Error Bar Choices in HERA PSPEC

pspec team

July 27, 2020

Abstract

In this memo we summarize the math behind several ways to derive error bars on power spectra which are available in the current pipeline of HERA PSPEC.

1 Foreground/systematics dependent variance

We begin with a general expression of the variance on power spectra with the existence of foregrounds or systematics, which includes both the noise variance and the signal-noise coupling term. Given two delay spectra $\tilde{x}_1 = \tilde{s} + \tilde{n}_1$ and $\tilde{x}_2 = \tilde{s} + \tilde{n}_2$, and we express $\tilde{s} = a + bi$, $\tilde{n}_1 = c_1 + d_1 i$ and $\tilde{n}_2 = c_2 + d_2 i$, the power spectra formed from $\tilde{x}_1^* \tilde{x}_2$ is

$$\begin{aligned}
P_{\tilde{x}_1\tilde{x}_2} &= \tilde{s}^*\tilde{s} + \tilde{s}^*\tilde{n}_2 + \tilde{n}_1^*\tilde{s} + \tilde{n}_1^*\tilde{n}_2 \\
&= \{a^2 + b^2 + a(c_1 + c_2) + b(d_1 + d_2) + c_1c_2 + d_1d_2\} \\
&\quad + \{a(d_2 - d_1) + b(c_1 - c_2) + d_2c_1 - d_1c_2\}i\,.
\end{aligned}$$

Here we consider $\langle s\rangle = s$, which means a and b are not random variables, but related to the true sky power spectrum by $P_{\tilde{s}\tilde{s}} = a^2 + b^2$, and c_1, d_1, c_2 and d_2 are i.i.d random normal variables. We then have

$$\begin{aligned}
\mathrm{var}\left[\mathrm{Re}(P_{\tilde{x}_1\tilde{x}_2})\right] &= \mathrm{var}\left[a^2 + b^2 + a(c_1 + c_2) + b(d_1 + d_2) + c_1c_2 + d_1d_2\right] \\
&= 2(a^2 + b^2)\langle c_1^2\rangle + 2\langle c_1^2\rangle^2 \\
&= \sqrt{2}P_{\tilde{s}\tilde{s}}P_{\mathrm{N}} + P_{\mathrm{N}}^2 \\
&= \sqrt{2}\langle \mathrm{Re}(P_{\tilde{x}_1\tilde{x}_2})\rangle P_{\mathrm{N}} + P_{\mathrm{N}}^2\,. \qquad (1)
\end{aligned}$$

In the equations above we have used the relation $\mathrm{var}(c_1c_2 + d_1d_2) = 2\langle c_1^2\rangle^2 = P_{\mathrm{N}}^2$, where P_{N} is the noise power spectrum we will refer to again later. We have also used the fact $\langle \mathrm{Re}(P_{\tilde{x}_1\tilde{x}_2})\rangle = P_{\tilde{s}\tilde{s}}$, therefore we can choose

$\sqrt{2}\mathrm{Re}(P_{\tilde{x}_1\tilde{x}_2})P_\mathrm{N} + P_\mathrm{N}^2$ as a general form of error bars with the existence of foregrounds or systematics. This will of course come with excess variance, which is not ideal, but may still be a good first-order approximation in the limit that we don't have good signal or residual systematic models.

If we consider the variance on the whole complex values of power spectrum, then

$$\begin{aligned}\mathrm{var}\left[P_{\tilde{x}_1\tilde{x}_2}\right] &= \mathrm{var}\left[a^2 + b^2 + a(c_1 + c_2) + b(d_1 + d_2) + c_1c_2 + d_1d_2\right] \\ &\quad + \mathrm{var}\left[a(d_2 - d_1) + b(c_1 - c_2) + d_2c_1 - d_1c_2\right] \\ &= 4(a^2 + b^2)\langle c_1^2\rangle + 4\langle c_1^2\rangle^2\,. \end{aligned} \tag{2}$$

It is just the form in Kolopanis et al. (2019), while they used the notation $P_\mathrm{N} = 2\langle c_1^2\rangle$, thus $\mathrm{var}\left[P_{\tilde{x}_1\tilde{x}_2}\right] = 2P_{\tilde{s}\tilde{s}}P_\mathrm{N} + P_\mathrm{N}^2$ there.

2 Analytic method from OQE formalism

The OQE formalism used in HERA PSPEC power spectrum estimation [1] naturally leads to an analytic expression of the output covariance between bandpowers.

In HERA PSPEC, an unnormalized estimator to αth bandpowers $\hat{q}_\alpha$ is defined as $\hat{q}_\alpha = \boldsymbol{x}_1^\dagger \boldsymbol{Q}^{12,\alpha}\boldsymbol{x}_2 = \sum_{ij}\boldsymbol{x}_{1,i}^*\boldsymbol{Q}_{ij}^{12,\alpha}\boldsymbol{x}_{2,j}$, where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are visibilities across frequencies. The key idea here is to propagate the input covariance on visibilities between frequencies into the output covariance on bandpowers between delays. We continue to define three sets of input covariance matrices $\boldsymbol{C}^{12}$, $\boldsymbol{U}^{12}$ and $\boldsymbol{S}^{12}$

$$\begin{aligned}\boldsymbol{C}_{ij}^{12} &\equiv \langle \boldsymbol{x}_{1,i}\boldsymbol{x}_{2,j}^*\rangle - \langle \boldsymbol{x}_{1,i}\rangle\langle \boldsymbol{x}_{2,j}^*\rangle \\ \boldsymbol{U}_{ij}^{12} &\equiv \langle \boldsymbol{x}_{1,i}\boldsymbol{x}_{2,j}\rangle - \langle \boldsymbol{x}_{1,i}\rangle\langle \boldsymbol{x}_{2,j}\rangle \\ \boldsymbol{S}_{ij}^{12} &\equiv \langle \boldsymbol{x}_{1,i}^*\boldsymbol{x}_{2,j}^*\rangle - \langle \boldsymbol{x}_{1,i}^*\rangle\langle \boldsymbol{x}_{2,j}^*\rangle\,, \end{aligned} \tag{3}$$

[1] http://reionization.org/wp-content/uploads/2020/04/HERA044_power_spectrum_normalization_v2.pdf

and we have

$$
\begin{aligned}
\langle \hat{q}_\alpha \hat{q}_\beta \rangle - \langle \hat{q}_\alpha \rangle \langle \hat{q}_\beta \rangle =& \sum_{ijkl} \langle \boldsymbol{x}_{1,i}^* \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{x}_{2,j} \boldsymbol{x}_{1,k}^* \boldsymbol{Q}_{kl}^{12,\beta} \boldsymbol{x}_{2,l} \rangle - \langle \boldsymbol{x}_{1,i}^* \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{x}_{2,j} \rangle \langle \boldsymbol{x}_{1,k}^* \boldsymbol{Q}_{kl}^{12,\beta} \boldsymbol{x}_{2,l} \rangle \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{Q}_{kl}^{12,\beta} (\langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{2,j} \boldsymbol{x}_{1,k}^* \boldsymbol{x}_{2,l} \rangle - \langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{2,j} \rangle \langle \boldsymbol{x}_{1,k}^* \boldsymbol{x}_{2,l} \rangle) \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{Q}_{kl}^{12,\beta} (\langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{1,k}^* \rangle \langle \boldsymbol{x}_{2,j} \boldsymbol{x}_{2,l} \rangle + \langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{2,l} \rangle \langle \boldsymbol{x}_{1,k}^* \boldsymbol{x}_{2,j} \rangle) \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{Q}_{kl}^{12,\beta} (\boldsymbol{S}_{ik}^{11} \boldsymbol{U}_{jl}^{22} + \boldsymbol{C}_{li}^{21} \boldsymbol{C}_{jk}^{21}) \\
=& \sum_{ijkl} (\boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{U}_{jl}^{22} \boldsymbol{Q}_{lk}^{21,\beta*} \boldsymbol{S}_{ki}^{11} + \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{C}_{jk}^{21} \boldsymbol{Q}_{kl}^{12,\beta} \boldsymbol{C}_{li}^{21}) \\
=& \mathrm{tr}(\boldsymbol{Q}^{12,\alpha} \boldsymbol{U}^{22} \boldsymbol{Q}^{21,\beta*} \boldsymbol{S}^{11}) + \mathrm{tr}(\boldsymbol{Q}^{12,\alpha} \boldsymbol{C}^{21} \boldsymbol{Q}^{12,\beta} \boldsymbol{C}^{21}) \,, \qquad (4)
\end{aligned}
$$

$$
\begin{aligned}
\langle \hat{q}_\alpha \hat{q}_\beta^* \rangle - \langle \hat{q}_\alpha \rangle \langle \hat{q}_\beta^* \rangle =& \sum_{ijkl} \langle \boldsymbol{x}_{1,i}^* \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{x}_{2,j} \boldsymbol{x}_{1,k} \boldsymbol{Q}_{kl}^{12,\beta*} \boldsymbol{x}_{2,l}^* \rangle - \langle \boldsymbol{x}_{1,i}^* \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{x}_{2,j} \rangle \langle \boldsymbol{x}_{1,k} \boldsymbol{Q}_{kl}^{12,\beta*} \boldsymbol{x}_{2,l}^* \rangle \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{Q}_{kl}^{12,\beta*} (\langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{2,j} \boldsymbol{x}_{1,k} \boldsymbol{x}_{2,l}^* \rangle - \langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{2,j} \rangle \langle \boldsymbol{x}_{1,k} \boldsymbol{x}_{2,l}^* \rangle) \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{Q}_{kl}^{12,\beta*} (\langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{2,l}^* \rangle \langle \boldsymbol{x}_{1,k} \boldsymbol{x}_{2,j} \rangle + \langle \boldsymbol{x}_{1,i}^* \boldsymbol{x}_{1,k} \rangle \langle \boldsymbol{x}_{2,j} \boldsymbol{x}_{2,l}^* \rangle) \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{Q}_{kl}^{12,\beta*} (\boldsymbol{S}_{il}^{12} \boldsymbol{U}_{kj}^{12} + \boldsymbol{C}_{ki}^{11} \boldsymbol{C}_{jl}^{22}) \\
=& \sum_{ijkl} (\boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{U}_{jk}^{21} \boldsymbol{Q}_{kl}^{12,\beta*} \boldsymbol{S}_{li}^{21} + \boldsymbol{Q}_{ij}^{12,\alpha} \boldsymbol{C}_{jl}^{22} \boldsymbol{Q}_{lk}^{21,\beta} \boldsymbol{C}_{ki}^{11}) \\
=& \mathrm{tr}(\boldsymbol{Q}^{12,\alpha} \boldsymbol{U}^{21} \boldsymbol{Q}^{12,\beta*} \boldsymbol{S}^{21}) + \mathrm{tr}(\boldsymbol{Q}^{12,\alpha} \boldsymbol{C}^{22} \boldsymbol{Q}^{21,\beta} \boldsymbol{C}^{11}) \,, \qquad (5)
\end{aligned}
$$

$$\begin{aligned}
\langle \hat{q}_\alpha^* \hat{q}_\beta^* \rangle - \langle \hat{q}_\alpha^* \rangle \langle \hat{q}_\beta^* \rangle =& \sum_{ijkl} \langle \boldsymbol{x}_{1,i} \boldsymbol{Q}_{ij}^{12,\alpha *} \boldsymbol{x}_{2,j}^* \boldsymbol{x}_{1,k} \boldsymbol{Q}_{kl}^{12,\beta *} \boldsymbol{x}_{2,l}^* \rangle - \langle \boldsymbol{x}_{1,i} \boldsymbol{Q}_{ij}^{12,\alpha *} \boldsymbol{x}_{2,j}^* \rangle \langle \boldsymbol{x}_{1,k} \boldsymbol{Q}_{kl}^{12,\beta *} \boldsymbol{x}_{2,l}^* \rangle \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha *} \boldsymbol{Q}_{kl}^{12,\beta *} (\langle \boldsymbol{x}_{1,i} \boldsymbol{x}_{2,j}^* \boldsymbol{x}_{1,k} \boldsymbol{x}_{2,l}^* \rangle - \langle \boldsymbol{x}_{1,i} \boldsymbol{x}_{2,j}^* \rangle \langle \boldsymbol{x}_{1,k} \boldsymbol{x}_{2,l}^* \rangle) \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha *} \boldsymbol{Q}_{kl}^{12,\beta *} (\langle \boldsymbol{x}_{1,i} \boldsymbol{x}_{1,k} \rangle \langle \boldsymbol{x}_{2,j}^* \boldsymbol{x}_{2,l}^* \rangle + \langle \boldsymbol{x}_{1,i} \boldsymbol{x}_{2,l}^* \rangle \langle \boldsymbol{x}_{2,j}^* \boldsymbol{x}_{1,k} \rangle) \\
=& \sum_{ijkl} \boldsymbol{Q}_{ij}^{12,\alpha *} \boldsymbol{Q}_{kl}^{12,\beta *} (\boldsymbol{S}_{jl}^{22} \boldsymbol{U}_{ik}^{11} + \boldsymbol{C}_{il}^{12} \boldsymbol{C}_{kj}^{12}) \\
=& \sum_{ijkl} (\boldsymbol{Q}_{ji}^{21,\alpha} \boldsymbol{U}_{ik}^{11} \boldsymbol{Q}_{kl}^{12,\beta *} \boldsymbol{S}_{lj}^{22} + \boldsymbol{Q}_{ji}^{21,\alpha} \boldsymbol{C}_{il}^{12} \boldsymbol{Q}_{lk}^{21,\beta} \boldsymbol{C}_{kj}^{12}) \\
=& \mathrm{tr}(\boldsymbol{Q}^{21,\alpha} \boldsymbol{U}^{11} \boldsymbol{Q}^{12,\beta *} \boldsymbol{S}^{22}) + \mathrm{tr}(\boldsymbol{Q}^{21,\alpha} \boldsymbol{C}^{12} \boldsymbol{Q}^{21,\beta} \boldsymbol{C}^{12})\,, \qquad (6)
\end{aligned}$$

where $\boldsymbol{Q}_{ij}^{12,\alpha *} = \boldsymbol{Q}_{ji}^{21,\alpha}$.

Therefore the covariance between the real part of $\hat{q}_\alpha$ and the real part of $\hat{q}_\beta$ is

$$\frac{1}{4} \left\{ (\langle \hat{q}_\alpha \hat{q}_\beta \rangle - \langle \hat{q}_\alpha \rangle \langle \hat{q}_\beta \rangle) + (\langle \hat{q}_\alpha \hat{q}_\beta^* \rangle - \langle \hat{q}_\alpha \rangle \langle \hat{q}_\beta^* \rangle) + (\langle \hat{q}_\alpha^* \hat{q}_\beta \rangle - \langle \hat{q}_\alpha^* \rangle \langle \hat{q}_\beta \rangle) + (\langle \hat{q}_\alpha^* \hat{q}_\beta^* \rangle - \langle \hat{q}_\alpha^* \rangle \langle \hat{q}_\beta^* \rangle) \right\}\,, \qquad (7)$$

and the covariance between the imaginary part of $\hat{q}_\alpha$ and the imaginary part of $\hat{q}_\beta$ is

$$\frac{1}{4} \left\{ (\langle \hat{q}_\alpha \hat{q}_\beta \rangle - \langle \hat{q}_\alpha \rangle \langle \hat{q}_\beta \rangle) - (\langle \hat{q}_\alpha \hat{q}_\beta^* \rangle - \langle \hat{q}_\alpha \rangle \langle \hat{q}_\beta^* \rangle) - (\langle \hat{q}_\alpha^* \hat{q}_\beta \rangle - \langle \hat{q}_\alpha^* \rangle \langle \hat{q}_\beta \rangle) + (\langle \hat{q}_\alpha^* \hat{q}_\beta^* \rangle - \langle \hat{q}_\alpha^* \rangle \langle \hat{q}_\beta^* \rangle) \right\}\,. \qquad (8)$$

$\hat{q}_\alpha$ should be normalized via multiplying a proper matrix $\boldsymbol{M}$ as

$$\hat{P}_\alpha = \sum_\beta \boldsymbol{M}_{\alpha\beta} \hat{q}_\beta\,. \qquad (9)$$

We then update the results above for $\hat{P}_\alpha$. The covariance between the real part of $\hat{P}_\alpha$ and the real part of $\hat{P}_\beta$ is

$$\begin{aligned}
&\frac{1}{4} \sum_{\gamma\delta} \Big\{ \boldsymbol{M}_{\alpha\gamma} \boldsymbol{M}_{\beta\delta} (\langle \hat{q}_\gamma q_\delta \rangle - \langle \hat{q}_\gamma \rangle \langle q_\delta \rangle) + \boldsymbol{M}_{\alpha\gamma} \boldsymbol{M}_{\beta\delta}^* (\langle \hat{q}_\gamma q_\delta^* \rangle - \langle \hat{q}_\gamma \rangle \langle q_\delta^* \rangle) + \\
&\boldsymbol{M}_{\alpha\gamma}^* \boldsymbol{M}_{\beta\delta} (\langle \hat{q}_\gamma^* q_\delta \rangle - \langle \hat{q}_\gamma^* \rangle \langle q_\delta \rangle) + \boldsymbol{M}_{\alpha\gamma}^* \boldsymbol{M}_{\beta\delta}^* (\langle \hat{q}_\gamma^* q_\delta^* \rangle - \langle \hat{q}_\gamma^* \rangle \langle q_\delta^* \rangle) \Big\}\,, \qquad (10)
\end{aligned}$$

and the covariance in the imaginary part of $\hat{P}_\alpha$ and the imaginary part of $\hat{P}_\beta$ is

$$\frac{1}{4}\sum_{\gamma\delta}\Big\{\boldsymbol{M}_{\alpha\gamma}\boldsymbol{M}_{\beta\delta}(\langle\hat{q}_\gamma q_\delta\rangle-\langle\hat{q}_\gamma\rangle\langle q_\delta\rangle)-\boldsymbol{M}_{\alpha\gamma}\boldsymbol{M}^*_{\beta\delta}(\langle\hat{q}_\gamma q^*_\delta\rangle-\langle\hat{q}_\gamma\rangle\langle q^*_\delta\rangle)-$$
$$\boldsymbol{M}^*_{\alpha\gamma}\boldsymbol{M}_{\beta\delta}(\langle\hat{q}^*_\gamma q_\delta\rangle-\langle\hat{q}^*_\gamma\rangle\langle q_\delta\rangle)+\boldsymbol{M}^*_{\alpha\gamma}\boldsymbol{M}^*_{\beta\delta}(\langle\hat{q}^*_\gamma q^*_\delta\rangle-\langle\hat{q}^*_\gamma\rangle\langle q^*_\delta\rangle)\Big\}\,. \tag{11}$$

Remarkably, the variance of the real part of $\hat{P}_\alpha$ is

$$\frac{1}{4}\sum_{\beta\gamma}\Big\{\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}_{\alpha\gamma}\big[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{U}^{22}\boldsymbol{Q}^{21,\gamma*}\boldsymbol{S}^{11})+\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{21}$$
$$\boldsymbol{Q}^{12,\gamma}\boldsymbol{C}^{21})\big]\;+2\times\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\big[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{U}^{21}\boldsymbol{Q}^{12,\gamma*}\boldsymbol{S}^{21})+$$
$$\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11})\big]+\boldsymbol{M}^*_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\big[\mathrm{tr}(\boldsymbol{Q}^{21,\beta}\boldsymbol{U}^{11}\boldsymbol{Q}^{12,\gamma*}$$
$$\boldsymbol{S}^{22})+\mathrm{tr}(\boldsymbol{Q}^{21,\beta}\boldsymbol{C}^{12}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{12})\big]\Big\}\,, \tag{12}$$

while the variance of the imaginary part of $\hat{P}_\alpha$ is

$$\frac{-1}{4}\sum_{\beta\gamma}\Big\{\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}_{\alpha\gamma}\big[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{U}^{22}\boldsymbol{Q}^{21,\gamma*}\boldsymbol{S}^{11})+\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{21}$$
$$\boldsymbol{Q}^{12,\gamma}\boldsymbol{C}^{21})\big]\;-2\times\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\big[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{U}^{21}\boldsymbol{Q}^{12,\gamma*}\boldsymbol{S}^{21})+$$
$$\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11})\big]+\boldsymbol{M}^*_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\big[\mathrm{tr}(\boldsymbol{Q}^{21,\beta}\boldsymbol{U}^{11}\boldsymbol{Q}^{12,\gamma*}$$
$$\boldsymbol{S}^{22})+\mathrm{tr}(\boldsymbol{Q}^{21,\beta}\boldsymbol{C}^{12}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{12})\big]\Big\}\,. \tag{13}$$

Therefore to get the final error bar on power spectrum, we should accurately model input covariance matrices on visibilities and propagate them into output covariance matrix on bandpowers. Especially, in the noise-dominated region, we have good models for the noise from the amplitudes of auto-correlation visibilities. We adopt a white noise model here, where the real and imaginary parts of noise signal are i.i.d., and uncorrelated between different frequency channels, so that we have non-zero diagonal $\boldsymbol{C}^{11}_\mathrm{n}$ and $\boldsymbol{C}^{22}_\mathrm{n}$, while $\boldsymbol{C}^{12}_\mathrm{n}$, $\boldsymbol{U}^{11}_\mathrm{n}$, $\boldsymbol{U}^{22}_\mathrm{n}$, $\boldsymbol{U}^{12}_\mathrm{n}$, $\boldsymbol{S}^{11}_\mathrm{n}$, $\boldsymbol{S}^{22}_\mathrm{n}$ and $\boldsymbol{S}^{12}_\mathrm{n}$ are all zeros! For a baseline $\boldsymbol{b}$ composed by two antennas a and b, we express $\boldsymbol{b}\equiv\{a,b\}$, and use the visibilities from auto-baseline $\{a,a\}$ and $\{b,b\}$ to estimate $\boldsymbol{C}_\mathrm{n}$ on baseline $\boldsymbol{b}$ as (Jacobs et al., 2015)

$$\begin{aligned}\boldsymbol{C}_{\mathrm{n},ii}(t)\equiv\;&\langle V_\mathrm{n}(\{a,b\},\nu_i,t)V^*_\mathrm{n}(\{a,b\},\nu_i,t)\rangle\\&-\langle V_\mathrm{n}(\{a,b\},\nu_i,t)\rangle\langle V^*_\mathrm{n}(\{a,b\},\nu_i,t)\rangle\\\approx\;&\left|\frac{V(\{a,a\},\nu_i,t)V(\{b,b\},\nu_i,t)}{N_\mathrm{nights}B\Delta t}\right|\,,\end{aligned} \tag{14}$$

where $B\Delta t$ is the product of the channel bandwidth and the integration time. Non-zero parts in Equation 12 or 13 give us the noise variance on either real or imaginary parts of power spectra as

$$\frac{1}{2}\sum_{\beta\gamma}\left\{\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\left[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{n}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{n}})\right\}. \tag{15}$$

In the following we will show it is an equivalent form of P_{N}^2 where P_{N} is what we call 'Analytic Noise Power Spectrum' estimated in another parallel way given a system temperature input.

If we also consider a 'Foreground/systematics dependent variance', Equation 12 reduces to

$$\begin{aligned}&\frac{1}{2}\sum_{\beta\gamma}\left\{\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\left[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{n}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{n}})\right.\right.\\&+\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{sky}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{n}})\\&\left.\left.+\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{n}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{sky}})\right]\right\},\end{aligned} \tag{16}$$

or

$$\begin{aligned}&\frac{1}{2}\sum_{\beta\gamma}\left\{\boldsymbol{M}_{\alpha\beta}\boldsymbol{M}^*_{\alpha\gamma}\left[\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{n}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{n}})\right.\right.\\&+\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{outer}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{n}})\\&\left.\left.+\mathrm{tr}(\boldsymbol{Q}^{12,\beta}\boldsymbol{C}^{22}_{\mathrm{n}}\boldsymbol{Q}^{21,\gamma}\boldsymbol{C}^{11}_{\mathrm{outer}})\right]\right\},\end{aligned} \tag{17}$$

$\boldsymbol{C}^{11}_{\mathrm{outer}}$ or $\boldsymbol{C}^{22}_{\mathrm{outer}}$ here is taken to be the outer product of input visibilities, like $\boldsymbol{C}^{11}_{\mathrm{outer},ij}=\boldsymbol{x}_{1,i}\boldsymbol{x}^*_{1,j}$. Equation 17 is an equivalent form to $\sqrt{2}\mathrm{Re}(P_{\tilde{x}_1\tilde{x}_2})P_{\mathrm{N}}+P_{\mathrm{N}}^2$.

2.1 Direct Noise Estimation By Differencing Visibility

For P_{N}, there are several other ways to calculate it. The sky signal (foregrounds and EoR signal) vary relatively slowly in time (or frequency), so that in a short time range we can assume the sky keeps almost constant. Thus by differencing the visibility between very close LST bins (or frequency channels), the residual is almost noise, like

$$\begin{aligned}V(\boldsymbol{b},\nu,t_1)-V(\boldsymbol{b},\nu,t_2)&\approx V_{\mathrm{n}}(\boldsymbol{b},\nu,t_1)-V_{\mathrm{n}}(\boldsymbol{b},\nu,t_2)\,,\\V(\boldsymbol{b},\nu_1,t)-V(\boldsymbol{b},\nu_2,t)&\approx V_{\mathrm{n}}(\boldsymbol{b},\nu_1,t)-V_{\mathrm{n}}(\boldsymbol{b},\nu_2,t)\,.\end{aligned} \tag{18}$$

With the visibility residual $[V_{\rm n}(\boldsymbol{b},\nu,t_1)-V_{\rm n}(\boldsymbol{b},\nu,t_2)]/\sqrt{2}$ (or $[V_{\rm n}(\boldsymbol{b},\nu_1,t)-V_{\rm n}(\boldsymbol{b},\nu_2,t)]/\sqrt{2}$), we can propagate it through the pipeline of power spectrum estimation and generate a "noise-like" power spectrum $\boldsymbol{P}_{\rm diff}$. These noise-like power spectra from differenced visibility, though highly scattered, can be seen as realizations of noise errors. For example, we take the time-differenced data to construct a noise-like power spectrum

$$\begin{aligned}\boldsymbol{P}_{\rm diff} &= \frac{(\tilde{n}_{1,t2}-\tilde{n}_{1,t_1})^*}{\sqrt{2}}\frac{(\tilde{n}_{2,t2}-\tilde{n}_{2,t_1})}{\sqrt{2}} \\ &= \{\frac{(c_{1,t2}-c_{1,t1})}{\sqrt{2}}\frac{(c_{2,t2}-c_{2,t1})}{\sqrt{2}}+\frac{(d_{1,t2}-d_{1,t1})}{\sqrt{2}}\frac{(d_{2,t2}-d_{2,t1})}{\sqrt{2}}\} \\ &\quad +\{\frac{(c_{1,t2}-c_{1,t1})}{\sqrt{2}}\frac{(d_{2,t2}-d_{2,t1})}{\sqrt{2}}-\frac{(c_{2,t2}-c_{2,t1})}{\sqrt{2}}\frac{(d_{1,t2}-d_{1,t1})}{\sqrt{2}}\}i\,, \end{aligned} \tag{19}$$

where we see $\langle\{{\rm Re}(\boldsymbol{P}_{\rm diff})\}^2\rangle\equiv\langle\{\frac{(c_{1,t2}-c_{1,t1})}{\sqrt{2}}\frac{(c_{2,t2}-c_{2,t1})}{\sqrt{2}}+\frac{(d_{1,t2}-d_{1,t1})}{\sqrt{2}}\frac{(d_{2,t2}-d_{2,t1})}{\sqrt{2}}\}^2\rangle=\langle c_1^2\rangle\langle c_2^2\rangle+\langle d_1^2\rangle\langle d_2^2\rangle=P_{\rm N}^2$. Therefore we could use $|{\rm Re}(\boldsymbol{P}_{\rm diff})|$ as a realization of error bars of $\boldsymbol{P}_{\tilde{x}_1\tilde{x}_2}$ in the noise-dominated region.

Intuitively, $\boldsymbol{P}_{\rm diff}$ can be computed from time-differenced or frequency differenced visibility. However, by differencing the neighbouring points in frequency, we actually apply a high-pass filter in the delay space which means we suppress the power at low delay modes. To illustrate it, we replace the original data vector $\boldsymbol{x}_i$ with the difference data vector $\boldsymbol{x}'_i\equiv V'(\boldsymbol{b},\nu_i)=\left[V(\boldsymbol{b},\nu_{i+1})-V(\boldsymbol{b},\nu_i)\right]/\sqrt{2}\equiv\left(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i\right)/\sqrt{2}\,(i=1,\cdots,N-1)$

and $\boldsymbol{x}'_N = \boldsymbol{x}_N$, and the new estimation of the same bandpower is

$$
\begin{aligned}
\hat{q}'_\alpha &\equiv \sum_{ij} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_i-\nu_j)} \boldsymbol{R}_{1,i}\boldsymbol{R}_{2,j}\boldsymbol{x}'^*_{1,i}\boldsymbol{x}'_{2,j} \\
&= \sum_{i=1,\cdots,N-1;j} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_i-\nu_j)} \boldsymbol{R}_{1,i}\boldsymbol{R}_{2,j} \frac{(\boldsymbol{x}_{1,i+1}-\boldsymbol{x}_{1,i})^*}{\sqrt{2}}\boldsymbol{x}'_{2,j} \\
&\quad + \sum_{j} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_N-\nu_j)} \boldsymbol{R}_{1,N}\boldsymbol{R}_{2,j}\boldsymbol{x}^*_{1,N}\boldsymbol{x}'_{2,j} \\
&= \sum_{i=1,\cdots,N-1;j} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_i-\nu_j)} \boldsymbol{R}_{1,i}\boldsymbol{R}_{2,j} \frac{\boldsymbol{x}^*_{1,i+1}}{\sqrt{2}}\boldsymbol{x}'_{2,j} \\
&\quad - \sum_{i=1,\cdots,N-1;j} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_i-\nu_j)} \boldsymbol{R}_{1,i}\boldsymbol{R}_{2,j} \frac{\boldsymbol{x}^*_{1,i}}{\sqrt{2}}\boldsymbol{x}'_{2,j} \\
&\quad + \sum_{j} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_N-\nu_j)} \boldsymbol{R}_{1,N}\boldsymbol{R}_{2,j}\boldsymbol{x}^*_{1,N}\boldsymbol{x}'_{2,j} \\
&\approx (e^{-i2\pi\eta_\alpha\Delta\nu}-1) \sum_{ij} \frac{1}{2} e^{i2\pi\eta_\alpha(\nu_i-\nu_j)} \boldsymbol{R}_{1,i}\boldsymbol{R}_{2,j} \frac{\boldsymbol{x}^*_{1,i}}{\sqrt{2}}\boldsymbol{x}'_{2,j} \\
&\approx (e^{i2\pi\eta_\alpha\Delta\nu}-1)(e^{-i2\pi\eta_\alpha\Delta\nu}-1) \\
&\quad \sum_{ij} \frac{1}{4} e^{i2\pi\eta_\alpha(\nu_i-\nu_j)} \boldsymbol{R}_{1,i}\boldsymbol{R}_{2,j}\boldsymbol{x}^*_{1,i}\boldsymbol{x}_{2,j} \\
&\approx \frac{(2\pi\eta_\alpha\Delta\nu)^2}{2}\hat{q}_\alpha \,(\text{when } \eta_\alpha\Delta\nu \ll 1)\,,
\end{aligned}
\tag{20}
$$

where we have assumed that the frequency channels are evenly spaced. If η_α is small (at low delays), we see $\hat{q}'_\alpha$ is highly suppressed from the original $\hat{q}_\alpha$, which introduce unphysical spectral structures. Thus time-differencing method is preferred to construct such noise-like power spectra.

2.2 Noise Power Spectrum

The 'analytic' noise power specturm can be also estimated from a system temperature input $T_{\rm sys}$ by (Cheng et al., 2018; Kern et al., 2020)

$$
P_{\rm N} = \frac{X^2 Y \Omega_{\rm eff} T^2_{\rm sys}}{t_{\rm int} N_{\rm coherent}\sqrt{2N_{\rm incoherent}}}\,, \tag{21}
$$

where X^2Y are conversion factors from sky angles and frequencies to cosmological coordinates, $\Omega_{\rm eff}$ is the effective beam area, $t_{\rm int}$ is the integration time, N_{coherent} is the number of samples averaged at the level of visibility while $N_{\text{incoherent}}$ is the numbers of samples averaged at the level of power spectrum.

Generally, $T_{\text{sys}} = T_{\text{sky}} + T_{\text{rcvr}}$. It can be estimated via the RMS of the differenced visibilities over samples, where we take the differences of raw visibilities in adjacent time and frequency channels to obtain the differenced visibilities first. By the relation

$$V_{\text{RMS}} = \frac{2k_b\nu^2}{c^2\Omega_p}\frac{T_{\text{sys}}}{B\Delta t}\,, \tag{22}$$

we could have a distinct system temperature on one baseline by taking RMS over all its time samples, or a baseline-time averaged system temperature over all times and baselines. Another way to estimate T_{sys} is also using the auto-correlation visibility, since itself is a good measure on the noise level on one antenna, by

$$\sqrt{V(\{a,a\})V(\{b,b\})} = \frac{2k_b\nu^2}{c^2\Omega_p}T_{\text{sys},\{\text{a,b}\}}\,. \tag{23}$$

Combining both Equation 22 and Equation 23 we derive a relation

$$V^2_{\text{RMS},\{\text{a,b}\}} = \frac{V(\{a,a\})V(\{b,b\})}{B\Delta t}\,, \tag{24}$$

which is equivalent to Equation 14 for the input noise covariance matrix. Thus the noise power spectrum estimated in this way essentially reduces to a special case of the analytic method we introduced earlier. While the analytic method is more preferred since in Equation 21 we actually use a spectral-window-averaged system temperature which might lose some information on frequency spectra during the averaging process.

References

C. Cheng, A. R. Parsons, M. Kolopanis, D. C. Jacobs, A. Liu, S. A. Kohn, J. E. Aguirre, J. C. Pober, Z. S. Ali, G. Bernardi, et al. Characterizing signal loss in the 21 cm reionization power spectrum: A revised study of paper-64. *The Astrophysical Journal*, 868(1):26, 2018.

D. C. Jacobs, J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. S. Ali, J. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, P. Klima, A. Liu, D. H. E. MacMahon, J. R. Manley, D. F. Moore, I. I.

Stefan, and W. P. Walbrugh. Multiredshift Limits on the 21 cm Power Spectrum from PAPER. , 801(1):51, Mar 2015. doi: 10.1088/0004-637X/801/1/51.

N. S. Kern, A. R. Parsons, J. S. Dillon, A. E. Lanman, A. Liu, P. Bull, A. Ewall-Wice, Z. Abdurashidova, J. E. Aguirre, P. Alexander, et al. Mitigating internal instrument coupling for 21 cm cosmology. ii. a method demonstration with the hydrogen epoch of reionization array. *The Astrophysical Journal*, 888(2):70, 2020.

M. Kolopanis, D. C. Jacobs, C. Cheng, A. R. Parsons, S. A. Kohn, J. C. Pober, J. E. Aguirre, Z. S. Ali, G. Bernardi, R. F. Bradley, et al. A simplified, lossless reanalysis of paper-64. *The Astrophysical Journal*, 883 (2):133, 2019.