Technical Notes on Null Stream Polarization Test

Isaac C. F. Wong, Peter T. H. Pang, Rico K. L. Lo

June 23, 2020

1 Null Stream

Consider a D-detector observation model in frequency domain,

$$\tilde{\mathbf{d}} = \mathbf{F}\tilde{\mathbf{h}} + \tilde{\mathbf{n}} \tag{1}$$

where $\tilde{\mathbf{d}} \in \mathbb{C}^{D\times K}$ is the observed data strain, $\mathbf{F} \in \mathbb{R}^{D\times M}$ is the antenna response function, $\tilde{\mathbf{h}} \in \mathbb{C}^{M\times K}$ is the polarization modes, $\tilde{\mathbf{n}} \in \mathbb{C}^{D\times K}$ is the detector noise, M is the number of polarization modes, and K is the number of frequency bins. Then divide each side of (4) by $\sqrt{\frac{1}{2\delta f}S[k]}$ where $S[k]$ is the one-sided noise power spectral density to obtain the noise-weighed quantities.

$$\tilde{\mathbf{d}}_w = \mathbf{F}_w\tilde{\mathbf{h}} + \tilde{\mathbf{n}}_w \tag{2}$$

where $\tilde{\mathbf{d}}_w$ is the noise-weighed data strain, $\mathbf{F}_w$ is the noise-weighed antenna response function, and $\tilde{\mathbf{n}}_w$ is the noise-weighed detector noise. Notice that $\mathbf{F}_w = \mathbf{F}_w(\alpha,\delta,\psi) \in \mathbb{R}^{D\times M\times K} = \mathbf{F}/\sqrt{\frac{1}{2\delta f}S}$ where α is the right ascension, δ is the declination, and ψ is the polarization angle. One can construct a **null projector $\mathbf{P}$** from $\mathbf{F}_w$

$$\mathbf{P}(\alpha,\delta) = \mathbf{I} - \mathbf{F}_w(\mathbf{F}_w^T\mathbf{F}_w)^{-1}\mathbf{F}_w^T \tag{3}$$

Notice that the construction of $\mathbf{P}$ is independent of ψ since ψ represents the rotation of the column vectors of $\mathbf{F}_w$ within the subspace spanned by the column vectors. One should notice that $\mathbf{F}_w \in \mathbf{R}^{D\times M\times K}$, and the matrix operation in (3) is done for each frequency index.
Suppose the source comes from $(\alpha_{\mathrm{true}},\delta_{\mathrm{true}})$ i.e.

$$\tilde{\mathbf{d}} = \mathbf{F}(\alpha_{\mathrm{true}},\delta_{\mathrm{true}})\tilde{\mathbf{h}} + \tilde{\mathbf{n}} \tag{4}$$

, then

$$\mathbf{P}(\alpha_{\mathrm{true}},\delta_{\mathrm{true}})\tilde{\mathbf{d}} = \mathbf{P}(\alpha_{\mathrm{true}},\delta_{\mathrm{true}})\tilde{\mathbf{n}} \tag{5}$$

with no signal content remains. We can conveniently define the null energy $E_{\rm null}$

$$E_{\rm null}(\alpha,\delta) = \sum_k \tilde{\mathbf{d}}^\dagger \mathbf{P}(\alpha,\delta)\tilde{\mathbf{d}} \tag{6}$$

where k is the frequency index. And

$$E_{\rm null}(\alpha_{\rm true},\delta_{\rm true}) = \sum_k \tilde{\mathbf{n}}^\dagger \mathbf{P}(\alpha_{\rm true},\delta_{\rm true})\tilde{\mathbf{n}} \tag{7}$$

where $2E_{\rm null}(\alpha_{\rm true},\delta_{\rm true})$ follows the χ^2 distribution with degree of freedom $\mathrm{DoF} = 2(D-M)K$. The likelihood function is thus in the form

$$p(\tilde{\mathbf{d}}|\alpha,\delta;\mathcal{H}) = \chi^2_{\rm DoF}(2E_{\rm null}(\alpha,\delta)) \tag{8}$$

where $\chi^2_{\rm DoF}(\cdot)$ denotes the χ^2 probability density function.

2 Polarization Test with Null Stream

The polarization hypotheses are encoded in the $\mathbf{F}$ assumed to construct null projectors. The different $\mathbf{F}$ being assumed and the implied time delay from the sky position together will resolve different polarization models regardless of the waveform $\tilde{\mathbf{h}}$. For example

$$\mathcal{H}_{\rm tensor} \to \mathbf{F} = \begin{bmatrix}\mathbf{f}_+ & \mathbf{f}_\times\end{bmatrix} \tag{9}$$

$$\mathcal{H}_{\rm vector} \to \mathbf{F} = \begin{bmatrix}\mathbf{f}_x & \mathbf{f}_y\end{bmatrix} \tag{10}$$

$$\mathcal{H}_{\rm scalar} \to \mathbf{F} = \begin{bmatrix}\mathbf{f}_b & \mathbf{f}_l\end{bmatrix} \tag{11}$$

For the case of scalar hypothesis, since $\mathbf{f}_b$ and $\mathbf{f}_l$ are collinear, we only need to pick either of them to construct the null projector. Then we can compute the evidence of each polarization hypothesis

$$Z_{\mathcal{H}} = \int p(\mathbf{d}|\vec{\theta};\mathcal{H})p(\vec{\theta})d\vec{\theta} \tag{12}$$

where $\vec{\theta} = (\alpha,\delta)$. The model selection is then done with the Bayes factor

$$\mathcal{B}^{\mathcal{H}_1}_{\mathcal{H}_0} = \frac{Z_{\mathcal{H}_1}}{Z_{\mathcal{H}_0}} \tag{13}$$

3 Resolvability beteen Different Polarization Hypotheses in a three-detector Network

Notice that the degree of freedom of the likelihood function is $\mathrm{DoF} = 2(D-M)K$. With $D=3$, the maximum value of M is hence 2, and then we would only be able to test for $\mathcal{H}_{\rm tensor}$ ($M=2$), $\mathcal{H}_{\rm vector}$ ($M=2$) and $\mathcal{H}_{\rm scalar}$ ($M=1$).

For brevity, let us call the signal containing M polarization modes as a rank M signal. This can be seen in

$$\tilde{\mathbf{s}}^{D\times K} = \mathbf{F}^{D\times M}\tilde{\mathbf{h}}^{M\times K} \tag{14}$$

. The $\tilde{\mathbf{s}}$ is indeed a rank M matrix.
Qualitatively, we can argue that in general scalar model can be better resolved from tensor model than vector model. This can be understood from the nature of null projector. The $\mathbf{P}_{\text{vector}}$ is constructed from $\begin{bmatrix}\mathbf{f}_x & \mathbf{f}_\times\end{bmatrix}$ which will project away two dimensions from the data space, while the $\mathbf{P}_{\text{scalar}}$ is constructed from either $\begin{bmatrix}\mathbf{f}_b\end{bmatrix}$ or $\begin{bmatrix}\mathbf{f}_l\end{bmatrix}$ which will project away only one dimension from the data space. Suppose the underlying truth is $\mathcal{H}_{\text{tensor}}$ which implies the signal is rank 2, then there is no way the $\mathbf{P}_{\text{scalar}}$ can completely cancel the signal content. However, it is possible that the $\mathbf{P}_{\text{vector}}$ which projects away two dimensions to almost cancel the signal at some (α, δ), but this seemingly near-degeneracy is resolved by the implied time delay at each sky position. And indeed, with more detectors, $\mathcal{H}_{\text{tensor}}$ and $\mathcal{H}_{\text{vector}}$ can be better resolved.

4 Residual Energy

We can also quantitatively verify the above argument. If we assume each polarization mode carries the same amount of energy i.e.

$$\sum_n h_+^2[n] = \sum_n h_\times^2[n] \tag{15}$$

, we can show that the ratio between the residual energy and the total signal energy is independent of $\tilde{\mathbf{h}}$, and this is convenient for us to examine the resolvability between different polarization hypotheses.

4.1 Derivation

Suppose a pure tensorial signal observed in a D-detector network $\mathbf{s} = \mathbf{F}_t\mathbf{h} \in \mathbb{R}^{D\times N}$, the total signal energy can be computed by

$$E(\mathbf{s}) = \mathrm{Tr}(\mathbf{s}^T\mathbf{s}) \tag{16}$$

, and

$$\begin{aligned} E(\mathbf{s}) &= \mathrm{Tr}(\mathbf{s}^T\mathbf{s}) && (17)\\ &= \mathrm{Tr}(\mathbf{h}^T\mathbf{F}_t^T\mathbf{F}_t\mathbf{h}) && (18) \end{aligned}$$

Perform singular value decomposition on $\mathbf{F}_t$, we have

$$\mathbf{F}_t = \mathbf{U}_t\mathbf{S}_t\mathbf{V}_t^T \tag{19}$$

then substitute it into (18) and then we could express it as follows

$$
\begin{aligned}
E(\mathbf{s}) &= \mathrm{Tr}(\mathbf{h}^T\mathbf{V}_t\mathbf{S}_t\mathbf{U}_t^T\mathbf{U}_t\mathbf{S}_t\mathbf{V}^T\mathbf{h}) &(20)\\
&= \mathrm{Tr}(\mathbf{S}_t^2\mathbf{V}^T\mathbf{h}\mathbf{h}^T\mathbf{V}) &(21)\\
&= \sum_i S_{t,i}^2(\mathbf{V}^T\mathbf{h}\mathbf{h}^T\mathbf{V})_{ii} &(22)
\end{aligned}
$$

Here $\mathbf{V^T h}$ can be interpreted as the polarization modes in the rotated polarization plane. If we assume each polarization mode carries the same energy, then

$$(\mathbf{V}^T\mathbf{h}\mathbf{h}^T\mathbf{V})_{ii} = (\mathbf{h}\mathbf{h}^T)_{ii} \tag{23}$$

and

$$(\mathbf{h}\mathbf{h}^T)_{ii} = (\mathbf{h}\mathbf{h}^T)_{jj} \tag{24}$$

then

$$E(\mathbf{s}) = (\mathbf{h}\mathbf{h}^T)_{jj}\sum_i S_{t,i}^2 \tag{25}$$

where $(\mathbf{h}\mathbf{h}^T)_{jj}$ is the energy of the j-th polarization mode.
The null stream $\mathbf{z}$ is contructed by projecting out the subspaces spanned by the antenna response function of given polarization modes. Denote $\mathbf{z}_{v/s}$ be the null stream constructed with vector antenna response function and scalar antenna response function respectively.

$$\mathbf{z}_{v/s} = \mathbf{P}_{v/s}\mathbf{s} \tag{26}$$

where

$$\mathbf{P}_{v/s} = \mathbf{I} - \mathbf{F}_{v/s}(\mathbf{F}_{v/s}^T\mathbf{F}_{v/s})^{-1}\mathbf{F}_{v/s}^T \tag{27}$$

The residual energy can also be expressed in the same form

$$E(\mathbf{z}_{v/s}) = \mathrm{Tr}(\mathbf{z}_{v/s}^T\mathbf{z}_{v/s}) \tag{28}$$

Similarly, the residual energy can be expressed in terms of singular values of $\mathbf{P}_{v/s}\mathbf{F}_t$ since

$$
\begin{aligned}
E(\mathbf{z}_{v/s}) &= \mathrm{Tr}(\mathbf{z}_{v/s}^T\mathbf{z}_{v/s}) &(29)\\
&= \mathrm{Tr}(\mathbf{h}^T(\mathbf{P}_{v/s}\mathbf{F}_t)^T(\mathbf{P}_{v/s}\mathbf{F}_t)\mathbf{h}) &(30)
\end{aligned}
$$

then perform singular value decomposition on $\mathbf{P}_{v/s}\mathbf{F}_t$

$$\mathbf{P}_{v/s}\mathbf{F}_t = \mathbf{U}_{v/s}\mathbf{S}_{v/s}\mathbf{V}_{v/s}^T \tag{31}$$

and hence by the result in (25), we have

$$E(\mathbf{z}_{v/s}) = (\mathbf{h}\mathbf{h}^T)_{jj} \sum_i S^2_{v/s,i} \tag{32}$$

Finally, the ratio between the residual energy and the total energy is then

$$R_{v/s} = \frac{E(\mathbf{z}_{v/s})}{E(\mathbf{s})} \tag{33}$$

$$= \frac{(\mathbf{h}\mathbf{h}^T)_{jj} \sum_i S^2_{t,i}}{(\mathbf{h}\mathbf{h}^T)_{jj} \sum_i S^2_{v/s,i}} \tag{34}$$

$$= \frac{\sum_i S^2_{v/s,i}}{\sum_i S^2_{t,i}} \tag{35}$$

which is independent of the waveform $\mathbf{h}$. One should keep in mind that the residual energy defined here is different from that in (6) since the null projector in (6) is constructed from the noise-weighed antenna response function $\mathbf{F}_w$ instead of $\mathbf{F}$ here, and there is no way to derive a waveform independent energy ratio from $\mathbf{F}_w$ unless we assume the noise power spectral densities of the detectors are the same. If we assume the noise power spectral densities are the same, we will obtain exactly the same result in (35). But here, (35) just serves as a simplified and convenient measure of the resolvability between different polarization hypotheses.

4.2 Plots

We assume the underlying truth is a tensorial signal, and then we construct the null projector $\mathbf{P}_{\text{vector}}$ and $\mathbf{P}_{\text{scalar}}$ at the true sky position to apply on the signal, but since we have shown that the ratio of residual energy and total signal energy is

$$R_{v/s} = \frac{\sum_i S^2_{v/s,i}}{\sum_i S^2_{t,i}} \tag{36}$$

, we just need to compute the ratio between the sums of the singular values with different (α, δ) over all sky in a HLV network.
Figure 1 shows the distribution of $R_{v/s}$. It shows there is a large fraction of sky points with a huge amount of residual energy after scalar projection, but for vector projection, there is a significant fraction of sky points having a low residual energy. Figure 2 shows the distribution of the ratio between vector and scalar residual energy with respect to each sky point. The residual energy of vector projector is smaller than that of scalar projection over a vast majority of sky positions. This would indicate a general behavior that scalar model can be better resolved from the tensor model than the vector model, which agrees with our qualitative argrument. Figure 3 shows the residual energy after vector projection over all sky. Figure 4 shows the residual energy after scalar projection over all sky.

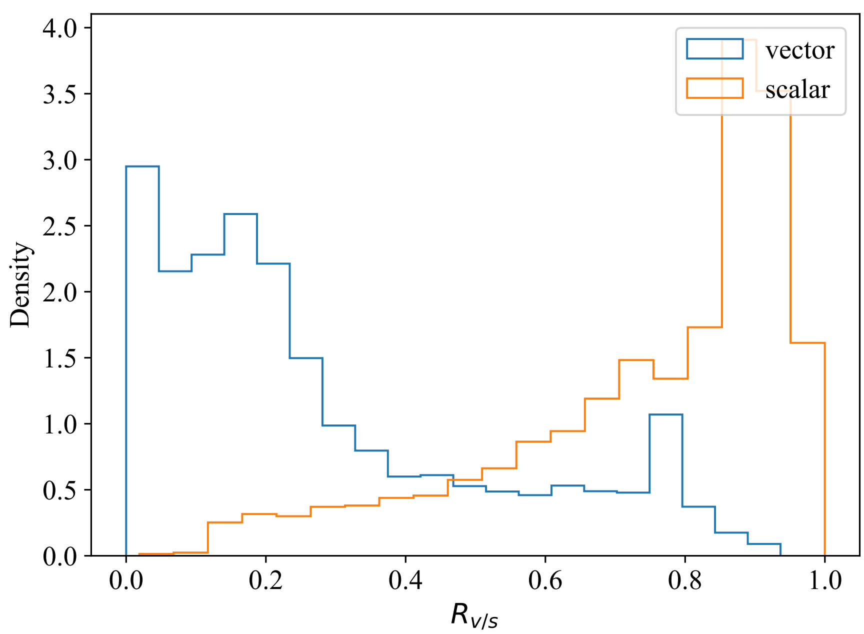

Figure 1: Distribution of residual energy.

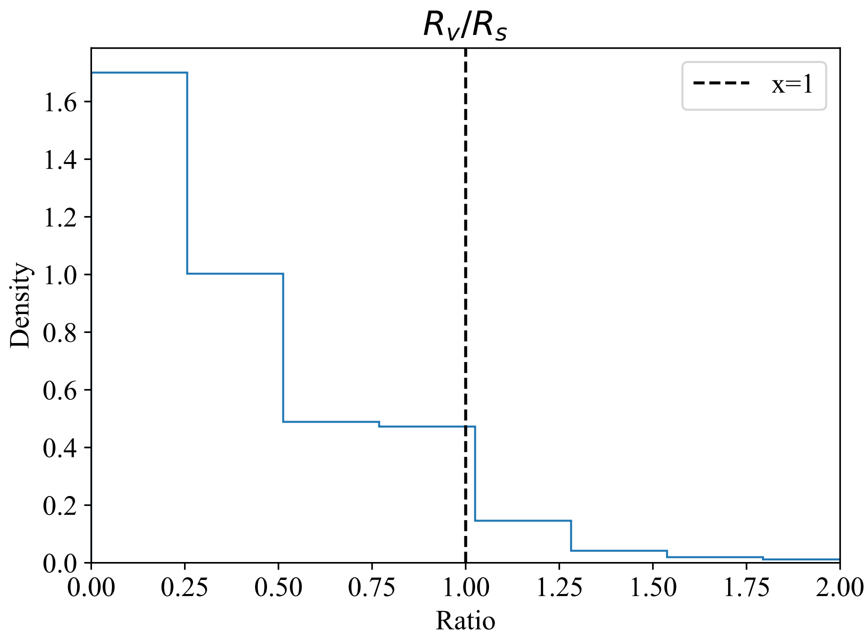

Figure 2: Distributon of of residual energy ratio.

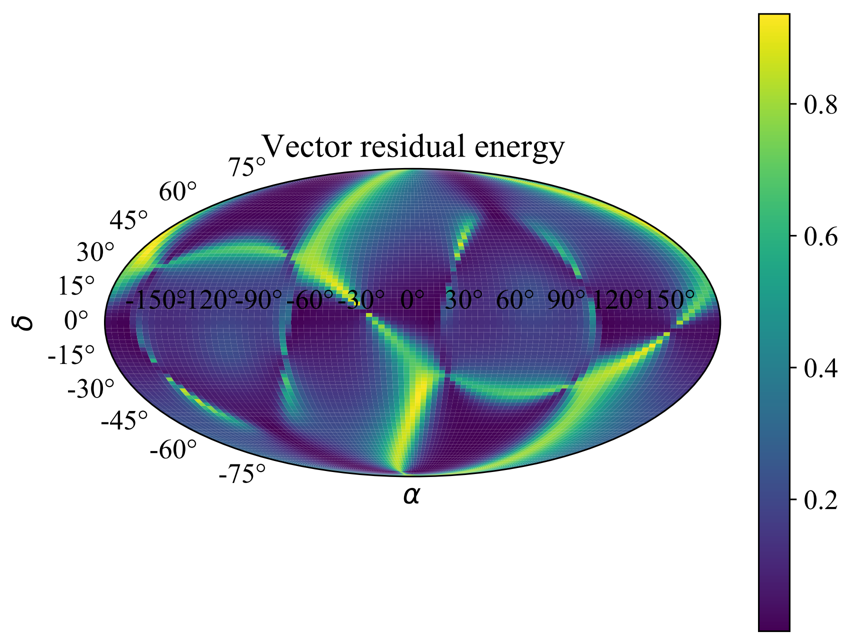

Figure 3: Skymap of vector residual energy.

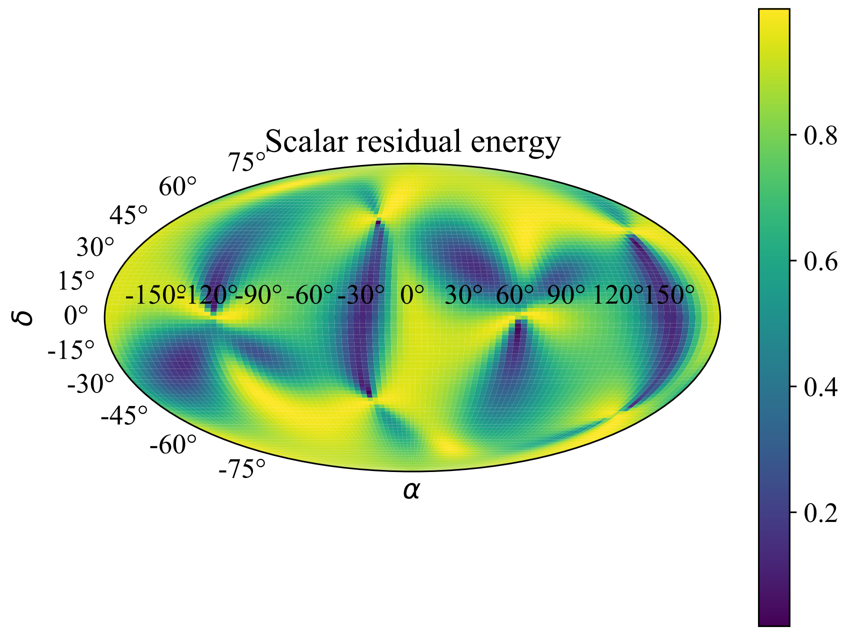

Figure 4: Skymap of scalar residual energy.

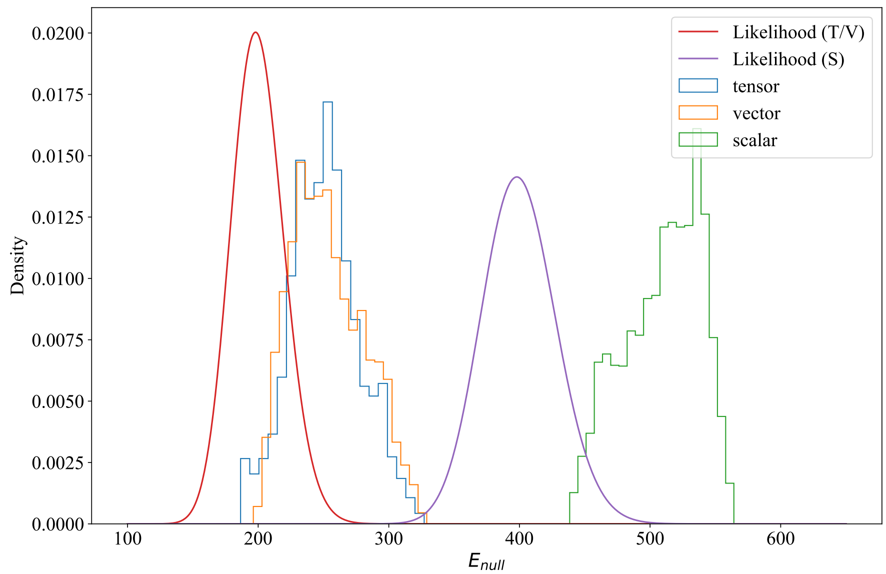

Figure 5: Distribution of E_{null} with different projectors.

5 Distribution of Null Energy over All Sky

We can also verify the greater distinguishing power of scalar model from the distribution of null enegy over all sky.

Here we inject a purely tensorial signal into HLV network with design densitivity, and then we compute the null energy over all sky by the tensor projector $\mathbf{P}_t$, vector projector $\mathbf{P}_v$ and scalar projector $\mathbf{P}_s$ respectively. The likelihood (T/V/S) (or $p(\mathbf{d}|\vec{\theta}; \mathcal{H}_{t/v/s})$) is the distribution of null energy when the assumed sky position and polarization hypothesis are correct such that there is no residual signal content. The likelihood for tensor and vector hypotheses here is the χ^2 probability density function with degree of freedom 200, while the likelihood for scalar hypothesis is the χ^2 probability density function with degree of freedom 400 (since in 3-detector network case, you will get double the degree of freedom in scalar hypothesis). Despite the difference in degree of freedom, one can observe in Figure 5 the E_{null} of scalar projection has much less overlap with its corresponding likelihood function i.e. Likelihood (S) which is attributed to the in general greater distinguishing power from the tensor model than the vector model.

If the underlying truth is purely tensorial, one should expect a more negative $\log_{10} \mathcal{B}_T^S$ than $\log_{10} \mathcal{B}_T^V$ in general due to the intrinsic nature of null projection explained in section 3.