Sensitivity of radio arrays to measuring 21-cm intensity power-spectrum

1 Basic definitions

The redshifted 21-cm signal is commonly represented using four equivalent quantities encoding the specific intensity (with respect to the CMB background): intensity as a function of the observed location in physical space $I(\vec{r})$, its Fourier transform $\widetilde{I}(\vec{k})$, and the scaled versions of these two functions, in sky and frequency coordinates: $\mathcal{I}(\theta_x, \theta_y, f)$, and $\widetilde{\mathcal{I}}(u, v, \eta)$.

Vector $\vec{k}$ (in comoving Mpc^{-1}) is a Fourier dual of $\vec{r}$ (in comoving Mpc), and likewise, $(\theta_x[rad], \theta_y[rad], f[Hz])$ are duals of $(u[rad^{-1}], v[rad^{-1}], \eta[sec])$. These two sets of coordinates are related through linear transformations in the following way

$$\begin{aligned} \theta_x &= \frac{r_x}{D_M(z)}, & u &= \frac{k_x D_M(z)}{2\pi}, \\ \theta_y &= \frac{r_y}{D_M(z)}, & v &= \frac{k_y D_M(z)}{2\pi}, \\ f &= \frac{H(z) f_{21,0}}{c(1+z)^2} r_z, & \eta &= \frac{c(1+z)^2}{2\pi H(z) f_{21,0}} k_z, \end{aligned} \tag{1}$$

where $f_{21,0}$ is the 21-cm frequency in the rest frame, $H(z)$ iz the Hubble parameter, $D_M(z)$ is the comoving distance in transverse direction, and z is the reference redshift in the middle of the observed data cube (where r_z and f intervals are evaluated). Note that the conditions of type $2\pi\theta_x u = r_x k_x$ are saticefied[1].

Now we can establish conventions for the Fourier transforms that relate the four intensity representation as follows

$$I(\vec{r}) = \int \widetilde{I}(\vec{k}) e^{i\vec{k}\cdot\vec{r}} d^3\vec{k}, \tag{2}$$

$$\widetilde{I}(\vec{k}) = \frac{1}{(2\pi)^3} \int I(\vec{k}) e^{-i\vec{k}\cdot\vec{r}} d^3\vec{r}, \tag{3}$$

$$\mathcal{I}(\theta_x, \theta_y, f) = \int\int\int \widetilde{\mathcal{I}}(u, v, \eta) e^{2\pi i(u\theta_x + v\theta_y + \eta f)} du dv d\eta, \tag{4}$$

$$\widetilde{\mathcal{I}}(u, v, \eta) = \int\int\int \mathcal{I}(\theta_x, \theta_y, f) e^{-2\pi i(u\theta_x + v\theta_y + \eta f)} d\theta_x d\theta_y df, \tag{5}$$

[1] The factor of 2π just comes from the Fourier-transform conventions in Eqs. 4 and 5.

such that the following scaling relation is saticefied

$$\widetilde{I}(\vec{k}) = \frac{H(z) f_{21,0}}{c(1+z)^2 D_M(z)^2} \widetilde{\mathcal{I}}(u, v, \eta), \tag{6}$$

where the proportionallity factor between these two functions is a Jacobian $\frac{d\theta_x d\theta_y df}{dr_x dr_y dr_z}$ (this scaling was obtained by substituting Eq. 1 into Eq. 5, and comparing the result to Eq. 3). Note also that there is no scaling between I and $\mathcal{I}$—they are the same function, up to change of variables.

Finally, since radio interferometers measure visibilities, we define the visibility for a pair of antennae as a Fourier-transform in frequency-coordinate only of the uv-plane specific intensity,

$$\widetilde{\mathcal{I}}(u, v, \eta) = \int V(u, v, f) e^{-2\pi f \eta} df, \tag{7}$$

where f is a discrete variable such that $f_{\max} - f_{\min} = \Delta f$ is the bandwidth of the observed data cube centered on z (see also Appendix A).

2 Power spectra definitions

We want to derive an expression for the power spectrum of noise intensity (including instrument noise, and noise from the sky) in $\vec{k}$ space. This power spectrum is given by

$$\langle \widetilde{I}(\vec{k}) \widetilde{I}^*(\vec{k}') \rangle = (2\pi)^3 P_{\widetilde{I}} \delta(\vec{k} - \vec{k}'). \tag{8}$$

On the other hand, the measurable quantity—the visibility—is a complex Gaussian variable with a zero mean, whose noise-induced component has a variance of[2]

$$\langle V(\vec{u}, f) V(\vec{u}', f')^* \rangle = \left(\frac{2 k_B T_{\rm sys}}{A_e \sqrt{\Delta f t_1}} \right)^2 \delta(\vec{u} - \vec{u}') \delta_{ff'}, \tag{9}$$

where t_1 is the total time a single baseline spent observing an element at the position $\vec{u} \equiv (u, v)$ in the uv plane.

3 From visibility to $P^N(\vec{k})$

The next step is to combine Eqs. 7, and 9, and take ensamble average to get

$$\begin{aligned} \langle \widetilde{\mathcal{I}}(u, v, \eta) \widetilde{\mathcal{I}}^*(u', v', \eta') \rangle &= \int\int \langle V(u, v, f) V^*(u', v', f') \rangle e^{2\pi i (f'\eta' - f\eta)} df df' \\ &= \frac{1}{t_1} \left(\frac{2 k_B T_{\rm sys}}{A_e} \right)^2 \delta(\vec{u} - \vec{u}') \delta(\eta - \eta'), \end{aligned} \tag{10}$$

where

$$\int e^{2\pi i f (\eta - \eta')} df = \delta(\eta - \eta'), \tag{11}$$

[2] See Appendix for a derivation of $V_{\rm rms}$.

is the periodic delta-function on the t_1 interval, and $D_{\rm band}$ is the total bandwidth of this data cube centered on z.

As the final step, we need to take into account the scaling relation of Eq. 6 and substitute it on the LHS of the above equation, and then introduce the power spectrum of Eq. 8. In addition, keeping in mind the property of the delta-function that $\delta(ax) = \frac{1}{a}\delta(x)$, we can substitute the relations between variables in Eq. 1 to cancel the delta functions in Eq. 10. We thus arrive at the following expression for the noise power spectrum, per $\vec{k}$ mode, per baseline,

$$P_1^N(\vec{k}) = \frac{1}{t_1}\frac{c(1+z)^2 D_M^2(z)}{H(z) f_{21,0}}\left(\frac{2k_B T_{\rm sys}}{A_e}\right)^2. \tag{12}$$

Note at the end how to compute t_1 from the total duration of the survey $t_{\rm obs}$. For a small beam size (much smaller than 2π), where telescopes scan the sky one beam width at a time, like for the case of radio dishes, t_1 is the total time spent observing a given uv element of size corresponding to the beam $\Omega_{\rm beam} = \lambda^2/A_e$, and is obtained by multiplying $t_{\rm obs}$ by the ratio of the solid angle of the survey $\Omega_{\rm survey}$ and the solid angle of the beam. However, in the case of simple dipole antennas, the beam is of the size of the survey, and coveres half the sky, and in this case t_1 just equals $t_{\rm obs}$.

The last step is to get from Eq. 12 to the expression for the noise power spectrum that corresponds to the observation with all the available baselines. To do that, we need to incorporate the knowledge about the array configuration and the coverage of the uv plane. In other words, we need to divide the expression in Eq. 12 by the number of baselines that see a given mode $\vec{k}$ at any given time $n_{\rm base}(\vec{k})$ (for a discussion of the uv coverage, see the following section). The final result for the noise power spectrum per mode $\vec{k}$ in the intensity units is then

$$P^N(\vec{k}) = \frac{\Omega_{\rm survey}}{t_{\rm obs}\lambda^2}\frac{c(1+z)^2 D_M^2(z)}{H(z) f_{21,0}}\frac{(2k_B T_{\rm sys})^2}{A_e n_{\rm base}(\vec{k})}, \tag{13}$$

and in temperature units

$$P^N(\vec{k}) = \frac{\lambda^2\Omega_{\rm survey}}{t_{\rm obs}}\frac{c(1+z)^2 D_M^2(z)}{H(z) f_{21,0}}\frac{T_{\rm sys}^2}{A_e n_{\rm base}(\vec{k})}. \tag{14}$$

Ta-dah!

4 The UV Coverage

Total number of baselines that can observe mode $\vec{k}$, $n_{\rm base}(\vec{k})$, is related to the (unitless) number density of baselines per element $dudv$, $n(u,v)$, as

$$n_{\rm base}(\vec{k}) = \frac{n(u,v)}{\Omega_{\rm beam}}, \tag{15}$$

where $\frac{1}{\Omega_{\rm beam}}$ represents an element in the uv plane, and the number density integrates to the total number of baselines $N_{\rm base}$ as

$$N_{\rm base} = \frac{1}{2}N_{\rm ant}(N_{\rm ant}+1) = \int_{\rm half} n(u,v)\,dudv, \tag{16}$$

where $N_{\rm ant}$ is the number of antennas in the array, and the integration is done on the half of the uv plane[3].

Let us now derive $n(\vec{k})$ for a specific array configuration that is of particualr interest to cosmology[4] —a tightly packed array of simple dipole antennas, tiling a squared-surface of the area $(\Delta L)^2$ with a filling fraction close to one (see Figure ??).[5] In this case, the beam is close to the size of the entire survey and coveres 1 sr, the affective area of a single dipole is $A_e = \lambda^2$, and the effective number of antennas is then $N_{\rm ant} = \frac{(\Delta L)^2}{\lambda^2}$.

A uniform tiling of a $(\Delta L)^2$ surface with dipoles then results in the following number density of baselines per uv element

$$n(u,v) = (\frac{\Delta L}{\lambda} - u)(\frac{\Delta L}{\lambda} - v). \tag{17}$$

After substituting the relation between $\vec{k} = (k, \theta_k, \phi_k)$ and (u, v),

$$\begin{aligned} u_\perp &\equiv \frac{D_A(z)}{2\pi} k \sin\theta_k, \\ u &= u_\perp \cos\phi_k, \\ v &= u_\perp \sin\phi_k, \end{aligned} \tag{18}$$

the corresponsing number of baselines observing a given $\vec{k}$ is then

$$n_{\rm base}(\vec{k}) = (\frac{\Delta L}{\lambda} - \frac{D_A(z)}{2\pi} k \sin\theta_k \cos\phi_k)(\frac{\Delta L}{\lambda} - \frac{D_A(z)}{2\pi} k \sin\theta_k \sin\phi_k). \tag{19}$$

5 MVQ estimator for uniform B

We will now derive an unbiased minimum-variance quadratic estimator $\widehat{B}$ for a uniform magnetic field (MF) $\vec{B}$, following a similar CMB formalism. We start by noting that the redshifted 21-cm brightness temperature contains contribution from the noise, and the signal, where the signal is generated both by the 21-cm signal with no magnetic field (null-case signal), and by $\vec{B}$, such that

$$\begin{aligned} T(\vec{k}) &= T_N(\vec{k}) + T_S(\vec{k}), \\ T_S(\vec{k}) &= T_{S,0}(\vec{k}) + B \frac{\partial T_S(\vec{k})}{\partial B}\bigg|_{B=0}, \end{aligned} \tag{20}$$

where the magnitude of the field B is a small expansion parameter. The derivative in the above equation is evaluated at $B = 0$, and similarly for the signal $T_{S,0}$ in the null case. We can now compute the following expectation value

$$\langle T(\vec{k}) T^*(\vec{k}')\rangle = P_{\rm null}(\vec{k})(2\pi)^3 \delta(\vec{k} - \vec{k}') + \langle T_{S,0}(\vec{k}) B \frac{\partial T_S^*(\vec{k}')}{\partial B}\bigg|_{B=0}\rangle + \langle T^*_{S,0}(\vec{k}') B \frac{\partial T_S(\vec{k})}{\partial B}\bigg|_{B=0}\rangle \tag{21}$$

[3] This is because the visibility has the following property $V(u, v, f) = V^*(-u, -v, f)$, and only half the plane contains independent samples.

[4] Note also that we assume that the array consists of many antennas, so that time-dependence of $n(u, v)$ is negligible; if this is not the case, one should compute its time average to account for Earth's rotation.

[5] This assumes a design such as the FFT telescope described in Tegmark and Zaldarriaga (2009).

where we introduced notation for the power spectrum in the null case,

$$P_{\rm null}(\vec{k}) \equiv P^N(\vec{k}) + P_0^S(\vec{k}), \tag{22}$$

such that P_0^S represents the 21-cm power spectrum in the absence of MFs. We have also assumed that the signal and the noise are uncorrelated, and kept only terms linear in B. To expand the RHS of Eq. 21 further, we first note that the only Gaussian random field the signal temperature is proportional to is the density fluctuation δ, with the proportionality being the transfer funtion G,

$$\begin{aligned} T_S(\vec{k}) &= G(\hat{k})\delta(k), \\ T_{S,0}(\vec{k}) &= G(\hat{k}, B=0)\delta(k), \end{aligned} \tag{23}$$

where $\hat{k}$ is a unit vector in the direction of $\vec{k}$. In this case, we note that

$$\frac{\partial T_S(\vec{k})}{\partial B} = \delta(k)\frac{\partial G(\vec{k})}{\partial B}, \tag{24}$$

where the derivative will be evaluated at $B = 0$. Eq. 21 now becomes

$$\langle T(\vec{k})T^*(\vec{k}')\rangle = \left(P_{\rm null}(\vec{k}) + 2BP_\delta(\vec{k})Re\left[G^*(\vec{k}, B=0)\frac{\partial G(\vec{k})}{\partial B}\bigg|_{B=0}\right]\right)(2\pi)^3\delta(\vec{k}-\vec{k}'). \tag{25}$$

Note here that the delta function evaluates to

$$\delta(\vec{k}-\vec{k}') = \frac{V}{(2\pi)^3}, \quad \text{for } \vec{k}=\vec{k}', \tag{26}$$

where V is the space volume of the survey, in our case.

The next step is to note that we observe only one universe, so the measured proxy for the ensamble average of Eq. 25 will be just the product $T(\vec{k})T^*(\vec{k})$. Using this fact and Eq. 26, and inverting Eq. 25 gives an estimate for B from a single $\vec{k}$-mode measurement,

$$\widehat{B}_{\vec{k}} = \frac{\frac{1}{V}T(\vec{k})T^*(\vec{k}) - P_{\rm null}(\vec{k})}{2P_\delta(\vec{k})Re\left[G^*(\vec{k}, B=0)\frac{\partial G(\vec{k})}{\partial B}\Big|_{B=0}\right]}. \tag{27}$$

This estimator is unbiased, $\langle \widehat{B}_{\vec{k}}\rangle = 0$, and the above expression can be used to calculate its covariance, from

$$C_{\vec{k},\vec{k}'} \equiv \langle \widehat{B}_{\vec{k}}\widehat{B}^*_{\vec{k}'}\rangle = \frac{\left\langle\left(\frac{1}{V}T(\vec{k})T^*(\vec{k}) - P_{\rm null}(\vec{k})\right)\left(\frac{1}{V}T^*(\vec{k}')T(\vec{k}') - P_{\rm null}(\vec{k}')\right)\right\rangle}{4P_\delta(\vec{k})P_\delta(\vec{k}')Re\left[G^*(\vec{k})\frac{\partial G(\vec{k})}{\partial B}\right]Re\left[G(\vec{k}')\frac{\partial G^*(\vec{k}')}{\partial B}\right]}, \tag{28}$$

where all derivatives and G are evaluated at $B = 0$, following, as usual, the null assumption. For simplicity, from now on, we drop the explicit notation for $B = 0$ but retain the null assumption in the entire derivation. The expectation value in the above equation involves temperature four-point correlation. If we enumerate factors of "T" in this correlation as $\langle 1\ 2\ 3\ 4\rangle$, the expansion of this correlation of four Gaussian random variables can be represented as a sum of the following contractions: $\langle T(\vec{k})T^*(\vec{k})T^*(\vec{k}')T(\vec{k}')\rangle = \langle 1\ 2\rangle\langle 3\ 4\rangle + \langle 1\ 4\rangle\langle 2\ 3\rangle + \langle 1\ 3\rangle\langle 2\ 4\rangle$. Keeping this order of summands, the correlation becomes

$$\langle T(\vec{k})T^*(\vec{k})T^*(\vec{k}')T(\vec{k}')\rangle = V^2 P_{\rm null}(\vec{k})^2 \left(1 + \delta_{\vec{k},\vec{k}'} + \delta_{\vec{k},-\vec{k}'}\right) \tag{29}$$

where we used Eq. 26, and the relation between delta function and Kronecker delta,

$$\delta_{\vec{k},\vec{k}'} = \frac{(2\pi)^3}{V}\delta(\vec{k} - \vec{k}'). \tag{30}$$

The rest of the terms in Eq. 28 are of the form

$$\frac{1}{V}\langle T(\vec{k})T^*(\vec{k})\rangle P_{\rm null}(\vec{k}') = P_{\rm null}(\vec{k})P_{\rm null}(\vec{k}'). \tag{31}$$

Finally, substituting Eqs. 29, 31, and 30, into Eq. 28, we get the following expression for the covariance

$$\langle \widehat{B}_{\vec{k}}\widehat{B}^*_{\vec{k}'}\rangle = \frac{P^2_{\rm null}(\vec{k})\left(\delta_{\vec{k},\vec{k}'} + \delta_{\vec{k},-\vec{k}'}\right)}{\left(2P_\delta(\vec{k})Re\left[G^*(\vec{k})\frac{\partial G(\vec{k})}{\partial B}\right]\right)^2}, \tag{32}$$

where the variance $\sigma^2_{\vec{k}} \equiv \langle \widehat{B}_{\vec{k}}\widehat{B}^*_{\vec{k}}\rangle$ represents diagonal elements of the covariance matrix.

This covariance matrix is singular, and the only non-vanishing entries are those relating the same mode with itself (or to the same mode in the opposite direction), which is a consequence of the reality of the temperature field, and the isotropy of space in the null-assumption case. The usual expression for a minimum-variance estimator,

$$\widehat{B} = \frac{\sum_{\vec{k},\vec{k}'} C^{-1}_{\vec{k},\vec{k}'}\widehat{B}_{\vec{k}}}{\sum_{\vec{k},\vec{k}'} C^{-1}_{\vec{k},\vec{k}'}}, \tag{33}$$

in this case reduces to

$$\widehat{B} = \frac{1}{2}\frac{\sum_{\vec{k}} \frac{\widehat{B}_{\vec{k}}}{\sigma^2_{\vec{k}}}}{\sum_{\vec{k}} \frac{1}{\sigma^2_{\vec{k}}}}, \tag{34}$$

where the factor of $\frac{1}{2}$ comes from the two Kronecker deltas in Eq. 32. The final expression for the estimator is then

$$\widehat{B} = \frac{\sum_{\vec{k}} \frac{\frac{1}{V}T(\vec{k})T^*(\vec{k}) - P_{\rm null}(\vec{k})}{P^2_{\rm null}(\vec{k})} P_\delta(\vec{k})Re\left[G^*(\vec{k})\frac{\partial G(\vec{k})}{\partial B}\right]}{\sum_{\vec{k}}\left(\frac{2P_\delta(\vec{k})Re\left[G^*(\vec{k})\frac{\partial G(\vec{k})}{\partial B}\right]}{P_{\rm null}(\vec{k})}\right)^2}, \tag{35}$$

Note at the end that the expression in the denominator is exactly the expression for the integrand for the inverse variance of the Fisher forecast, as expected.

6 Fisher Analysis

We will first discuss the unsatured case, where the strength of $\vec{B}$ produces less than 1 rad of precession at all z of interest, and then move on to discussing detectability in the saturated case, where B is strong in this sense.

If an experiment measures redshifted 21-cm brightness-temperature power spectrum $P(k, \theta_k, \phi_k)$, which is a function of a parameter B (in this case, the strength of a uniform MF that evolves as $B = B_0(1+z)^2$), then this experiment's sensitivity to recovering B is given by the Fisher matrix, which combines measurements at every (k, θ_k, ϕ_k) mode,

$$\sigma_B^{-2}(z) = \int dV_{\text{patch}}(z) \frac{k^2 dk d\phi_k \sin\theta_k d\theta_k}{(2\pi)^3} \left(\frac{\frac{\partial P_S}{\partial B}}{P_N + P_S} \right)^2$$
$$= \int dV_{\text{patch}}(z) \frac{k^2 dk d\phi_k \sin\theta_k d\theta_k}{2(2\pi)^3} \left(\frac{2P_\delta(k,z) G(B=0, \theta_k, \phi_k) \frac{\partial G}{\partial B}(\theta_k, \phi_k, z)\Big|_{B=0}}{P_N(k, \theta_k) + P_\delta(k,z) G^2(B=0, \theta_k, \phi_k)} \right)^2, \tag{36}$$

where V_{patch} is the comoving volume of the survey, such that

$$dV_{\text{patch}} = \frac{c}{H(z)} D_A^2 \Omega_{\text{survey}}. \tag{37}$$

Note that we operate under the null assumtion of small B, so every summand in the above equation is evaluated for a fiducial case of $B = 0$. In this case, the above expression can be used to compute amgnitude of B detectable at 1-σ level for a given noise level (and cosmology). The power spectrum of the 21-cm signal is in the denominator of the summands in order to account for sample variance. In our case, it is computed for a given cosmology, as

$$P_S = G^2(\theta_k, \phi_k) P_\delta(k), \tag{38}$$

where P_δ is the power spectrum of density fluctuations at a given redshift, for a given cosmology, and G is the transfer function (a function of B computed from microphysical calculations of Paper I).[6] The integration limits in Eq. 36 are: $\phi_k \in [0, 2\pi]$; $\theta_k \in [0, \pi]$; and $k \in [2\pi u_{\min}/(d_A \sin\theta_k), 2\pi u_{\max}/(d_A \sin\theta_k)]$, where $u_{\min,\max} = \frac{L_{\text{min, max}}}{\lambda}$ correspond to the maximum and minimum baseline, respectively.

For the saturated case, we are unable to measure the exact magnitude of B, but a slightly different inquiry becomes interesting. Namely, it is useful to know how sensitive are future experiments to distinguishing staturated case from zero MF. To make a forecast for this question, let us write the signal power spectrum as a sum of the contributions from very weak field (a.k.a. $B = 0$), and very strong field (denoted as infinity),

$$P = (1-\xi) P\Big|_{B\to 0} + \xi P\Big|_{B\to\infty}, \tag{39}$$

[6] Note here that we do not take the average of the noise over ϕ_k, because we are interested in reconstructing the sensitivity of a given (B, θ_B, ϕ_B).

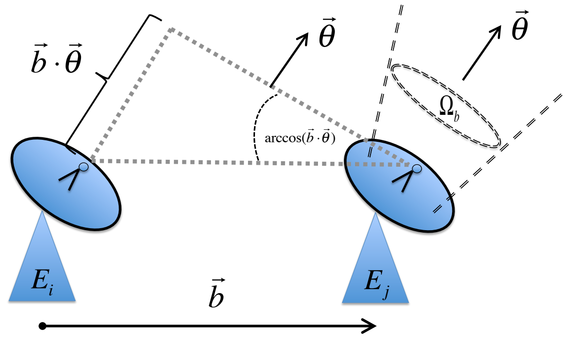

Fig. 1: Two-antennae interferometer.

where we have not explicitly written the indicies and dependencies, to maximize clarity of the equation. We can then perform the standard Fisher analysis, completely analogous to the unsaturated case, in order to understand constraints on parameter ξ. In this case, σ_ξ would give a 1-σ sensitivity to detecting the maximum change in G, due to the rise to the saturation ceiling (see Figure **??**).

A Visibility-variance derivation

Here we derive the variance of the visibility for an interferometric array of two antennas separated by a baseline $\vec{b}$, each with an effective collecting area A_e, observing a single element in uv plane for time duration t_1, in the total bandwidth $\Delta f = f_{\max} - f_{\min}$ of N_f discrete frequencies. This setup is shown in Figure 1.

The quantity such interferometer measures is the correlation between electric fields E_i and E_j induced in the two antennae, as a function of the frequency. Since the observation time t_1 is finite, frequencies measured during t_1 are discrete,

$$f_n = n/t_1, \tag{40}$$

where $n \in Z$, such that

$$E(t) = \sum_n \widetilde{E}(f_n) e^{2\pi i f_n t}. \tag{41}$$

The correlation coefficient of the electric fields is defined as

$$\rho_{ij}(f) \equiv \frac{\langle \widetilde{E}_i^*(f)\widetilde{E}_j(f)\rangle}{\sqrt{\langle|\widetilde{E}_i(f)|^2\rangle\langle|\widetilde{E}_j(f)|^2\rangle}}, \tag{42}$$

for each measured frequency f. We now assume that $var(Re[\widetilde{E}_i]) = var(Im[\widetilde{E}_i]) = \sigma^2$, such that the real (or imaginary) part of ρ has the following variance

$$\begin{aligned} var(Re[\rho_{ij}(f)]) &= \frac{1}{(2\sigma^2)^2} var(\langle Re[\widetilde{E}_i] Re[\widetilde{E}_j] + Im[\widetilde{E}_i] Im[\widetilde{E}_j]]\rangle) \\ &= \frac{2\sigma^2\sigma^2}{(2\sigma^2)^2} = \frac{1}{2N_f} = \frac{1}{2t_1\Delta f}, \end{aligned} \tag{43}$$

where the number of observed frequencies in time t_1 is derived from Eq. 40 to be,

$$N_f = t_1 \Delta f. \tag{44}$$

Before continuing, let us take a brief digression to show that the above formula implicitly assumes that the electric fields in the two antennas $\widetilde{E}_i$ and $\widetilde{E}_j$ have a very weak correlation, $\rho \ll 1$. Namely, suppose x and y are random Gaussian variables, where $var(x) \equiv \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$, and similarly for y, and their correlation coefficient is $\rho \equiv \frac{\langle xy \rangle}{\sqrt{\langle x^2 \rangle \langle y^2 \rangle}}$. In this case, the following is true

$$\begin{aligned} var(xy) &= \langle x^2 y^2 \rangle - \langle xy \rangle^2 = \langle x^2 \rangle \langle y^2 \rangle + \langle xy \rangle^2 \\ &= \langle x^2 \rangle \langle y^2 \rangle + \rho \langle x \rangle^2 \langle y \rangle^2 = var(x) var(y)(1 + \rho^2), \end{aligned} \tag{45}$$

so that when ρ is small, then $var(xy) = var(x)var(y)$, which was assumed in the first equality of Eq. 43.

Resuming the derivation, if different frequencies are uncorrelated, the result of Eq. 43 implies

$$\langle |\rho_{ij}(f)|^2 \rangle = \frac{1}{t_1 \Delta f}. \tag{46}$$

The final step in this derivation requires the relation between intensity in the sky $\mathcal{I}(\vec{\theta}, f)$ (within the beam of size Ω_b in steradians) and the measured electric fields,

$$\langle \widetilde{E}_i^*(f) \widetilde{E}_j(f) \rangle = C \int_{\Omega_b} d^2\vec{\theta} \mathcal{I}(\vec{\theta}, f) e^{2\pi i \frac{f}{c} \vec{b}_{ij} \cdot \vec{\theta}} R(\vec{\theta}), \tag{47}$$

where $\vec{\theta} = (\theta_x, \theta_y)$ is a unit vector that defines a direction in the sky, C is a constant, and R is the antenna response function (the shape of the beam in the sky); $\frac{2\pi f}{c} \vec{b}_{ij} \cdot \vec{\theta}$ is the phase delay between two antennae. From the above formula and Eq. 42, it follows[7]

$$\rho_{ij}(f) = \frac{C \int_{\Omega_b} d^2\vec{\theta} \mathcal{I}(\vec{\theta}, f) e^{2\pi i (u\theta_x + v\theta_y)}}{C \int_{\Omega_b} d^2\vec{\theta} \mathcal{I}(\vec{\theta}, f)}, \tag{48}$$

where the denominator in the above formula approximately integrates to (for a small beam)

$$\int_{\Omega_b} d^2\vec{\theta} \mathcal{I}(\vec{\theta}, f) \approx \Omega_b \mathcal{I}(\vec{\theta}, f). \tag{49}$$

[7] Note that a position in the uv plane measures the phase lag between the two dishes in wavelenghts, such that $u\theta_x + v\theta_y \equiv \frac{\vec{b} \cdot \vec{\theta}}{\lambda}$.

We can now use the approximate expression for the resolution of a single dish,

$$\Omega_b = \frac{\lambda^2}{A_e}, \tag{50}$$

the Reyliegh-Jeans law (or the definition of the brightness temperature),

$$\mathcal{I}(\vec{\theta}, f) = \frac{2k_B T_{\rm sys}}{\lambda^2}, \tag{51}$$

and note that the numerator in Eq. 48 matches the definition of visibility (Eq. 7), to get

$$\rho_{ij}(f) = \frac{A_e}{2k_B T{\rm sys}} V(u, v, f), \tag{52}$$

Substituting Eq. 46 into the above expression, we get the final result of this derivation,

$$\langle |V(u, v, f)|^2 \rangle = \left(\frac{2k_B T_{\rm sys}}{A_e \sqrt{t_1 \Delta f}} \right)^2 \delta(\vec{u} - \vec{u}') \delta_{ff'}, \tag{53}$$

where V is a complex Gaussian variable, centered at zero, and uncorrelated for different values of its arguments.

It should be noted at the end that we were calculating the contribution to the visibility from the noise only (the system in the absence of a signal), so we used system temperature for brightness temperature (this could contain the signal from foregrounds and from the instrument). In case we want to repeat the computation in the presence of a signal, $T_{\rm sys}$ should instead be the sum of the signal and the noise temperatures.

What about polarization?