_{03} & M_{00} - M_{01} \end{bmatrix} && (39)
\end{aligned}$$

The two eigenvalues of any 2x2 matrix $\boldsymbol{H}$ can be expressed generally as

$$\lambda_{\pm} = \frac{1}{2}\left(\mathrm{Tr}(\boldsymbol{H}) \pm \sqrt{\mathrm{Tr}(\boldsymbol{H})^2 - 4\det(\boldsymbol{H})}\right) \tag{40}$$

So the eigenvalues of $\boldsymbol{J}\boldsymbol{J}^\dagger$ are

$$\alpha_{\pm} = \frac{1}{2}\left(\mathrm{Tr}\big(\boldsymbol{J}\boldsymbol{J}^\dagger\big) \pm \sqrt{\mathrm{Tr}(\boldsymbol{J}\boldsymbol{J}^\dagger)^2 - 4\det(\boldsymbol{J}\boldsymbol{J}^\dagger)}\right) \tag{41}$$

and the eigenvalues of $\boldsymbol{J}^\dagger\boldsymbol{J}$ are

$$\beta_{\pm} = \frac{1}{2}\left(\mathrm{Tr}\big(\boldsymbol{J}^\dagger\boldsymbol{J}\big) \pm \sqrt{\mathrm{Tr}(\boldsymbol{J}^\dagger\boldsymbol{J})^2 - 4\det(\boldsymbol{J}^\dagger\boldsymbol{J})}\right) \tag{42}$$

Then since

$$\frac{1}{2}\,\mathrm{Tr}\big(\boldsymbol{J}\boldsymbol{J}^\dagger\big) = \frac{1}{2}\,\mathrm{Tr}\big(\boldsymbol{J}^\dagger\boldsymbol{J}\big) = M_{00} \tag{43}$$

and

$$\det\big(\boldsymbol{J}\boldsymbol{J}^\dagger\big) = \det\big(\boldsymbol{J}^\dagger\boldsymbol{J}\big) = |\det(\boldsymbol{J})|^2 \tag{44}$$

we have that $\alpha_{\pm} = \beta_{\pm}$. On the other hand by direct evaluation from the matrices in equations 37 and 39, we have

$$\alpha_{\pm} = M_{00} \pm M_p \tag{45}$$
$$\beta_{\pm} = M_{00} \pm M_u \tag{46}$$

and hence

$$\alpha_{+} = \beta_{+} \tag{47}$$
$$M_{00} + M_p = M_{00} + M_u \tag{48}$$
$$M_p = M_u \tag{49}$$

□

Why might this be useful? The components M_{01}, M_{02}, M_{03} are important because they determine the extent to which polarized foregrounds are able to contaminate the power spectrum, but the degree to which these matrix components might be measured directly is not well understood. On the other hand, while the components M_{10}, M_{20}, M_{30} do not enter the power spectrum directly, the fact that they describe the response to unpolarized emission means that they might be measured using the same (or similar) techniques that are used to characterize the power pattern of an instrument. This means that we can obtain a limit on polarization leakage by measuring the response of an instrument to unpolarized sources, which is probably a more practical/at-all-possible kind of experiment to do.

4 Relation to the IXR

The Intrinsic Cross-polarization Ratio (IXR) was developed in [1] with the precise measurement of the polarization state of an astronomical source in mind as the goal of the instrument. As such, the authors based the metric on the ability to invert a Jones or Mueller matrix. Our goal is different, and consequently it was not immediately obvious that the IXR would constrain the instrument in a way that was optimal for reducing potential polarization contamination in the power spectrum. By going back to the definition of the visibility and working forwards, we can be sure that our figure of merit is properly constraining the instrument in the way that is optimal for our measurements. But of course, our desire for a metric that is independent of the polarization state of the sky turns out to be mathematically similar to the desire to produce a basis independent, point-by-point metric of matrix invertibility. As such, the point-leakage bound is closely related to the IXR_J for the Jones matrix. Recall that the IXR_J is defined in as

$$IXR_J = \left(\frac{s_1 + s_2}{s_1 - s_2}\right)^2 \tag{50}$$

where s_1 and s_2 are the singular values of the Jones matrix $\boldsymbol{J}$. But the singular values of $\boldsymbol{J}$ are equivalent to $\sqrt{\lambda_+}$ and $\sqrt{\lambda_-}$, the eigenvalues of the Hermitian matrix $\boldsymbol{J}^\dagger\boldsymbol{J}$ (and of $\boldsymbol{J}\boldsymbol{J}^\dagger$, as shown above) which means we can rewrite the IXR_J as

$$IXR_J = \left(\frac{\sqrt{\lambda_+} + \sqrt{\lambda_-}}{\sqrt{\lambda_+} - \sqrt{\lambda_-}}\right)^2 \tag{51}$$

$$= \left(\frac{\sqrt{M_{00}} + \sqrt{M_{00}}}{\sqrt{M_{00} + M_p} - \sqrt{M_{00} - M_p}}\right)^2 \tag{52}$$

$$= 4M_{00}\left(\frac{1}{\sqrt{M_{00} + M_p} - \sqrt{M_{00} - M_p}}\right)^2. \tag{53}$$

So the IXR_J is a function that is M_{00} (the sum of the power-patterns of the two feeds), times a function that goes to infinity in places where the Jones matrix $\boldsymbol{J}$ is proportional to the identity matrix, that is, when $M_p = M_u = 0$. However, this is not a suitable function to integrate over the sphere as the integral will formally diverge for any instrumental Jones matrix that is defined to be the identity somewhere on the sky (usually zenith). At the same time it is "good" in precisely the places that M_p is "good". Its just that the values that correspond to "good" and "bad" are inversely related between the two quantities.

The point here is that optimizing M_p for polarization rejection will generally lead to a similarly optimized IXR, but the IXR is not necessarily a useful metric for the delay spectrum analysis method.

References

[1] Carozzi, T. D., and G. Woan. "A fundamental figure of merit for radio polarimeters." IEEE Transactions on Antennas and Propagation 59.6 (2011): 2058-2065.