HERA Memo 128
Received 27 Feb 2024

Delay window functions for the HERA beams

Adélie Gorce

February 9, 2024

The point of this memo is to present the derivation of exact window function for the analysis of HERA's power spectrum estimates and to compare results between the dipole and the Vivaldi feeds. The H1C IDR3 full season results were published using this derivation of the window functions, for the dipole feeds (Fagnoni et al. 2021a). We expect the H4C results to be published with the same method, but this time using the beam simulations corresponding to the Vivaldi feeds (Fagnoni et al. 2021b).

The exact window functions include accurate information about the instrument and the data, outside of the framework of the delay approximation. The derivation has been detailed in Gorce et al. (2023) and included to the `hera_pspec` code through the `uvwindow` module. References will be made throughout to functions included in this class.

We present results for the two frequency bands considered for the analysis of H1C data, that is:

- Band 1, $117.1 \leq \nu/\mathrm{MHz} \leq 132.6$ (spectral window indices: 175-334), centred on $z = 10.4$, and
- Band 2, $150.3 \leq \nu/\mathrm{MHz} \leq 167.8$ (spectral window indices: 515-694), centred on $z = 7.9$.

1 HERA beam

This calculation requires a model for the beam. We consider the HERA dipole beam simulations by Nicolas Fagnoni (`https://github.com/HERA-Team/HERA-Beams`) for different polarisation channels: pseudo-Stokes (`pI,pQ,pU,pV`) and power (`xx,xy,yx,yy`). The E-field allows us to obtain the beams for the four pseudo-Stokes parameters (pI,pQ,pU,pV). Using the linear relations below, we derive the beams for the measurement polarisation:

$$\begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{pmatrix} \begin{pmatrix} V_{XX} \\ V_{XY} \\ V_{YX} \\ V_{YY}. \end{pmatrix} \tag{1}$$

or

$$\begin{pmatrix} V_{XX} \\ V_{XY} \\ V_{YX} \\ V_{YY} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} V_I \\ V_Q \\ V_U \\ V_V \end{pmatrix} \tag{2}$$

The beam is given in azimuthal-zenith (ϕ, z) coordinates, which we convert into Cartesian coordinates (x, y) by assuming the sky is flat on the considered area:

$$\begin{cases} x = z\cos\phi \\ y = z\sin\phi. \end{cases} \tag{3}$$

We obtain the beam illustrated in Fig. 1 for the `pI` polarisation.

There are differences in the two beams. First, in terms of spatial structure (upper panels): The side lobes have a different spatial structure, with the side lobes of the Vivaldi feeds being larger in amplitude.

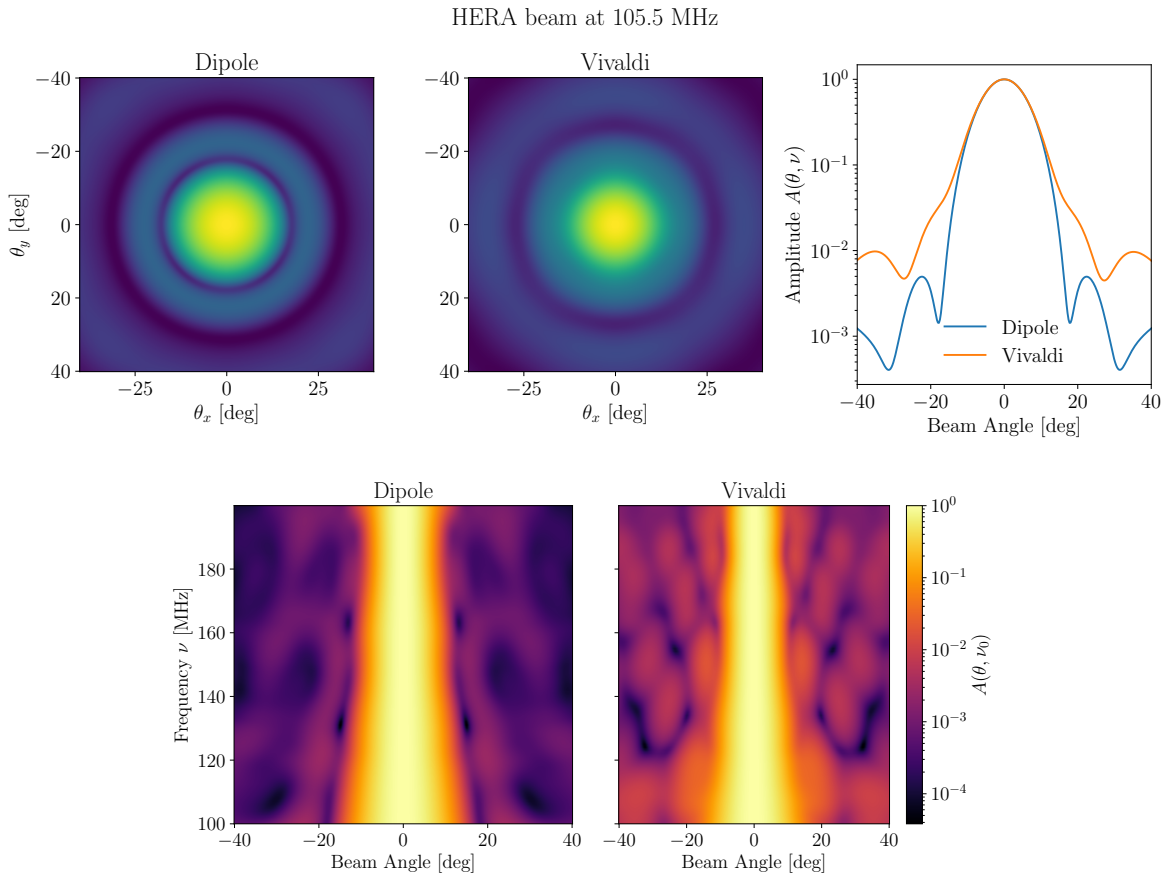

Figure 1: Comparison of the HERA dipole and Vivaldi beams amplitudes $A(\theta, \nu)$ for the `pI` polarisation, obtained with a simulation and normalised to one (Fagnoni et al. 2021a,b). *Top: Spatial structure.* HERA beams projected on a flat sky surface at 105.5 MHz . This figure is similar to Fig. 19b of Fagnoni et al. (2021b), although at a different frequency. *Bottom: Chromaticity.* Frequency evolution of the HERA beams along the H1C bandwidth.

Because of this, we expect differences in the $k_\perp$ structure of the window functions computed with both beams. Second, in terms of frequency dependence: in the lower panels of the figures, we see the evolution of the beam's spatial structure with frequency. Because of this, we expect differences in the $k_\parallel$ structure of the window functions. Note that the Vivaldi feeds give access to a larger bandwidth than the dipole feeds but, here, we only consider the dipole bandwidth (100-200 MHz).

2 Derivation of the delay window functions

We want to derive window functions which allow for a better mapping between measurement space ($\mathbf{b}$, τ) and cosmological space ($k_\perp$, $k_\parallel$) by considering the impact of the delay approximation.

2.1 Spatial Fourier transform

Take the visibility equation:

$$V(\mathbf{b},\nu) = \int \mathrm{d}^2\theta\; T(\theta,\nu) A(\theta,\nu) \mathrm{e}^{-2i\pi\nu\mathbf{b}\cdot\theta/c} \tag{4}$$

with $A(\theta,\nu)$ the HERA beam, illustrated on Fig. 1, and $T(\theta,\nu)$ the sky temperature which we can re-write in Cartesian coordinates in the flat-sky approximation as

$$T(\mathbf{r}_\perp, r_\parallel) \equiv T(\theta,\nu), \tag{5}$$

with $\mathbf{r}_\perp \equiv d_c(z)\theta$ and $r_\parallel \equiv \alpha(z)\nu$ where

$$\alpha(z) \equiv \frac{c(1+z)^2}{\nu_{21} H(z)} \tag{6}$$

for ν_{21} the rest-frame 21cm frequency, $H(z)$ the Hubble function. Letting its Fourier transform $T(k_\perp, k_\parallel)$, one can write

$$T(\mathbf{r}_\perp, r_\parallel) = \int \frac{\mathrm{d}^2\mathbf{k}_\perp \mathrm{d}k_\parallel}{(2\pi)^3}\, T(\mathbf{k}_\perp, k_\parallel)\; \mathrm{e}^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp + k_\parallel r_\parallel)}, \tag{7}$$

which, in turn, leads to

$$V(\mathbf{b},\nu) = \frac{1}{(2\pi)^3} \int \mathrm{d}^2\theta \int \mathrm{d}^2\mathbf{k}_\perp\; \mathrm{d}k_\parallel\, A(\theta,\nu)\, T(\mathbf{k}_\perp, k_\parallel)\; \mathrm{e}^{i\theta\cdot[d_c(z)\mathbf{k}_\perp - 2\pi\mathbf{b}\nu/c]} \mathrm{e}^{i\alpha(z)\nu k_\parallel}. \tag{8}$$

We define the delay transform as the Fourier transform of the visibility measured by one baseline, along the frequency direction:

$$V(\mathbf{b},\tau) \equiv \int \mathrm{d}\nu\, V(\mathbf{b},\nu)\; \mathrm{e}^{-2i\pi\nu\tau}. \tag{9}$$

Following the equations above, we have

$$\begin{aligned} V(\mathbf{b},\tau) &= \frac{1}{(2\pi)^3} \int \mathrm{d}^2\mathbf{k}_\perp\; \mathrm{d}k_\parallel\, T(\mathbf{k}_\perp, k_\parallel) \int \mathrm{d}\nu \int \mathrm{d}^2\theta\; A(\theta,\nu)\; \mathrm{e}^{i\theta\cdot[d_c(z)\mathbf{k}_\perp - 2\pi\mathbf{b}\nu/c]}\; \mathrm{e}^{i\nu[\alpha(z)k_\parallel - 2\pi\tau]} \\ &= \frac{1}{(2\pi)^3} \int \mathrm{d}^2\mathbf{k}_\perp\; \mathrm{d}k_\parallel\, T(\mathbf{k}_\perp, k_\parallel)\; \chi(\mathbf{k}_\perp, k_\parallel; \mathbf{b}, \tau), \end{aligned} \tag{10}$$

where we have defined the function χ which describes the mapping between Fourier space and measurement space. The delay spectrum can then be written as

$$\begin{aligned} \langle |V(\mathbf{b},\tau)|^2 \rangle &= \frac{1}{(2\pi)^6} \int \mathrm{d}^3\mathbf{k} \int \mathrm{d}^3\mathbf{k}'\, \langle T(\mathbf{k}) T(\mathbf{k}') \rangle\, \chi(\mathbf{k};\mathbf{b},\tau) \chi^\star(\mathbf{k}';\mathbf{b},\tau) \\ \hat{P}(k_\perp, k_\parallel) &= \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{k}\, P(\mathbf{k})\, |\chi(\mathbf{k};\mathbf{b},\tau)|^2. \end{aligned} \tag{11}$$

We see that the power spectrum estimator at a point $(k_\perp, k_\parallel)$ is a weighted sum of the true power spectrum at this point. The weighted sum is the window function:

$$W(\mathbf{k}; \mathbf{b}, \tau) = |\chi(\mathbf{k}; \mathbf{b}, \tau)|^2. \tag{12}$$

This function will be later normalised such that all $(\mathbf{b}, \tau)$ bins are given equal weight, that is, for each $(\mathbf{b}, \tau)$ bin, $\int \mathrm{d}k_\parallel \, \mathrm{d}k_\perp \, W(k_\perp, k_\parallel) = 1$.

From a mathematical point of view, for each pair $(\mathbf{b}, \tau)$, we have a 2D matrix W of the window function and a 2D matrix P of the true power spectrum, then the estimator for this bin is $\hat{P} = \mathrm{Tr}(W^\mathrm{T} P)$.

Above, we have defined

$$\begin{aligned} \chi(\mathbf{k}; \mathbf{b}, \tau) &= \int \mathrm{d}\nu \, \mathrm{e}^{2i\pi\nu[\alpha(z)k_\parallel/2\pi - \tau]} \int \mathrm{d}^2\theta \; A(\theta, \nu) \; \mathrm{e}^{2i\pi\theta\cdot[d_c(z)\mathbf{k}_\perp/2\pi - \mathbf{b}\nu/c]} \\ &= \int \mathrm{d}\nu \, \mathrm{e}^{2i\pi\nu[\alpha(z)k_\parallel/2\pi - \tau]} \tilde{A}(\mathbf{q}_\perp, \nu) \end{aligned} \tag{13}$$

The latter term is the Fourier transform of $A(\theta)$ for the Fourier duals

$$\mathbf{q}_\perp \equiv \frac{\nu}{c}\mathbf{b} - \frac{d_c(z)}{2\pi}\mathbf{k}_\perp. \tag{14}$$

In the case of the delay approximation, the frequency-dependent term vanishes and the Fourier dual of θ is simply $\frac{d_c(z)}{2\pi}\mathbf{k}_\perp$. We recover the fact that the delay approximation is valid for short baselines ($b \sim 0$). For each frequency, the FT of the beam shifts by $\nu b/c$, as illustrated on Fig 2[1]. Note that this effect strongly depends on the taper that is applied along the frequency axis: Most of the contribution to the resulting window function will come from the frequency at which the taper is maximal. For a Blackman-Harris taper, the window function will be determined by the FT of the beam at the central frequency along the spectral window considered.

In Fig. 2, we compare the spatial Fourier transforms of the dipole and Vivaldi beams. As expected, they follow the same shift with frequency but their spatial structure is different: for the Vivaldi beam, the Fourier transform is narrower (as the beam had stronger side lobes in real space), such that we expect the window functions to be narrower and the leakage from neighbouring modes into a given spherical bin to be slightly weaker.

Code integration: The $\tilde{A}(\mathbf{q}_\perp, \nu)$ object corresponds to the `FTBeam` class of the `UVWindow` module. To define such an object, there are three options:

1. Compute from a beam simulation file with the `FTBeam.from_beam` method. This task is computationally intensive.
2. Load pre-computed the Fourier transforms along the whole HERA bandwidth with the `FTBeam.from_file` method. We have stored these files in the `h5py` format on lustre. Different files correspond to different polarisation channels, indicated as a suffix to the file (e.g., `"FT_beam_HERA_dipole_xx.hdf5"`).
3. Compute for a Gaussian beam with the `FTBeam.gaussian` method. Then all derivations are analytical and computing times are greatly reduced.

This Fourier transform is performed numerically, in pixel space, and therefore the arrays stored do not correspond to any physical units. The conversion from numerical to physical units is given by the `kperp4bl_freq` method and depends on the spectral window and baseline length b considered.

[1] For a Gaussian field, the FT is another Gaussian, and the centre of the Gaussian shifts with frequency. The beam is a Gaussian in real space $A(\theta) = \exp[-\theta^2/\sigma^2]$. Integrating by parts, we find the corresponding Fourier transform to be (in two dimensions), $\tilde{A}(q) = \pi\sigma^2/L \, \exp[-(\pi\sigma)^2 q^2]$.

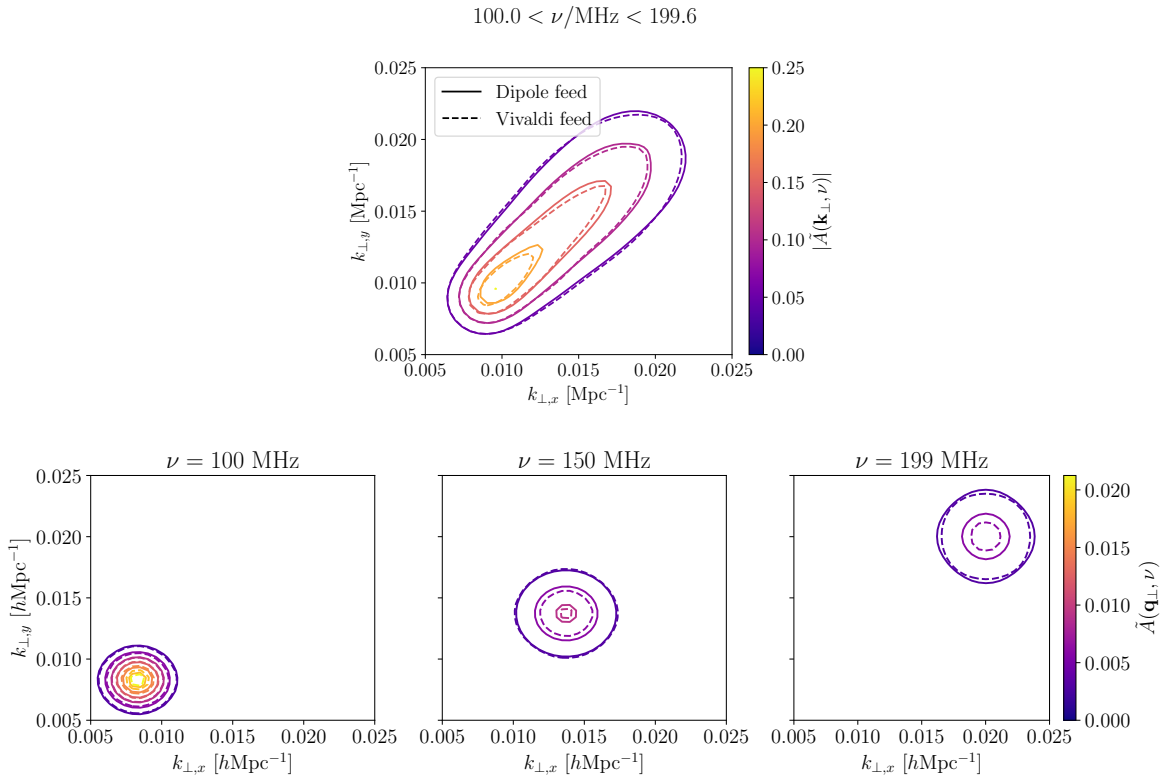

Figure 2: Spatial Fourier transform of the HERA beam, represented for different frequencies along the HERA bandwidth, to illustrate the chromaticity of the instrument. The Fourier transform of the beam shifts along the $k_\perp$ space as frequency increases, illustrating the impact of beam chromaticity on the cosmological scales probed by the instrument. In all panels, we compare the results obtained for the dipole feed (solid lines) and the Vivaldi feed (dashed lines).

2.2 Fourier transform along the frequency axis

The final integral over the frequency in equation (13) can be considered a Fourier transform, where η is the Fourier dual of ν such that

$$\eta \equiv \tau - \frac{\alpha(z)}{2\pi} k_\parallel. \tag{15}$$

Now, we have

$$\chi(\mathbf{k}_\perp, k_\parallel; \mathbf{b}, \tau) = \tilde{A}\left(\frac{\nu}{c}\mathbf{b} - \frac{d_c(z)}{2\pi}\mathbf{k}_\perp, \tau - \frac{\alpha(z)}{2\pi}k_\parallel\right). \tag{16}$$

The object $\tilde{A}$ (and so χ) is maximal at $\eta = 0$ or $k_\parallel = 2\pi\tau/\alpha(z)$, which is the centre of the corresponding window function.

> *Code integration:* The Fourier transform along the frequency direction is performed inside the `UVWindow` class, with the `_take_freq_FT` method, which takes as inputs the $\tilde{A}(\mathbf{q}_\perp, \nu)$ array (`FTBeam` object), interpolated on a regular $(k_{\perp,x}, k_{\perp,y})$ grid (via `interpolate_FT_beam`) and the channel width $\delta\nu$. We do not recommend using this function in a standalone way but rather as part of the method `uvwindow.get_cylindrical_wf` which, given a `UVWindow` object, computes the cylindrical window function along a ($k_\perp$, $k_\parallel$) grid for a given baseline length (and all the delays along the spectral window considered).

2.3 Cylindrical window functions

After performing cylindrical average, one gets

$$W(k_\perp, k_\parallel; b, \tau) = \int k_\perp \mathrm{d}\phi_k \, |\tilde{A}(\mathbf{q}_\perp, \eta)|^2, \tag{17}$$

where ϕ_k is the azimuthal angle in k-space. Note that if the polarisation channels cross-correlated to obtain the bandpower are different, for example `xx` and `yy`, the equation above writes $W(k_\perp, k_\parallel; b, \tau) = \int k_\perp \mathrm{d}\phi_k \, \tilde{A}_{xx}(\mathbf{q}_\perp, \eta)\tilde{A}^\star_{yy}(\mathbf{q}_\perp, \eta)$.

> *Code integration:* The option to cross-correlate different polarisation channel is not currently implemented in the `UVWindow` class.

These window functions are represented in Fig. 3 for the simulated HERA beams and for four different baseline lengths, at different delay times. Most of the structure seen along $k_\parallel$ is due to the taper (here, a Blackman-Harris function has been used) and disappears in favour of a `sinc` when using a step function. The bins are centred at

$$\begin{cases} k_\perp = |\mathbf{k}_\perp| = \dfrac{2\pi}{d_c(z)} \dfrac{\nu b}{c}, \\ k_\parallel = \dfrac{2\pi|\tau|}{\alpha(z)}. \end{cases} \tag{18}$$

Note that these window functions are symmetrical with respect to τ, that is

$$W(k_\perp, k_\parallel; \mathbf{b}, \tau) = W(k_\perp, k_\parallel; \mathbf{b}, -\tau). \tag{19}$$

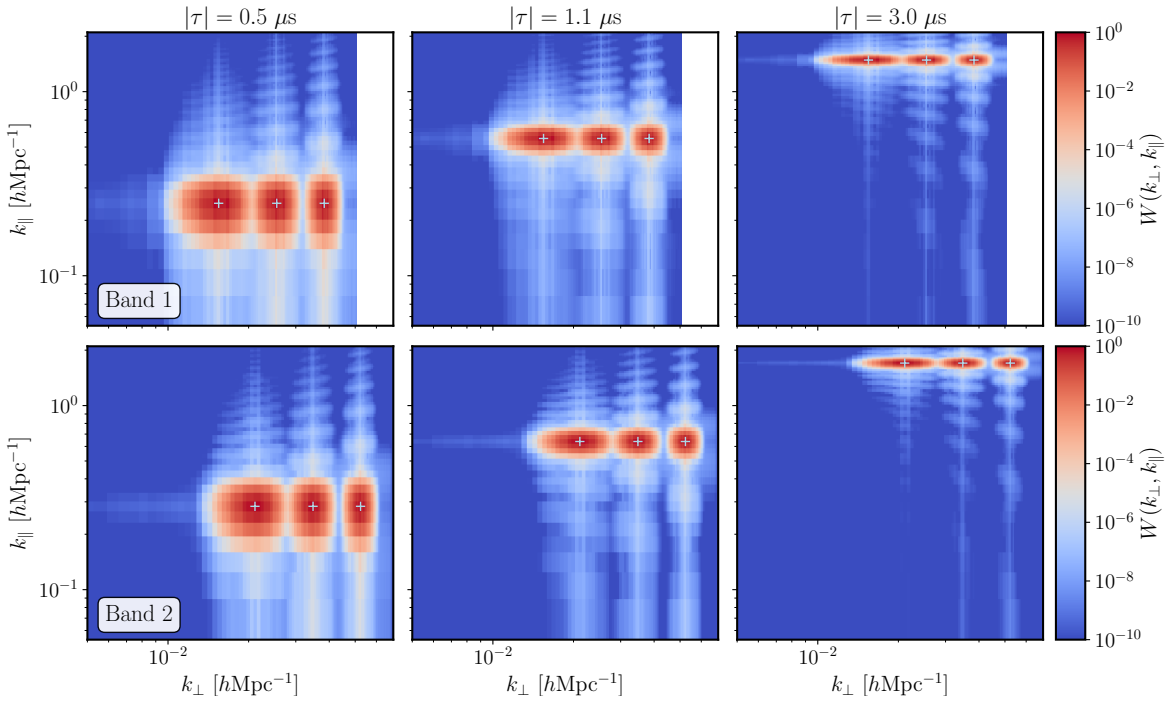

Figure 3: Cylindrical window functions obtained according to equation (12) for different baseline lengths (38.7, 63.7, 95.8 m) and for the simulated HERA dipole beam along Band 1. Each panel corresponds to a different delay. Figure from Gorce et al. (2023).

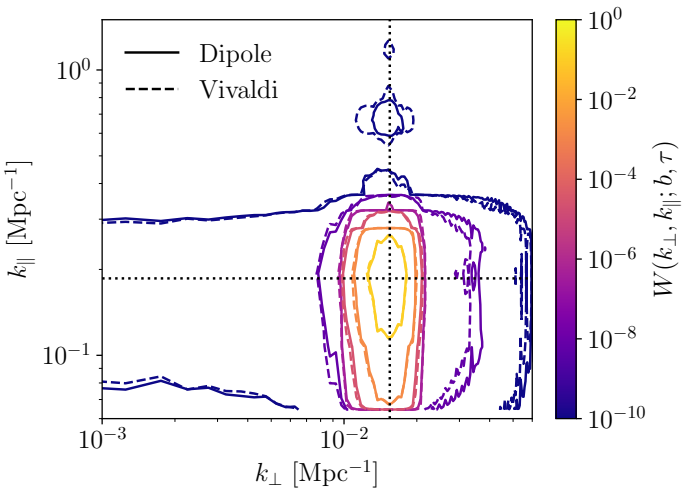

Figure 4: Comparison of the cylindrical window function obtained for the dipole and the Vivaldi beam simulations for $b = 38.6$m and $\tau = 375$ns along Band 1.

Code integration: The cylindrical average is performed in function `UVWindow.get_cylindrical_wf`, which takes as inputs the baseline length b, the polarisation, the FT of the beam $\tilde{A}(\mathbf{q}_\perp, \eta)$ and optional bins along $k_\perp$ and $k_\parallel$.

In Fig. 4, we compare the cylindrical window function obtained for both HERA beams along Band 1. We see that they generally present the same structure along the line-of-sight and perpendicular modes. As observed in Fig. 2, the Vivaldi feed leads to slightly narrower window functions along the $k_\perp$ axis. However, this is only true for the first mode. Despite the amplitude of the window function being modulated by the Fourier transform of the Blackman-Harris taper used, we see that the higher order modes are larger for the Vivaldi than for the dipole feed, which (see Fig. 1) had a different chromaticity. If these differences are easily noticed on Fig. 4 because of the choice of a logarithmic colour scale, we do not expect them to propagate to large difference in the spherical window functions or in the theoretical interpretation of the power spectrum, as they are only at the $\sim 10^{-10}$ to 10^{-8} level.

2.4 Spherical window functions

We perform a spherical average to obtain the window functions as a function k and compare them to the fiducial ones. For each spherical k-bin, we identify the (b, τ) pairs that contribute (in general, there is more than one), and take a weighted average of them. We then take a spherical average of the remaining function in $(k_\perp, k_\parallel)$ space. That is, for each k-bin α:

$$W_\alpha(k) = \int \mathrm{d}\Omega_k \frac{1}{\sum_i w_i} \sum_{(b,\tau)_i} w_i W_i(k_\perp, k_\parallel) \tag{20}$$

such that the centre $(k_\perp, k_\parallel)_i$ of the window function associated with $(b, \tau)_i$ is within the α k-bin. The weights w_i correspond to the observational significance of the $(b, \tau)_i$ pair, and depend for example on their noise. These weights will depend on choices made in the earlier analysis of the data - for example, limits on the maximum baselines considered (see H1C 1DR3 memo).

The exact spherical window function for bin $k = 0.61\,h\mathrm{Mpc}^{-1}$ obtained with the HERA beam along Band 1 is shown on Fig. 5 and compared to the one obtained in the delay approximation. We see that the

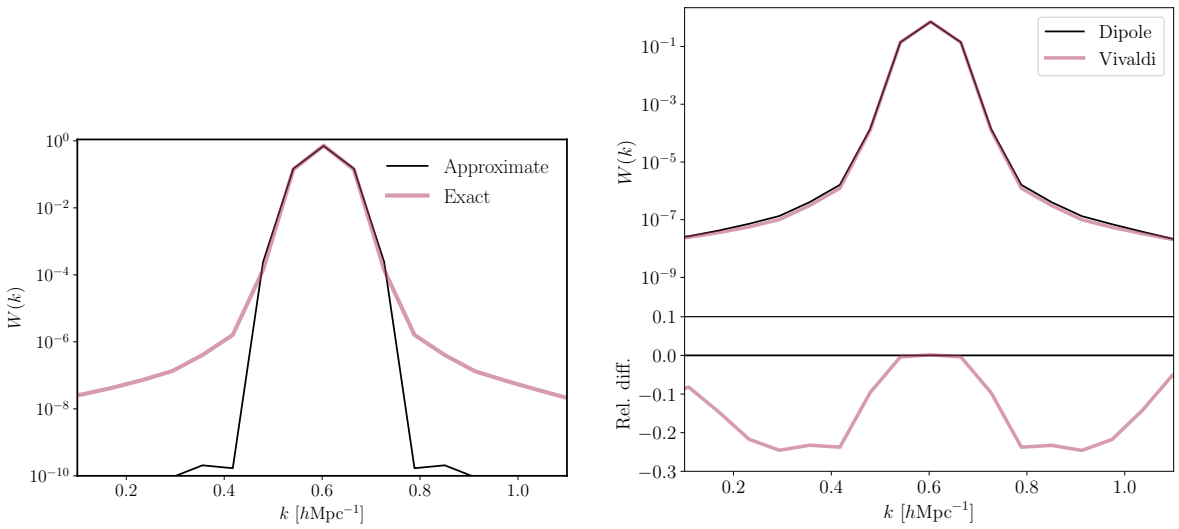

Figure 5: Window function for spherical bin $k = 0.61\,h\text{Mpc}^{-1}$ *Left:* Comparison the window function computed according to equation (12) for the HERA dipole beam along Band 1 with its equivalent, obtained in the delay approximation. See also Figure 7 of Gorce et al. (2023). *Right:* Comparison of the exact window function obtained for the dipole and the Vivaldi feeds.

approximation leads to strongly underestimating the tails of the window functions, and therefore ignoring the overlap between bins. However, the difference only becomes significant when the window function reaches a 10^{-4} amplitude, which means it will only become noticeable in a power spectrum estimator for high amplitude signals

The right-hand side of the figure presents the same window function, computed with the simulated beams for the dipole and the Vivaldi feeds. The Vivaldi window function has smaller tails than its dipole counterpart on almost all scales but the centre of the bin with up to a 20% difference at $\Delta k = 0.2\,h\text{Mpc}^{-1}$ from the centre. Despite this difference being small in absolute value (about 10^{-4}), it could partly limit leakage of very bright foregrounds into the cosmological power spectrum near the wedge limit.

> *Code integration:* The spherical average is performed in function `UVWindow.get_spherical_wf`, which takes as inputs the spherical k-bins to be used, a list of baseline lengths to compute the window functions for, and potentially associated weights. Note that the spherical average itself is performed, from the cylindrical window functions, with `UVWindow.cylindrical_to_spherical`.

2.5 Window functions for the Vivaldi feeds

In the previous section, we compare the window functions obtained for the dipole and Vivaldi feeds on the H1C bandwith and spectral windows ($100 \leq \nu/\text{MHz} \leq 200$, channel width 90 kHz). However, installing the Vivaldi feeds has given us access to a wider range of frequencies. Data files for H4C cover frequencies 46.9 to 234.3 MHz for a channel width of 122 kHz. In this section, we will look at the window functions in the different bands considered for the analysis of H4C data, as the choice of bandwidth and the channel width have an impact on the shape of the window functions (see Sec. 4.2.2 of Gorce et al. (2023)). We considered the following bands:

- Band 1: $51 \leq \nu/\text{MHz} \leq 62$
- Band 2: $63 \leq \nu/\text{MHz} \leq 75$
- Band 3: $75 \leq \nu/\text{MHz} \leq 87$
- Band 4: $110 \leq \nu/\text{MHz} \leq 124$
- Band 5: $126 \leq \nu/\text{MHz} \leq 136$

- Band 6: $139 \leq \nu/\mathrm{MHz} \leq 145$
- Band 7: $146 \leq \nu/\mathrm{MHz} \leq 153$
- Band 8: $156 \leq \nu/\mathrm{MHz} \leq 165$
- Band 9: $165 \leq \nu/\mathrm{MHz} \leq 175$
- Band 10: $176 \leq \nu/\mathrm{MHz} \leq 182$
- Band 11: $198 \leq \nu/\mathrm{MHz} \leq 209$
- Band 12: $209.5 \leq \nu/\mathrm{MHz} \leq 220$.

Note that not all the bands considered will result in an upper limit.

Because the beam was simulated only on a range $50 \leq \nu/\mathrm{MHz} \leq 250$, i.e. the 46.9 to 50 MHz range present in the data is missing in the simulation, we extrapolate the beam below 50 MHz to its value at 50 MHz such that **we assume no frequency dependence of the beam below 50 MHz.** This is fine as long as the bands considered for upper limits do not go below 50 MHz.

We show in figure 6 the cylindrical window functions for different sets of (τ, b) values, for all the 12 bands considered in the H4C analysis. The differences we see between the different bands are mostly due to the spectral differences: The window functions are centred at different $(k_\parallel, k_\parallel)$ for the same (b, τ) between bands because equations (18) are redshift-dependent, whilst some window functions are wider than others along the longitudinal direction because the band is narrower (e.g., band 6 compared to band 5). As their shape is dominated by the Fourier transform along the frequency direction, this effect propagates to the spherical window functions, shown in figure 7. Their width is directly proportional to the width of the band. Note that because the shape depends mostly on the spectral characteristics, we don't expect the window functions to change much when applying data weights.

3 Impact of data characteristics

3.1 Tapering the spectral window

In practice, the visibilities are not measured on an infinite range of frequencies, and when defining the delay transform in Eq. 9, the integral over frequency is limited to the spectral window considered. We write

$$V(\mathbf{b}, \tau) = \int \mathrm{d}\nu \, V(\mathbf{b}, \nu) \, \mathrm{e}^{-2i\pi\nu\tau} \times \Phi_0(\nu), \tag{21}$$

if ν_0 is the central frequency of the spectral window and B_0 is its width, that is the bandwidth. The tapering function $\Phi_0(\nu)$ can be chosen by the data analyst and is zero outside of the spectral window, maximal at its middle. By default, we consider a Blackman-Harris filter (see upper panel of Fig. 8). An additional term is added to the definition of χ in equation (13):

$$\begin{aligned} \chi(\mathbf{k}; \mathbf{b}, \tau) &= \int \mathrm{d}\nu \, \mathrm{e}^{2i\pi\nu[\alpha(z)k_\parallel/2\pi - \tau]} \tilde{A}(\mathbf{q}_\perp, \nu) \times \Phi_0(\nu) \\ &= \tilde{A}\left(\frac{\nu}{c}\mathbf{b} - \frac{d_c(z)}{2\pi}\mathbf{k}_\perp, \tau - \frac{\alpha(z)}{2\pi}k_\parallel\right) * \tilde{\Phi}_0(\eta), \end{aligned} \tag{22}$$

where $*$ denotes a convolution and $\tilde{\Phi}_0$ the Fourier transform of the taper. Hence, the window functions will be modulated by the Fourier transform of the taper along longitudinal modes.

We show different figures to understand the impact of the taper on our window functions in Fig. 8. First, we see that applying the taper reduces the frequency-dependence of the Fourier transform of the beam along the spectral window as it concentrates the contribution of the central frequency (middle panel). This is equivalent to reducing by half the width of the spectral window and results in wider window function tails along the $k_\parallel$ direction. However, the stronger effect comes from the modulation of the window function amplitude by the Fourier transform of the taper. If no taper (or a step-like taper) is applied, the window function is modulated by a η^{-2} power-law corresponding to the Fourier transform of the step function – the sinc function. This modulation is clearly visible in the lower panel, both for

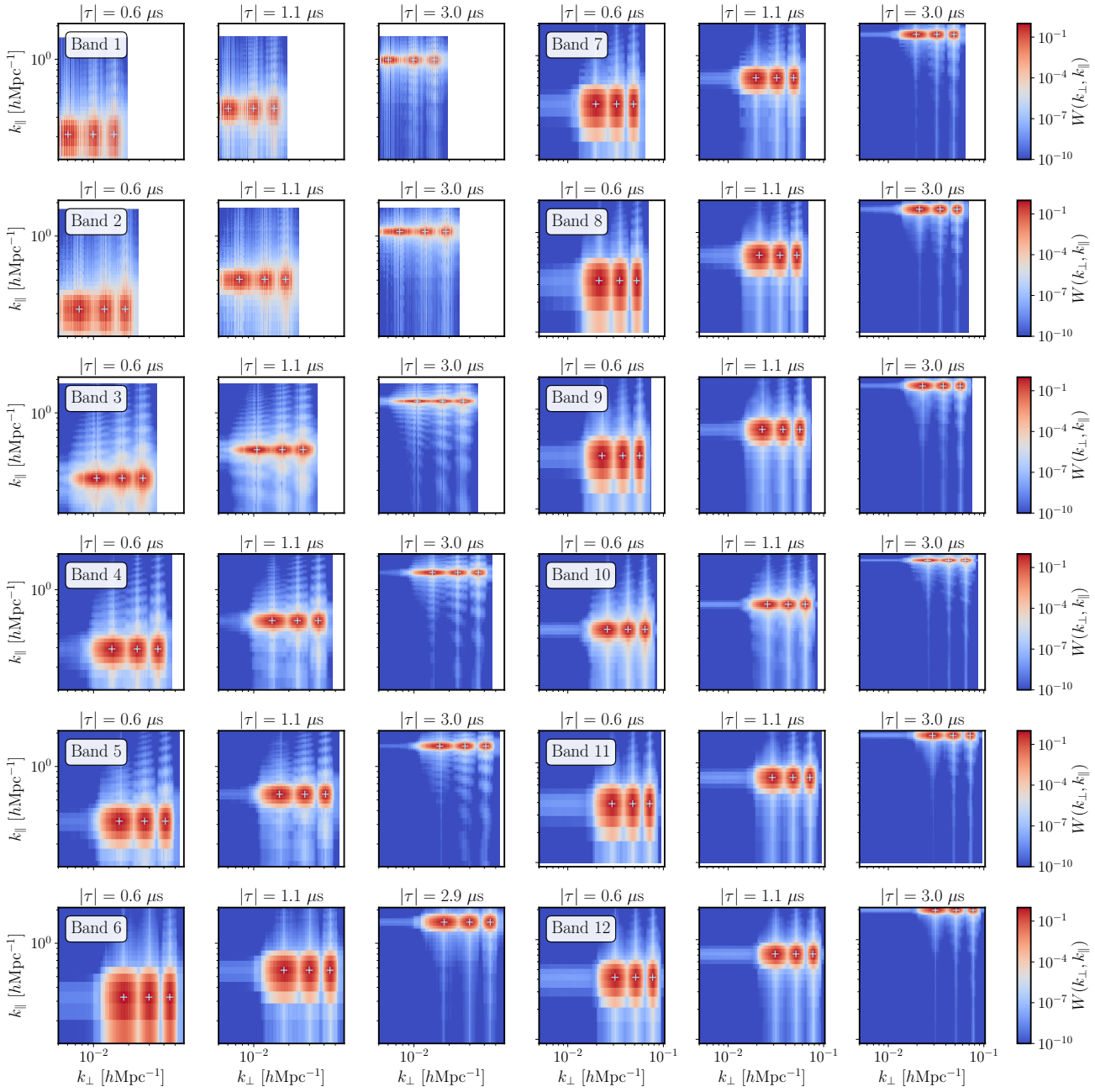

Figure 6: Cylindrical window functions obtained according to equation (12) for different baseline lengths (38.7, 63.7, 95.8 m) and for the simulated HERA Vivaldi beam along all the bands considered in the H4C analysis. Each panel corresponds to a different delay.

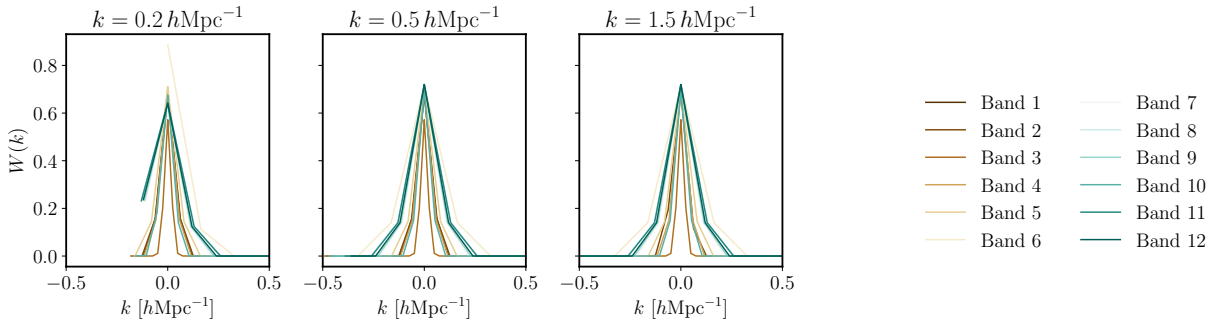

Figure 7: Spherical window functions obtained according to equation (12) for the simulated HERA Vivaldi beam along all the bands considered in the H4C analysis but for a dataset corresponding to H1C IDR3 baselines and no data weights applied. Each panel corresponds to a different spherical bin.

cylindrical and spherical window functions. For a discussion of the impact of tapering on foreground leakage, see Thyagarajan et al. (2013).

> *Code integration:* The taper is an attribute of the `UVWindow` object and can be defined when initialising the object. It must be readable by `uvtools.dspec.gen_window`. It is applied to $\tilde{A}(\mathbf{q}_\perp, \nu)$ when taking the FT along frequency, in `_take_freq_FT`. Default value is a Blackman-Harris taper.

3.1.1 Accounting for channel width

In practice, the visibility is not measured at an exact frequency, but over a frequency bin called a frequency channel of width $\Delta\nu$. Hence, the measured visibility is an integral such that

$$V^{\text{meas.}}(\mathbf{b}, \nu) = \int_{\nu-\Delta\nu/2}^{\nu+\Delta\nu/2} \mathrm{d}\nu'\ V(\mathbf{b}, \nu'), \tag{23}$$

which we can also write in terms of a top-hat function (or taper) ϕ which will be one inside the bin and zero outside:

$$V^{\text{meas.}}(\mathbf{b}, \nu) = \int \mathrm{d}\nu'\ V(\mathbf{b}, \nu')\ \phi\left[\frac{2(\nu - \nu')}{\Delta\nu}\right] \tag{24}$$

which can be written as a convolution:

$$V^{\text{meas.}}(\mathbf{b}, \nu) = V(\mathbf{b}, \nu) * \phi(2\nu/\Delta\nu). \tag{25}$$

By definitition, the delay visbility is the Fourier transform of the measured visibility along frequency. As the Fourier transform of a convolution, it is simply the product of the two Fourier transforms:

$$V(\mathbf{b}, \tau) = \frac{\Delta\nu}{2}\tilde{\phi}\left[\frac{\Delta\nu}{2}\tau\right] \times \int \mathrm{d}\nu\, V(\mathbf{b}, \nu)\ \mathrm{e}^{-2i\pi\nu\tau}, \tag{26}$$

where we used the fact that the Fourier transform of $\phi(\nu/\alpha)$ is $\alpha\tilde{\phi}(\tau\alpha)$. The window functions are modified accordingly:

$$W(\mathbf{k}; \mathbf{b}, \tau) = |\chi(\mathbf{k}; \mathbf{b}, \tau)|^2 \times \left|\frac{\Delta\nu}{2}\tilde{\phi}\left[\frac{\Delta\nu}{2}\tau\right]\right|^2 \tag{27}$$

Hence, the taper acts as a normalisation factor for the window function and will have no impact here.

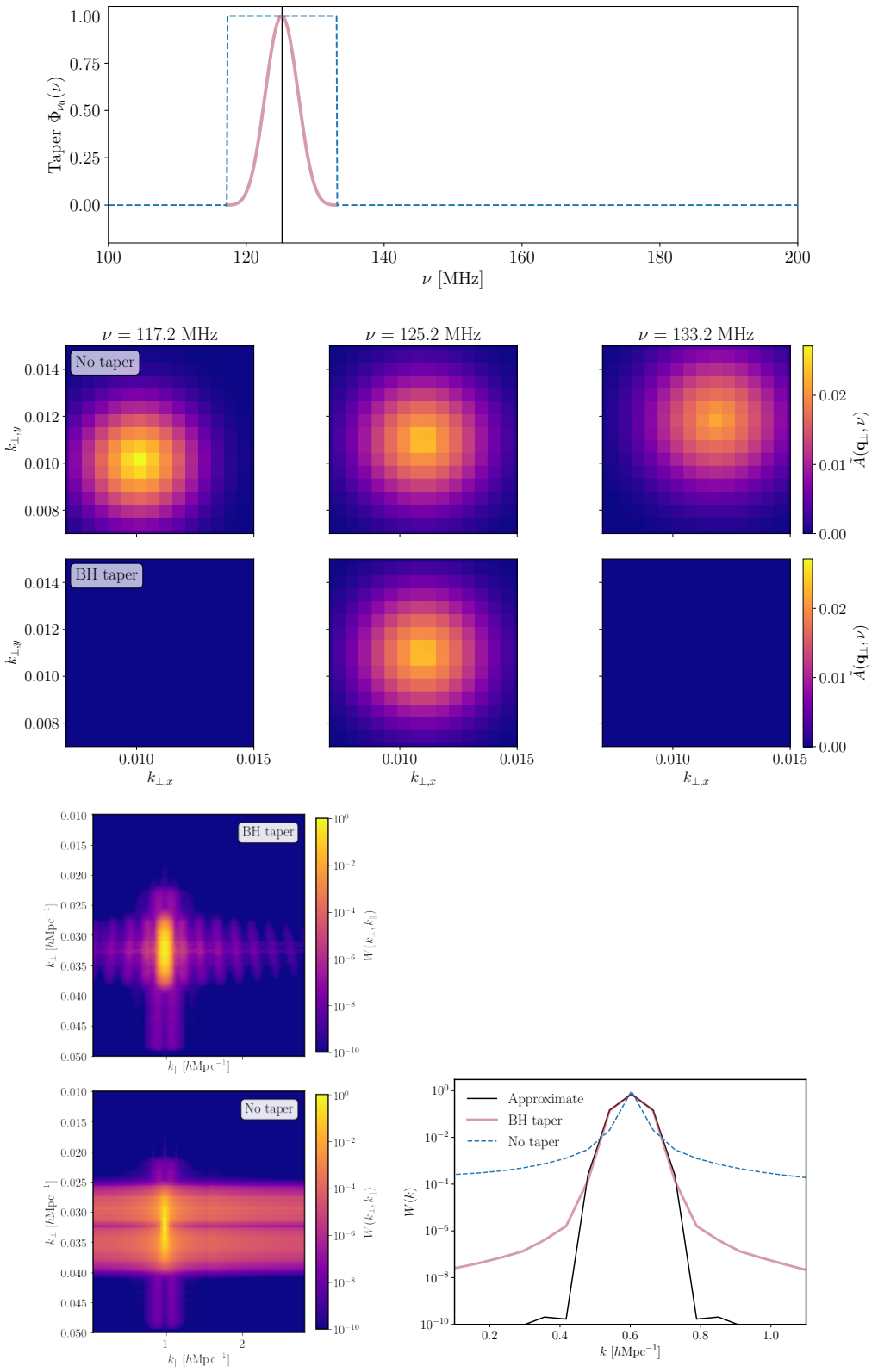

Figure 8: Impact of spectral tapering on the exact window functions. *Top panel:* Frequency tapers considered, illustrated along Band 1: No taper (step function, dashed line) or Blackman-Harris taper (BH, solid line). *Middle panel:* Frequency-dependence of the spatial Fourier transform of the beam along Band 1 with and without applying a taper along the frequency axis. *Bottom panel:* Comparison of the window functions obtained for the HERA beam with and without taper along Band 1, in cylindrical space for $b = 81.3\,\mathrm{m}$ (left) and in spherical space (right).

3.2 Binning in k-space

The technical specifications of HERA will tell us how to choose the ranges used for the cylindrical binning of the window functions. In the Phase I upper limit paper, the k-bins have a width $\Delta k = 0.053$ Mpc^{-1}.

In the sky plane The angular resolution θ of the instrument is roughly given by $1/\Delta\theta \sim b/\lambda$, such that the best resolution is given by the largest baseline in the interferometer and the smallest wavelength (or largest frequency) measured by the instrument. HERA's maximum baseline has length $b_{\rm max} = 139.36$ m and the maximum frequency probed is $\nu_{\rm max} = 199.90$ MHz (corresponding to $z_{\rm min} = 6.11$), hence the best resolution one can achieve with the instrument is

$$\Delta\theta_{\rm max} = \frac{c}{\nu_{\rm max} b_{\rm max}} \sim 0.011 \text{ rad} = 37.8'.$$

Ideally, we want a beam simulation with resolution around $40'$ or lower. However, the currently available beam simulation has a resolution (in azimuth and zenith angles) of about $60'$. If we approximate the HERA beam by a Gaussian beam, any resolution can be achieved.

Numerically, a map of a given resolution $\Delta\theta$ and length L in real space will given to Fourier modes ranging from 0 to $k_{\rm max}$ such that $2\pi/k_{\rm max} = d_c(z)\Delta\theta$, with resolution $\Delta k = 1/L$. Since we can extend the beam simulation (the beam being zero at the borders) to any size, we are not limited in our choice of bin width for $k_\perp$. However, we will be limited in the maximum value of $k_\perp$ that we can probe. We have

$$k_{\rm max}(b,\nu) = \frac{2\pi}{d_c(z)\ \Delta\theta} = \frac{2\pi\nu b}{d_c(\nu)c},$$

and the largest $k_\perp$ probed by the telescope will be by the longest baseline, at the maximum frequency and is $k_{\rm max} = 0.100$ Mpc^{-1}. Given the resolution of the beam simulation, we can probe

$$k_\perp \leq 0.063 \text{ Mpc}^{-1}.$$

For limited spectral window, this value differs. For Band 1 and 2, we have, respectively, $k_{\rm max} = 0.056$ Mpc^{-1} and $k_{\rm max} = 0.059$ Mpc^{-1}.

Finally, we want to make sure that the binning can resolve the shift induced by the chromaticity of the instrument. For each frequency, the maximum of the beam is located at the centre of the image, that is at $\mathbf{r} = 0$. The maximum of the FT will also be at the centre of the FT image, that is at $\mathbf{q} = 0$, that is, according to equation (14), at $\mathbf{k}_\perp = \frac{2\pi}{d_c(z)} \times \frac{b\nu}{c}$ as illustrated on Fig. 2. A minimal requirement is therefore to resolve the shift along the whole bandwidth for any baseline b, that is

$$\Delta k_\perp \leq \frac{2\pi}{d_c(z)} \times \frac{bB_0}{c}$$

for B_0 the bandwidth. Considering the smallest baseline, we obtain for band 1 and band 2, respectively, $\Delta k_\perp \leq 0.0007$ Mpc^{-1} and $\Delta k_\perp \leq 0.0009$ Mpc^{-1}. At best, there is one pixel shift between each frequency bin:

$$\Delta k_\perp = \frac{2\pi}{d_c(z)} \times \frac{b\Delta\nu}{c}$$

for $\Delta\nu$ the frequency resolution of the instrument. Such a condition requires a resolution of $\Delta k_\perp \leq 5 \times 10^{-6}$ Mpc^{-1}, which will lead to oversampling artefacts.

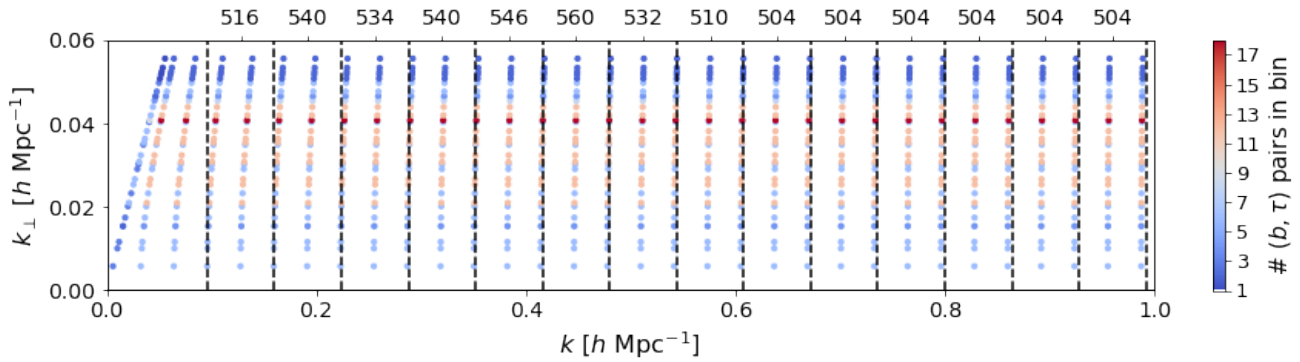

Figure 9: Distribution of the $k_\perp$ and k values sampled by the baselines in HERA along Band 1 (see H1C IDR3.2 memo). Each point corresponds to a set of baseline-delay (b, τ) pairs, their number being given by the colour scale. Vertical lines represent the k-bins considered for the H1C IDR2 analysis. Each bin includes the number of baselines written above the figure (before flagging).

Along the line of sight On the other hand, the $k_\parallel$ binning is limited by the frequency resolution of the instrument ($\Delta\nu = 97.7$ kHz) and the bandwidth B_0 of the spectra window considered. For band 1 and 2, the bandwidth are, respectively, $B_1 = 20.5$ MHz and $B_2 = 15.5$ MHz. The resolution is given by

$$\Delta k_\parallel \equiv \frac{2\pi}{\alpha(z)} \times \frac{1}{B_0},$$

which gives, for Band 1 and 2: $\Delta k_\parallel = 0.0216$ Mpc^{-1}.

The maximum value of $k_\parallel$ that can be sampled by the instrument is then

$$k_\parallel \leq \frac{2\pi}{\alpha(z)} \times \frac{1}{\Delta\nu},$$

that is $k_{\rm max} = 3.407$ Mpc^{-1} for band 1 and $k_{\rm max} = 3.849$ Mpc^{-1} for band 2.

Maximum $k_\perp$ and $k_\parallel$ probed Because the instrument has a given set of baselines and the spectra window chosen limits the range of delays probed when constructing the power spectrum, the range of $k_\perp$ and $k_\parallel$ will also be limited. Computing the ($k_\perp$, $k_\parallel$) bin centres associated with all possible combinations of τ and b, we find:

- For Band 1, $k_\perp < 0.08$ Mpc^{-1} and $k_\parallel < 1.70$ Mpc^{-1}, which correspond to $k < 1.70$ Mpc^{-1}.
- For Band 2, $k_\perp < 0.11$ Mpc^{-1} and $k_\parallel < 1.92$ Mpc^{-1}, which correspond to $k < 1.92$ Mpc^{-1}.

Knowing this, we choose to bin $k_\parallel$ on a range $0.0215 \leq k_\perp/\text{Mpc}^{-1} \leq 3.40$ with 158 bins of width $\Delta k_\parallel = 0.0215$ Mpc^{-1} and $k_\perp$ on a range $0.0005 \leq k_\perp/\text{Mpc}^{-1} \leq 0.088$ with 158 bins of width $\Delta k_\perp = 0.0005$ Mpc^{-1}.

> *Code integration:* These considerations are taken into account when the default $(k_\perp, k_\parallel)$ bins are used for cylindrical binning, through the `UVWindow.get_kpara_bins` and the `UVWindow.get_kperp_bins` methods.

3.3 Limitations and approximations

Different elements of this derivation and of the code can be improved on:
- The flat-sky approximation and
- The projection of the baseline vector (i.e. identifying $\mathbf{b} \equiv b$).

This will be the focus of future work.

References

Fagnoni, N., de Lera Acedo, E., DeBoer, D. R., et al. 2021a, MNRAS, 500, 1232

Fagnoni, N., de Lera Acedo, E., Drought, N., et al. 2021b, IEEE Transactions on Antennas and Propagation, 69, 8143

Gorce, A., Ganjam, S., Liu, A., et al. 2023, MNRAS, 520, 375

Thyagarajan, N., Udaya Shankar, N., Subrahmanyan, R., et al. 2013, ApJ, 776, 6