Antenna Weighting for Reducing Channel Time Variation in High-Mobility Massive MIMO

Yinghao Ge, Weile Zhang, Feifei Gao, Shun Zhang, and Xiaoli Ma

Abstract

In this paper, we derive the exact channel power spectrum density (PSD) for the high-mobility massive multiple-input multiple-output (MIMO) uplink from a high-speed terminal (HST) to a base station (BS) with angle-domain Doppler shifts compensation. A large-scale linear antenna array is equipped at the HST to separate multiple Doppler shifts in angle domain via high-resolution transmit beamforming. Each beamforming branch comprises a dominant Doppler shift, which can be compensated to suppress the time variation. We derive the PSD and the Doppler spread as a measure of assessing the residual time variation of the resultant uplink channels. Interestingly, the derived channel PSD can be exactly expressed as the product of a beam function and a window function. The beam function reflects the impact of array configuration and is in fact the scaled radiation pattern of the matched-filter beamformer towards the normal direction, while the window function depends on how the beamforming directions are configured. Such characteristic of the PSD facilitates investigating the impact of some essential parameters such as antenna spacing and beamforming directions on the PSD. Inspired by the PSD analysis, we further propose an antenna weighting technique to reduce the time variation of the channel by equivalently modifying the beam function. The optimal antenna weights that minimize the Doppler spread is derived in closed-form. Numerical results are provided to corroborate both the PSD analysis and the superiority of antenna weighting technique.

Y. Ge and W. Zhang are with the MOE Key Lab for Intelligent Networks and Network Security, School of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an, Shaanxi, 710049, China. (Email: ge_yinghao_jacques@163.com, wlzhang@mail.xjtu.edu.cn).

F. Gao is with the State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Automation, Tsinghua University, Beijing, 100084, China (Email: feifeigao@ieee.org).

S. Zhang is with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi’an, Shaanxi, 710071, China (Email: zhangshunsdu@gmail.com).

X. Ma is with Georgia Institute of Technology, Atlanta, GA 30332, USA (Email: xiaoli@gatech.edu).
Index Terms

High-mobility communication, time-varying channel, power spectrum density (PSD), Doppler spread, angle-domain massive MIMO, antenna weighting technique.

I. INTRODUCTION

Over the past few decades, high-mobility communications have drawn exploding interests from researchers [1]–[4]. The relative motion between transceivers can pose great challenges for communication, including frequent handovers and multiple Doppler shifts. The multiple Doppler shifts superimpose at the receiver, resulting in fast time fluctuations of channel and bringing severe inter-carrier interference (ICI) to the orthogonal frequency division multiplexing (OFDM) systems [5].

The fast time-varying feature of channel makes the direct channel estimation quite complicated and even infeasible. Some works employ the basis expansion model (BEM) [6]–[10] to approximately represent the fast time-varying channel, such that the parameters to be estimated are significantly reduced. Another kind of frequently adopted approaches approximate the channel autocorrelation as the weighted summation of two monochromatic plane waves [11]–[13]. Considering that each Doppler shift is related to an angle-of-arrival (AoA) for downlink or angle-of-departure (AoD) for uplink, the multiple Doppler shifts can be separated in angle domain. Such concept could be found in [14], [15], where the small-scale uniform circular antenna array (UCA) and uniform linear antenna array (ULA) are adopted to separate the multiple Doppler shifts and eliminate ICI via array beamforming. However, due to the limited spatial resolution, their work only applies to high-mobility scenarios with a few dominating paths, such as viaducts and rural areas.

In order to deal with the richly scattered high-mobility scenarios including tunnels or urban areas, researchers resort to the large-scale antenna array, which is considered as a promising technique for the next generation wireless systems owing to its enhanced spectral and energy efficiency as well as high spatial resolution [16]–[20]. The authors of [21] propose to separate the
multiple downlink Doppler shifts in angle domain by a pre-designed beamforming network with a large-scale ULA at the base station (BS). After estimating and compensating the Doppler shift in each branch, the resultant channel turns to be quasi time-invariant and can be estimated with conventional channel estimation approaches. The array imperfection is further taken into account in [22], and the multi-Doppler shift separation via array beamforming can be done after array calibration. Unlike [21] and [22] which address the multiple downlink Doppler shifts, [23] focuses on the uplink from the high-speed terminal (HST) to BS, where the Doppler shifts are related to AoDs instead of AoAs. As a result, a large-scale ULA is configured at the HST to perform high-resolution transmit beamforming, and the multi-branch signal is emitted after compensating the multiple Doppler shifts in angle domain to suppress the time variation of channel. In practice, however, the number of antennas may not be sufficiently large to generate beamformers with infinite spatial resolution. Thus, the Doppler shifts intermingled via the sidelobes cannot be completely compensated, resulting in the residual time variation of the uplink channel. The Doppler spread is further derived in [23] as a measure of assessing the time variation caused by the residual Doppler shifts. However, the analysis in [23] is approximative and only valid contingent on 1) the array is a large-scale ULA, 2) the channel is Jakes’ channel [24], [25] and 3) the beamforming directions are evenly configured.

In this paper, we apply the multi-branch transmit beamforming and angle-domain Doppler shifts compensation scheme to a far more generalized high-mobility scenario and analyze the power spectrum density (PSD) and Doppler spread to assess the residual time variation of the uplink equivalent channel. Based on the PSD analysis, an optimal antenna weighting technique is further proposed to reduce the time variation of channel. The main contributions of this paper can be summarized as follows:

- Explicit PSD expression with wider applicability and clearer insights: Unlike [23], we do not require the channel to follow Jakes’ model; neither should the antenna array be ULA. The most interesting observation is that the channel PSD can always be expressed as the product of a beam function and a window function. The former can be uniquely determined
by the antenna array configuration and in fact corresponds to the radiation pattern obtained
with the matched-filter (MF) beamformer pointing to the normal direction, while the latter
depends on the AoD region and the beamforming directions. This allows us to observe how
the antenna spacing and beamforming directions influence the PSD.

- Reduction of Doppler spread through antenna weighting: The PSD being expressed as the
  product of two independent functions facilitates to reduce the Doppler spread. By carefully
designing the optimal antenna weights, we can equivalently modify the beam function and
minimize the Doppler spread to suppress the residual time variation. The numerical results
demonstrate the substantial superiority of the proposed antenna weighting technique.

The rest of this paper is organized as follows. The transmit array beamforming and Doppler
shifts compensation scheme under a more generalized high-mobility scenario is briefly described
in Section II. Section III gives the detailed derivation of the channel PSD and Doppler spread,
based on which the impact of antenna spacing and beamforming directions is investigated. The
antenna weighting technique, especially the computation of optimal antenna weights, is presented
in Section IV. Simulation results are provided in Section V. Section VI concludes the paper.

**Notations:** Superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$ and $E\{\cdot\}$ represent conjugate, transpose, Her-
mitian, inverse and expectation, respectively; $j = \sqrt{-1}$ is the imaginary unit; $|\cdot|$ denotes the
absolute value operator; $\|\cdot\|_2$ denotes the Euclidean norm of a vector or Frobenius norm of a
matrix; $\otimes$ denotes the Kronecker product operator; $\text{diag}(\mathbf{x})$ is a diagonal matrix with vector $\mathbf{x}$ as
the main diagonal; $\mathbb{C}^{m \times n}$ defines the vector space of all $m \times n$ complex matrices; $\mathbf{I}_N$ stands for
the $N \times N$ identity matrix. For the positive semi-definite Hermitian matrix $\mathbf{X}$, $\lambda_r(\mathbf{X})$ and $\mathbf{v}_r(\mathbf{X})$
denote the $r$th eigenvalue (in descending order) and the corresponding eigenvector, respectively.

**II. System Model**

Consider the OFDM uplink transmission in a high-mobility scenario where the signal trans-
mitted from the HST arrives at the BS along a number of independent subpaths, as illustrated in
Fig. 1. The HST is equipped with an $M$-elements linear antenna array. Assume that the direction
of the linear array coincides with that of HST motion. Then, the array response vector pointing
to direction $\theta$ can be expressed as $a(\theta) = \left[ a_1(\theta), \ a_2(\theta), \ \ldots, \ a_M(\theta) \right]^T$, where the $r$th element is given by $a_r(\theta) = e^{j2\pi \Delta d_r \cos \theta}$. Here, $\Delta d_r$ is the antenna spacing between the $r$th antenna and the first antenna, and we have $\Delta d_1 = 0$. By denoting the velocity of HST as $v$, the maximum Doppler shift $f_d$ can be defined as $f_d = \frac{v}{\lambda}$, where $\lambda$ is the carrier wavelength.

The channel between the $r$th antenna and BS is modeled as $L$ taps. Denote $d_l$ as the relative delay of the $l$th tap. For each channel tap, the departure angles are known to be constrained within $(\theta_L, \theta_R)$, with $\theta_L$ and $\theta_R$ being the bounds of the AoD region. Similar to [26], [27], we denote $\kappa_l(\theta)$ as the complex-valued channel gain at the $l$th tap corresponding to the AoD $\theta$. The channels with different AoDs are assumed uncorrelated, i.e., $E\{\kappa_l(\theta)\kappa_l^*(\theta')\} = \frac{1}{\pi} \rho_l(\theta) \delta(\theta - \theta')$, where $\rho_l(\theta)$ represents the channel power angle spectrum (PAS) which models the channel power distribution in the angular domain [28], [29]. Moreover, there holds $\int_{\theta_L}^{\theta_R} \rho_l(\theta) d\theta = 1, \forall l = 1, 2, \ldots, L$ such that the total channel gain is normalized to 1. The maximum AoD region $(\theta_L, \theta_R)$ is assumed known at the transmitter, while the knowledge about the channel PAS might be unavailable.

Denote $s_m = \left[ s_m(0), \ s_m(1), \ \ldots, \ s_m(N-1) \right]$ as the length-$N$ transmitted time domain symbols in the $m$th OFDM block. The cyclic prefix (CP) of length $N_{cp}$ is appended to $s_m$, which implies that $s_m(-n) = s_m(N-n)$ for $n = 1, 2, \ldots, N_{cp}$. Then, the transmitted signal matrix
at the transmit antenna array after delay of $d_l$ can be expressed as $S_m (d_l) = \mathbf{1}_{M \times 1} \otimes s_m (d_l) / \sqrt{M}$, where $s_m (d_l) = \left[ s_m (-d_l), s_m (1-d_l), \ldots, s_m (N-1-d_l) \right]$ corresponds to the right circular shift of $s_m$ by a factor of $d_l$. Here, the divisor $\sqrt{M}$ is added to keep the total transmit power per symbol to 1. Moreover, define $N_s = N + N_{cp}$ as the length of a whole OFDM block.

Let the transmitted signal pass through the above-described channel. The received signal in the $m$th block (after CP removal) at the BS without Doppler shifts compensation can be expressed as the following $1 \times N$ vector

$$
y_m = \sum_{l=1}^{L} \int_{\theta_l}^{\theta_R} \kappa_l(\theta) \mathbf{a}^T(\theta) S_m (d_l) \Phi_m (\theta) d\theta + n_m,
$$

where $\Phi_m (\theta) = \text{diag} \left( \beta_{m,0}(\theta), \beta_{m,1}(\theta), \ldots, \beta_{m,N-1}(\theta) \right)^T$ represents the phase rotation matrix induced by Doppler shift and $\beta_{m,n}(\theta) = e^{j2\pi f_d \cos \theta (mN_s + n - d_l) T_s}$. Here, $T_s$ is the sampling interval. Besides, $n_m \in \mathbb{C}^{1 \times N}$ is the zero-mean complex additive white Gaussian noise (AWGN) in the $m$th block at the BS with $E\{n_m^H n_m\} = \sigma_n^2 I_N$, where $\sigma_n^2$ is the noise power.

Since the signal AoDs are constrained within $(\theta_L, \theta_R)$, we perform the multi-branch transmit beamforming towards a set of $Q$ directions $\vartheta_q \in (\theta_L, \theta_R), q = 1, 2, \ldots, Q$. Moreover, assume that the maximum Doppler shift is perfectly known at the HST. Then, the transmit beamforming and Doppler shifts compensation can be performed by substituting $S_m (d_l)$ with $\tilde{S}_{m,q} (d_l) = b^*(\vartheta_q) s_m (d_l) \Psi_{m,q} (d_l)$, where $b(\vartheta_q) = \frac{\eta}{M \sqrt{Q}} \mathbf{a}(\vartheta_q) e^{j \phi(\vartheta_q)}$ represents the $q$th beamformer. Here, $\phi(\vartheta_q)$ denotes the random phase introduced at $b(\vartheta_q)$, and $\eta = \frac{1}{\sqrt{Q} \left\| \sum_{q=1}^{Q} \mathbf{a}(\vartheta_q) e^{j \phi(\vartheta_q)} \right\|}$ is the normalization coefficient to keep the total transmit power per symbol to 1. The Doppler shift compensation matrix is $\Psi_{m,q} (d_l) = \text{diag} \left( \tilde{\beta}_{m,0,q}(d_l), \tilde{\beta}_{m,1,q}(d_l), \ldots, \tilde{\beta}_{m,N-1,q}(d_l) \right)^T$, where $\tilde{\beta}_{m,n,q}(d_l) = e^{-j2\pi f_d \cos \theta_q (mN_s + n - d_l) T_s}$.

Substituting $S_m (d_l)$ with $\tilde{S}_{m,q} (d_l)$ in (1), we arrive at

$$
r_{m,q} = \sum_{l=1}^{L} \int_{\theta_l}^{\theta_R} \kappa_l(\theta) \mathbf{a}^T(\theta) b^*(\vartheta_q) s_m (d_l) \Psi_{m,q} (d_l) \Phi_m (\theta) d\theta + n_m,
$$

$$
= \frac{\eta}{\sqrt{Q}} e^{-j \vartheta_q} \sum_{l=1}^{L} \int_{U(\vartheta_q)} \kappa_l(\theta) s_m (d_l) d\theta
$$

where $U(\vartheta_q)$ is the desired signal
\[
\sum_{l=1}^{L} \int_{(\theta_L, \theta_R) \cap U(\vartheta_q)} \kappa_l(\theta) \mathbf{b}^H(\vartheta_q) \, \mathbf{a}(\theta) \, s_m(d_l) \, \Psi_{m,q}(d_l) \, \Phi_m(\theta) \, d\theta + n_{m, \text{noise}}, \tag{2}
\]

where \(U(\vartheta_q) = \{\theta \mid \vartheta_q - \Delta\theta_1(\vartheta_q) \leq \theta \leq \vartheta_q + \Delta\theta_2(\vartheta_q)\}\) denotes the neighbourhood of \(\theta\) around \(\vartheta_q\), which is small enough for \(\mathbf{a}^T(\theta) \, \mathbf{a}^*(\vartheta_q) = M\) and \(\Psi_{m,q}(d_l) \, \Phi_m(\theta) = I_N\) to approximately hold true. Here, \(\Delta\theta_i(\vartheta_q), i = 1, 2\) are two appropriately chosen small values depending on \(\vartheta_q\).

When the number of antennas \(M\) is massive, the interference in (2) tends to vanish, and the time-varying channel can be decomposed into a set of parallel time-invariant channels. However, the number of antennas may not be sufficiently large in practice, in which case there will still be uncompensated Doppler shifts due to limited spatial resolution while a thorough time-invariant equivalent channel cannot be achieved for each beamforming branch. The Doppler spread \([30], [31]\) could be employed here as a metric to measure the time variation of the equivalent channel.

The derivation of Doppler spread requires the channel PSD, which is the Fourier Transform of the channel autocorrelation. Since different channel taps are independent and have identical statistical properties, we can consider only one tap for simplicity, i.e., \(L = 1\), \(d_1 = 0\). By ignoring the noise item, the signal at the BS obtained after Doppler shifts compensation and multi-branch beamforming can be expressed as

\[
r_m = \sum_{q=1}^{Q} r_{m,q} = \sum_{q=1}^{Q} \int_{\theta_L}^{\theta_R} \kappa_1(\theta) \mathbf{a}^T(\theta) \mathbf{b}^*(\vartheta_q) \, s_m(d_l) \, \Psi_{m,q}(d_l) \, \Phi_m(\theta) \, d\theta,
\]

\[
= \frac{\eta}{\sqrt{Q}} \sum_{q=1}^{Q} \int_{\theta_L}^{\theta_R} \kappa(\theta) \, e^{-j\varphi(\vartheta_q)} \frac{1}{M} \mathbf{a}^H(\vartheta_q) \, \mathbf{a}(\theta) \, s_m(\Psi_{m,q}(d_l) \, \Phi_m(\theta)) \, d\theta, \tag{3}
\]

where \(\kappa_1(\theta)\) has been denoted as \(\kappa(\theta)\) for conciseness. Besides, the complex-valued channel gain \(\kappa(\theta)\) can be further rewritten as \(\kappa(\theta) = \alpha(\theta) e^{j\varphi(\theta)}\), where \(\alpha(\theta) \sim \mathcal{N}(0, \rho(\theta))\) and \(\varphi(\theta) \sim \mathcal{U}(0, \pi)\) denote the random channel gain and phase for the path with AoD \(\theta\), respectively.

### III. Analysis of the Channel PSD and Doppler Spread

#### A. Derivation of the Channel PSD

The equivalent uplink channel of (3) can be expressed in continuous-time form as

\[
g(t) = \frac{1}{\sqrt{Q}} \sum_{q=1}^{Q} \int_{\theta_L}^{\theta_R} \alpha(\theta) \, G(\cos \theta, \cos \vartheta_q) \, e^{j2\pi f_d t \cos \theta - j2\pi f_d t \cos \vartheta_q + j\varphi(\vartheta_q) - j\varphi(\vartheta_q)} \, d\theta,
\]
where \( \omega_d = 2\pi f_d \) and \( G(\cos \theta, \cos \vartheta_q) = \frac{1}{M} \mathbf{a}^H(\vartheta_q) \mathbf{a}(\theta) = \frac{1}{M} \sum_{r=1}^{M} e^{j2\pi \Delta \vartheta_q (\cos \theta - \cos \vartheta_q)} \). Note that the normalization coefficient \( \eta \) is omitted in the continuous-form channel \( (4) \) for simplicity, since it does not affect the following PSD analysis. Besides, \(|G(\cos \theta, \cos \vartheta_q)|^2\) is in fact the radiation pattern at direction \( \theta \) with \( \frac{1}{M} \mathbf{a}(\vartheta_q) \) as beamformer. Moreover, by fixing \( \vartheta_q = \frac{\pi}{2} \) and varying \( \theta \), \(|G(\cos \theta, \cos \vartheta_q)|^2\) will be exactly the radiation pattern obtained with the MF beamformer pointing towards the normal direction of the linear array.

To simplify the derivation of PSD, we will assume the uniform PAS, i.e., \( E\{\rho(\theta)\} = \frac{1}{\theta_R - \theta_L}, \forall \theta \in (\theta_L, \theta_R) \) in the following. Nevertheless, the proposed analysis in this paper could be easily extended to general channel PAS. In the case of uniform PAS, denote \( \theta_{as} = \frac{\theta_R - \theta_L}{2} \) and \( \bar{\theta} = \frac{\theta_L + \theta_R}{2} \) as the mean AoD and angular spread (AS), respectively. Furthermore, the described channel model reduces to Jakes’ channel model \([25]\) at \( \theta_L = 0, \theta_R = \pi \). The autocorrelation for the equivalent continuous channel \( g(t) \) is given by

\[
R_g(\tau) = E\{g(t) g^*(t + \tau)\},
\]

\[
= \frac{1}{Q} \sum_{q=1}^{Q} \sum_{k=1}^{Q} \int_{\theta_L}^{\theta_R} \int_{\theta_L}^{\theta_R} E\{\alpha(\theta) \alpha^*(\bar{\theta}) e^{j[\phi(\theta) - \phi(\bar{\theta})]} e^{j[\phi(\vartheta_q) - \phi(\bar{\theta})]}G(\cos \theta, \cos \vartheta_q) G^*(\cos \bar{\theta}, \cos \vartheta_k) e^{j\omega_d(\cos \theta - \cos \vartheta_q) t} e^{-j\omega_d(\cos \bar{\theta} - \cos \vartheta_k)(t + \tau)}\} d\theta d\bar{\theta},
\]

\[
= \frac{1}{Q} \sum_{q=1}^{Q} \int_{\theta_L}^{\theta_R} E\{\alpha(\theta) \}^2 |G(\cos \theta, \cos \vartheta_q)|^2 e^{-j\omega_d(\cos \theta - \cos \vartheta_q) \tau} d\theta,
\]

\[
= \frac{1}{(\theta_R - \theta_L) Q} \sum_{q=1}^{Q} \int_{\theta_L}^{\theta_R} |G(\cos \theta, \cos \vartheta_q)|^2 e^{-j\omega_d(\cos \theta - \cos \vartheta_q) \tau} d\theta,
\]

where \( = \) employs the properties

\[
E\{e^{j[\phi(\theta) - \phi(\bar{\theta})]}\} = \begin{cases} 1, & \theta = \bar{\theta} \\ 0, & \theta \neq \bar{\theta} \end{cases}, \quad E\{e^{j[\phi(\vartheta_q) - \phi(\bar{\theta})]}\} = \begin{cases} 1, & q = k \\ 0, & q \neq k \end{cases}.
\]

and \( \approx \) comes from \( E\{\alpha(\theta) \}^2\).

The channel PSD is the Fourier transform of the channel autocorrelation \( R_g(\tau) \) and the explicit expression of channel PSD is provided by the following Lemma.
Lemma 1: Let \( \omega \) be the Doppler frequency and denote \( \tilde{\omega} = \frac{\omega}{\omega_d} = \frac{\omega}{2\pi f_d} \) as the normalized Doppler frequency with respect to the maximum Doppler shift. Then, for the given channel autocorrelation \( R_g(\tau) \) in (5), the channel PSD can be expressed in the form of

\[
P(\omega) = \frac{1}{\omega_d} |G(\tilde{\omega})|^2 W(\tilde{\omega}),
\]

where

\[
|G(\tilde{\omega})|^2 = \left| \frac{1}{M} \sum_{r=1}^{M} e^{-j2\pi \Delta \varphi_r} \right|^2,
\]

and

\[
W(\tilde{\omega}) = \frac{2\pi}{(\theta_R - \theta_L)Q} \sum_{q=1}^{Q} \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta_q)^2}} T_q(\tilde{\omega}),
\]

are named as beam function and window function, respectively. Here, \( T_q(\tilde{\omega}) = \begin{cases} 1, & q \in S(\tilde{\omega}) \\ 0, & q \notin S(\tilde{\omega}) \end{cases} \) is the binary-value indicator function indicating whether the \( q \)th beamforming branch contributes to the PSD \( P(\omega) \) at \( \omega = \omega_d \tilde{\omega} \), with \( S(\tilde{\omega}) \) being defined as

\[
S(\tilde{\omega}) = \{ q \mid \tilde{\omega} + \cos \theta_R \leq \cos \vartheta_q \leq \tilde{\omega} + \cos \theta_L, \ \cos \theta_R \leq \cos \vartheta_q \leq \cos \theta_L \},
\]

\[
= \begin{cases} 
\{ q \mid \cos \theta_R \leq \cos \vartheta_q \leq \tilde{\omega} + \cos \theta_L \}, & -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\
\{ q \mid \tilde{\omega} + \cos \theta_R \leq \cos \vartheta_q \leq \cos \theta_L \}, & 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R) 
\end{cases},
\]

where \( \mu(\theta_L, \theta_R) = \cos \theta_L - \cos \theta_R \).

Proof: See Appendix A.

From Lemma 1, the following observations can be made:

1) From (9), it can be seen that the PSD is nonzero only for \( |\tilde{\omega}| \leq \mu(\theta_L, \theta_R) = \cos \theta_L - \cos \theta_R \). Obviously, the maximum Doppler frequency \( \omega_{\text{max}} \), or equivalently the nonzero region of the PSD, is uniquely determined by the AoD region.

2) The most interesting observation from (6) is that the channel PSD can be fully characterized by \( |G(\tilde{\omega})|^2 \) and \( W(\tilde{\omega}) \). Taking \( \vartheta_q = \frac{\pi}{2} \) and \( -\tilde{\omega} = \cos \theta - \cos \vartheta_q = \cos \theta \), we can arrive at \( G(\cos \theta, \cos \frac{\pi}{2}) = G(\tilde{\omega}) \), which implies that \( |G(\tilde{\omega})|^2 \) corresponds to the radiation pattern obtained with the MF beamformer pointing to the normal direction. This explains why \( |G(\tilde{\omega})|^2 \) is named as beam function. Besides, \( T_q(\tilde{\omega}) \) is the binary-value indicator function indicating whether the \( q \)th
beamforming branch contributes to the PSD at \( \tilde{\omega} \). Therefore, \( \mathcal{W}(\tilde{\omega}) \) reflects the comprehensive impact of different beamformers on the PSD, and is named as window function to highlight its distortion effect on the beam function \( |G(\tilde{\omega})|^2 \). Moreover, the beam function \( |G(\tilde{\omega})|^2 \) only depends on the antenna array configuration (more specifically, antenna spacings \( \Delta d_n \)), and for a given AoD region, the window function \( \mathcal{W}(\tilde{\omega}) \) is determined by the configuration of beamforming directions \( \vartheta_q, q = 1, 2, \cdots, Q \).

3) The PSD in (6) can be equivalently written as \( P(\omega) = \frac{1}{\omega_d} \left| G\left(\frac{\omega}{\omega_d}\right)\right|^2 \mathcal{W}\left(\frac{\omega}{\omega_d}\right) \). Evidently, increasing \( \omega_d \), i.e., the maximum Doppler shift \( f_d \), will preserve the shape of the PSD, except that the resulting PSD will be linearly stretched in frequency and reversely decreased in amplitude.

Nevertheless, the integral of \( P(\omega) \) with respect to \( \omega \) is independent of \( \omega_d \), because of

\[
\int_{\Delta(\theta_L, \theta_R)\omega_d}^{\mu(\theta_L, \theta_R)\omega_d} P(\omega) d\omega = \int_{\mu(\theta_L, \theta_R)\omega_d}^{-\mu(\theta_L, \theta_R)\omega_d} \left| G\left(\frac{\omega}{\omega_d}\right)\right|^2 \mathcal{W}\left(\frac{\omega}{\omega_d}\right) d\omega = \int_{-\mu(\theta_L, \theta_R)\omega_d}^{\mu(\theta_L, \theta_R)\omega_d} \left| G\left(\frac{\tilde{\omega}}{\omega_d}\right)\right|^2 \mathcal{W}\left(\frac{\tilde{\omega}}{\omega_d}\right) d\tilde{\omega}.
\]

4) The Doppler spread can be calculated as

\[
\sigma_{DS} = \sqrt{\frac{\int_{-\mu(\theta_L, \theta_R)\omega_d}^{\mu(\theta_L, \theta_R)\omega_d} \omega^2 P(\omega) d\omega}{\int_{-\Delta(\theta_L, \theta_R)\omega_d}^{\mu(\theta_L, \theta_R)\omega_d} P(\omega) d\omega}} = \sqrt{\frac{\int_{\mu(\theta_L, \theta_R)}^{-\mu(\theta_L, \theta_R)} \tilde{\omega}^2 |G(\tilde{\omega})|^2 \mathcal{W}(\tilde{\omega}) d\tilde{\omega}}{\int_{-\mu(\theta_L, \theta_R)}^{\mu(\theta_L, \theta_R)} |G(\tilde{\omega})|^2 \mathcal{W}(\tilde{\omega}) d\tilde{\omega}}}.
\]

Considering that the two integrals with respect to \( \tilde{\omega} \) in (10) does not depend on \( \omega_d \), we know that the Doppler spread \( \sigma_{DS} \) is linearly proportional to \( \omega_d \), i.e., the maximum Doppler shift \( f_d \).

In other words, the higher the HST velocity is, the larger the Doppler spread \( \sigma_{DS} \) will be.

B. Impact of Beamforming Directions and Antenna Array Configuration on PSD

In this section, we will discuss how the beamforming directions and antenna array configuration influence the channel PSD. Note that the signal AoD region \((\theta_L, \theta_R)\) reflects the intrinsic feature of the channel, which is uncontrollable.

1) Impact of beamforming directions: As the number of selected beamformers \( Q \) tends to infinity, the window function given in (8) can be transformed into the following integral form

\[
\mathcal{W}(\tilde{\omega}) = \frac{2\pi}{\theta_R - \theta_L} \frac{1}{Q} \sum_{q=1}^{Q} \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta_q)^2}} I_q(\tilde{\omega}),
\]

\[
= \frac{2\pi}{\theta_R - \theta_L} \int_{\theta_L}^{\theta_R} \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta)^2}} f(\vartheta) I(\vartheta, \tilde{\omega}) d\vartheta,
\]

(11)
where \( \vartheta \) is the continuous counterpart of \( \vartheta_q \), \( f(\vartheta) \) is the density function of \( \vartheta \), and \( \mathcal{I}(\vartheta, \tilde{\omega}) = \begin{cases} 1, & \vartheta \in S(\vartheta, \tilde{\omega}) \\ 0, & \vartheta \notin S(\vartheta, \tilde{\omega}) \end{cases} \) is the binary-value indicator function, with

\[
S(\vartheta, \tilde{\omega}) = \begin{cases}
\{ \vartheta \mid \cos \theta_R \leq \cos \vartheta \leq \tilde{\omega} + \cos \theta_L \}, & -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\
\{ \vartheta \mid \tilde{\omega} + \cos \theta_R \leq \cos \vartheta \leq \cos \theta_L \}, & 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R) \\
\{ \vartheta \mid \arccos(\tilde{\omega} + \cos \theta_L) \leq \vartheta \leq \theta_R \}, & -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\
\{ \vartheta \mid \theta_L \leq \vartheta \leq \arccos(\tilde{\omega} + \cos \theta_R) \}, & 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R) \end{cases}.
\]

\tag{12}

Note that \( S(\vartheta, \tilde{\omega}) \) can be directly derived from \( S(\tilde{\omega}) \), by substituting \( \vartheta_q \) in (9) with \( \vartheta \).

Next, we further derive a more explicit form of window function, under two typical configurations of beamforming directions: First, the beamforming directions are configured such that \( \cos \vartheta_q, q = 1, 2, \cdots, Q \) are evenly distributed between \( (\cos \theta_R, \cos \theta_L) \); second, the beamforming directions \( \vartheta_q, q = 1, 2, \cdots, Q \) themselves are evenly configured between \( (\theta_L, \theta_R) \). We refer to the two configurations of beamforming directions as ‘Equi-cos’ and ‘Equi-angle’, respectively.

Note that ‘Equi-cos’ is considered since the multi-branch beamforming with such configured beamformers can be implemented efficiently with fast Fourier transform (FFT) [20].

Case 1: In the case of ‘Equi-cos’, i.e., \( \cos \vartheta_q, q = 1, 2, \cdots, Q \) are evenly distributed between \( (\cos \theta_R, \cos \theta_L) \), the density function can be expressed as

\[
f(\vartheta) = \frac{\sin \vartheta}{\mu(\theta_L, \theta_R)}.
\]

\tag{13}

Here, the density function (13) should be in sinusoidal form since ‘Equi-cos’ distribution implies \( d \cos \vartheta = \sin \vartheta d\vartheta \), and the normalization term \( \mu(\theta_L, \theta_R) \) comes from \( \int_{\theta_L}^{\theta_R} \sin \vartheta d\vartheta = \mu(\theta_L, \theta_R) \).

As a result, the window function can be expressed as

\[
W(\tilde{\omega}) = 2\pi \frac{1}{\theta_R - \theta_L} \sum_{q=1}^{Q} \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta_q)^2}} \mathcal{I}_q(\tilde{\omega}),
\]

\[
= 2\pi \frac{1}{\theta_R - \theta_L} \mu(\theta_L, \theta_R) \int_{\theta_L}^{\theta_R} \frac{\sin \vartheta}{\sqrt{1 - (\cos \vartheta - \tilde{\omega})^2}} \mathcal{I}(\vartheta, \tilde{\omega}) d\vartheta.
\]

\tag{14}

To further simplify (14), we take a variable substitution of \( x = \arccos(\cos \vartheta - \tilde{\omega}) \), i.e., \( \vartheta = \arccos(\cos x + \tilde{\omega}) \). Then, the indicator function \( \mathcal{I}(\vartheta, \tilde{\omega}) \) becomes \( \mathcal{I}(\arccos(\cos x + \tilde{\omega}), \tilde{\omega}) = \).
The derived window function $I(\vartheta, \tilde{\omega})_{\vartheta = \arccos(\cos x + \tilde{\omega})}$, with the beamformer set $S(\vartheta, \tilde{\omega})$ being transformed into

$$S(\arccos(\cos x + \tilde{\omega}), \tilde{\omega}) = \begin{cases} \{ x \mid \theta_L \leq x \leq \arccos(\cos \theta_R - \tilde{\omega}) \}, & -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\ \{ x \mid \arccos(\cos \theta_L - \tilde{\omega}) \leq x \leq \theta_R \}, & 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R) \end{cases}. \quad (15)$$

After the variable substitution, (14) can be finally expressed in closed-form as

$$W(\tilde{\omega}) = \frac{2\pi}{\theta_R - \theta_L} \frac{1}{\mu(\theta_L, \theta_R)} \int_{\arccos(\cos \theta_R - \tilde{\omega})}^{\arccos(\cos \theta_L - \tilde{\omega})} I(\arccos(\cos x + \tilde{\omega}), \tilde{\omega}) \, dx,$$

$$= \begin{cases} 2\pi \frac{1}{\theta_R - \theta_L} \mu(\theta_L, \theta_R) \int_{\theta_L}^{\arccos(\cos \theta_R - \tilde{\omega})} I(\arccos(\cos x + \tilde{\omega}), \tilde{\omega}) \, dx, & -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\ 2\pi \frac{1}{\theta_R - \theta_L} \mu(\theta_L, \theta_R) \int_{\arccos(\cos \theta_L - \tilde{\omega})}^{\theta_R} I(\arccos(\cos x + \tilde{\omega}), \tilde{\omega}) \, dx, & 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R) \end{cases}. \quad (16)$$

Furthermore, if the channel follows Jakes’ channel model, we have $\theta_L = 0^\circ, \theta_R = 180^\circ$ and $\mu(\theta_L, \theta_R) = 2$. Then, the above derived window function (16) can be simplified as

$$W(\tilde{\omega}) = \arccos(|\tilde{\omega}|-1), \quad |\tilde{\omega}| \leq 2. \quad (17)$$

Case 2: In the case of ‘Equi-angle’, i.e., $\vartheta_q, q = 1, 2, \ldots, Q$ are evenly selected between $(\theta_L, \theta_R)$, the density function can be given by

$$f(\vartheta) = \frac{1}{\theta_R - \theta_L}. \quad (18)$$

As a result, the window function can be expressed as

$$W(\tilde{\omega}) = \frac{2\pi}{\theta_R - \theta_L} \frac{1}{Q} \sum_{q=1}^{Q} \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta_q)^2}} I_q(\tilde{\omega}),$$

$$= \frac{2\pi}{\theta_R - \theta_L} \int_{\theta_L}^{\theta_R} \frac{1}{\theta_R - \theta_L} \frac{1}{\sqrt{1 - (\cos \vartheta - \tilde{\omega})^2}} \int_{\theta_R}^{\theta_R + \cos \vartheta} I(\vartheta, \tilde{\omega}) \, d\vartheta, \quad \begin{cases} 2\pi \frac{1}{(\theta_R - \theta_L)^2} \int_{\arccos(\tilde{\omega} + \cos \vartheta)}^{\theta_R} \frac{1}{\sqrt{1 - (\cos \vartheta - \tilde{\omega})^2}} \, d\vartheta, & -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\ 2\pi \frac{1}{(\theta_R - \theta_L)^2} \int_{\theta_R}^{\theta_R + \cos \vartheta} \frac{1}{\sqrt{1 - (\cos \vartheta - \tilde{\omega})^2}} \, d\vartheta, & 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R) \end{cases}. \quad (19)$$

If the channel follows Jakes’ channel model, we have $\theta_L = 0^\circ, \theta_R = 180^\circ$ and $\mu(\theta_L, \theta_R) = 2$. Then, the above derived window function (19) can be further simplified as

$$W(\tilde{\omega}) = \frac{2}{\pi} \int_{\arccos(|\tilde{\omega}|-1)}^{\arccos(|\tilde{\omega}|-1)} \frac{1}{\sqrt{1 - (\cos \vartheta - |\tilde{\omega}|)^2}} \, d\vartheta, \quad |\tilde{\omega}| \leq 2. \quad (20)$$
In fact, for $0 < |\tilde{\omega}| < 2$, the window function (20) can be equivalently transformed into the following elliptic integral

$$W(\tilde{\omega}) = \frac{2}{\pi} \int_0^{\arccos(|\tilde{\omega}| - 1)} \frac{1}{\sqrt{1 - (\cos \vartheta - |\tilde{\omega}|)^2}} \ d\vartheta = \frac{2}{\pi} \int_{|\tilde{\omega}| - 1}^{1} \frac{1}{\sqrt{1 - x^2}} \sqrt{1 - (x - |\tilde{\omega}|)^2} \ dx,$$

which employs the property of equation (3.147-4) in [32], $\nu = \sqrt{1 - |\tilde{\omega}|^2}$ and $F(\psi, k)$ is the elliptic integral of the first kind defined as

$$F(\psi, k) = \int_0^{\psi} \frac{1}{\sqrt{1 - k^2 \sin^2 \alpha}} \ d\alpha.$$

Based on (21), we can obtain that as $|\tilde{\omega}|$ tends to 2, $\nu$ approaches to 0 and there holds

$$\lim_{|\tilde{\omega}| \to 2} W(\tilde{\omega}) = \frac{2}{\pi} \int_0^\pi \lim_{\nu \to 0} \frac{1}{\sqrt{1 - \nu^2 \sin^2 \xi}} \ d\xi = \frac{2}{\pi} \int_0^\pi 1 \ d\xi = 1. \ (22)$$

As for $\tilde{\omega} = 0$, there holds

$$W(0) = \frac{2}{\pi} \int_0^\pi \frac{1}{\sqrt{1 - \cos^2 \vartheta}} \ d\vartheta = \frac{2}{\pi} \int_0^\pi \frac{1}{\sin \vartheta} \ d\vartheta = \frac{4}{\pi} \int_0^\pi \frac{1}{\sin \vartheta} \ d\vartheta = \frac{4}{\pi} \int_0^\pi \frac{1}{\sin \frac{\vartheta}{2} \cos \frac{\vartheta}{2}} \ d\frac{\vartheta}{2} = \frac{4}{\pi} \int_0^\pi \frac{1}{\tan \vartheta} \ d\vartheta = \frac{4}{\pi} \int_0^\pi \frac{1}{\tan \vartheta} \ d\vartheta = \frac{4}{\pi} \ln |\tan \vartheta| |\vartheta| = +\infty. \ (23)$$
In summary, the window function given in \( (20) \) for Jakes’ channel with ‘Equi-angle’ beamforming direction distribution can be re-expressed as

\[
\mathcal{W}(\tilde{\omega}) = \begin{cases} 
\frac{2}{\pi} \int_0^{\pi/2} \frac{1}{\sqrt{1 - \tilde{\omega}^2 \sin^2 \xi}} \ d\xi, & 0 < |\tilde{\omega}| < 2 \\
\inf, & \tilde{\omega} = 0 \\
1, & |\tilde{\omega}| = 2.
\end{cases}
\]  

(24)

For comparison, different forms of window functions under different channel assumptions and beamforming directions are summarized in Table I. The window functions given in \( (16) \), \( (17) \) and \( (20) \) are depicted in Fig. 2. Note that \( (16) \) and \( (17) \) adopt ‘Equi-cos’ while \( (20) \) adopts ‘Equi-angle’. Besides, Jakes’ channel model is assumed for \( (17) \) and \( (20) \), whereas we take \( \theta_L = 0^\circ \), \( \theta_R = 90^\circ \) for \( (16) \). All the three window functions are nonnegative and decrease with increasing \( |\tilde{\omega}| \). Hence, they all attain the maximum at \( \tilde{\omega} = 0 \). Apart from this, the following observations can be made:

First, unlike \( (17) \) and \( (20) \), \( (16) \) yields a window function which is asymmetric about \( \tilde{\omega} = 0 \). Such asymmetry is due to the fact that the mean AoD \( \bar{\theta} = \pi/4 \) deviates from \( \pi/2 \). Moreover, \( \mathcal{W}(\tilde{\omega}) \) in \( (16) \) remains zero for \( |\tilde{\omega}| > 1 \), due to \( \mu(\theta_L, \theta_R) = 1 \). Second, comparing \( (17) \) and \( (20) \), we observe that the window function in \( (20) \) is more concentrated around \( \tilde{\omega} = 0 \), while that in \( (17) \) better attenuates the high Doppler frequencies as \( |\tilde{\omega}| \) approaches 2. Third, \( \mathcal{W}(\tilde{\omega}) \) in \( (20) \) is unbounded above at \( \tilde{\omega} = 0 \) and converges to 1 as \( |\tilde{\omega}| \) tends to 2, which matches with the former analysis in \( (22) \) and \( (23) \).

2) Impact of antenna array configuration: The choice of antenna spacing \( \Delta d_r \) is crucial, and we specifically focus on the ULA, i.e., \( \Delta d_r = (r-1)d \), to facilitate the discussion about the

| \( \mathcal{W}(\omega) \) | Non-Jakes’ channel | Jakes’ channel |
|----------------|-----------------|----------------|
| **Equi-cos** | \[ \begin{array}{l}
\frac{2\pi}{\theta_R - \theta_L} \arccos(\cos \theta_R - \tilde{\omega}), \quad -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\
\frac{2\pi}{\theta_R - \theta_L} \arccos(\cos \theta_L - \tilde{\omega}), \quad 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R)
\end{array} \] | \[ \begin{array}{l}
\arccos(|\tilde{\omega}| - 1), \quad |\tilde{\omega}| \leq 2
\end{array} \] |

**Equi-angle** | \[ \begin{array}{l}
\frac{2\pi}{\theta_R - \theta_L} \int_{\theta_L}^{\theta_R} \frac{1}{\sqrt{1 - (\sin \xi)^2}} d\xi, \quad -\mu(\theta_L, \theta_R) \leq \tilde{\omega} < 0 \\
\frac{2\pi}{\theta_R - \theta_L} \int_{\theta_L}^{\theta_R} \frac{1}{\sqrt{1 - (\cos \xi)^2}} d\xi, \quad 0 \leq \tilde{\omega} \leq \mu(\theta_L, \theta_R)
\end{array} \] | \[ \begin{array}{l}
\frac{2\pi}{\theta_R - \theta_L} \int_0^{\pi/2} \frac{1}{\sqrt{1 - (\tilde{\omega}^2 \sin^2 \xi)}} d\xi, \quad 0 < |\tilde{\omega}| < 2 \\
\inf, \quad \tilde{\omega} = 0 \\
1, \quad |\tilde{\omega}| = 2
\end{array} \] |
impact of antenna spacing \(d\) on the PSD and Doppler spread. In such case, we have
\[
|G(\tilde{\omega})|^2 = \left| \frac{1}{M} \sum_{r=1}^{M} e^{-j2\pi(r-1)\frac{d}{\lambda}} \right|^2 = \frac{\sin^2(\chi M \tilde{\omega})}{M^2 \sin^2(\chi \tilde{\omega})},
\]
where \(\chi = \pi \frac{d}{\lambda}\). It is seen that the beam function \(|G(\tilde{\omega})|^2\) turns into a periodic function of \(\tilde{\omega}\), which repeats itself with period \(\tilde{\Omega} = \frac{\pi}{\chi}\).

Note that the antenna spacing \(d\) can be set a bit larger to gain higher beamforming resolution, but it cannot exceed \(d_{\text{max}} = \frac{\lambda}{2}\) to avoid aliasing. Moreover, \(d = \frac{\lambda}{2}\) will also incur the aliasing between \(0^\circ\) and \(180^\circ\). Therefore, we limit the range of antenna spacing as \(0 < d < \frac{\lambda}{2}\), and the optimal antenna spacing should be compromised between beamforming resolution and aliasing avoidance.

Fig. 3 compares the beam function \(|G(\omega)|^2\), window function \(W(\omega)\) and PSD \(P(\omega)\), when the antenna spacings are taken as \(d = 0.3\lambda\) and \(d = 0.5\lambda\), respectively. Note that the absolute values of \(|G(\omega)|^2\), \(W(\omega)\) and \(P(\omega)\) have been scaled such that their maximums are all 1 (e.g., the depicted window function is in fact \(W(\omega)/\max\{|W(\omega)|\}\)). The maximum Doppler shift is taken as \(f_d = 1000\text{ Hz}\). The signal AoDs are assumed to follow uniform distribution between \((0, \pi)\) (namely Jakes’ model) and the beamforming directions are configured such that \(\cos \vartheta_q, q = 1, 2, \cdots, Q\) are uniformly distributed between \((-1, 1)\). Since the signal AoD region and the beamforming directions are exactly the same, both the two cases share the same window function \(W(\omega)\). Hence, only the choice of antenna spacing accounts for the difference between Fig. 3(a) and Fig. 3(b). As anticipated, the beam function \(|G(\omega)|^2\) accomplishes a full period within \(\omega \in (-2\omega_d, 2\omega_d)\),
which implies $|G(\pm 2\omega_d)|^2 = |G(0)|^2$. Therefore, despite the attenuation effect of the window function $W(\omega)$, the PSD at $\frac{d}{\lambda} = 0.5$ would be much larger than that at $\frac{d}{\lambda} = 0.3$, for large $\omega$. Since a PSD concentrated around low Doppler frequencies is more favorable for reducing the time variation of the equivalent channel, the antenna spacing $d = 0.5\lambda$ should be avoided. This graphically explains from another perspective why the tradeoff between beamforming resolution and aliasing avoidance needs to be taken when determining the antenna spacing $d$.

IV. Antenna Weighting Technique for Reducing Doppler Spread

According to the above discussion, the channel PSD is determined by both the beam function $|G(\tilde{\omega})|^2$ and window function $W(\tilde{\omega})$. Hence, it is possible to reduce the Doppler spread by equivalently modifying the beam function $|G(\tilde{\omega})|^2$, which in fact corresponds to the radiation pattern of the array. We propose to minimize the Doppler spread through antenna weighting, as illustrated in Fig. 4. Specifically, the $r$th antenna is weighted by a complex weight $u_r$ after multi-branch transmit beamforming and the weights $u = [u_1, u_2, \ldots, u_M]^T$ could be optimized to minimize the Doppler spread, which is equivalent to reduce the residual time variation of the channel.

The array response vector $a(\theta)$ remains the same as in the previous section, and the MF beamformers are taken as $b_{AW}(\theta_q) = \eta_{AW} \sqrt{Q} a(\theta_q) e^{j\phi(\theta_q)}$, with $\eta_{AW} = \frac{1}{\sqrt{\sum_{q=1}^{Q} 1 \diag(u) a(\theta_q) e^{j\phi(\theta_q)}}}$ being the normalization coefficient to keep the total transmit power per symbol as 1. Then, similar to (3), the signal received at the BS after Doppler shifts compensation, multi-branch
transmit beamforming and antenna weighting can be re-expressed as (the noise item is ignored and only one channel tap is considered)

\[
\mathbf{r}_{m, AW} = \sum_{q=1}^{Q} \int_{\theta_l}^{\theta_R} \kappa(\theta) \mathbf{a}^T(\theta) \text{diag}(\mathbf{u}) \mathbf{b}_{AW}^*(\vartheta_q) \mathbf{s}_m(d_1) \mathbf{\Psi}_{m,q}(d_1) \mathbf{\Phi}_m(\theta)d\theta,
\]

\[
= \frac{\eta_{AW}}{\sqrt{Q}} \sum_{q=1}^{Q} \int_{\theta_l}^{\theta_R} \kappa(\theta)\exp(-j\vartheta_q) \frac{1}{M} \mathbf{a}^H(\vartheta_q) \text{diag}(\mathbf{u}) \mathbf{a}(\theta) \mathbf{s}_m \mathbf{\Psi}_{m,q}(d_1) \mathbf{\Phi}_m(\theta)d\theta.
\]  

(25)

By ignoring the real scalar \(\eta_{AW}\) which does not affect the PSD analysis, the equivalent uplink channel of (25) can be expressed in continuous-time form as

\[
g_{AW}(t) = \frac{1}{\sqrt{Q}} \sum_{q=1}^{Q} \int_{\theta_l}^{\theta_R} \alpha(\theta) G_{AW}(\cos \theta, \cos \vartheta_q) e^{j\omega_d(\cos \theta - \cos \vartheta_q)t + j\vartheta(\theta) - j\vartheta_q} d\theta,
\]

(26)

where

\[
G_{AW}(\cos \theta, \cos \vartheta_q) = \frac{1}{M} \mathbf{a}^H(\vartheta_q) \text{diag}(\mathbf{u}) \mathbf{a}(\theta) = \frac{1}{M} \sum_{r=1}^{M} u_r e^{j2\pi \frac{\Delta f_r}{\lambda}(\cos \theta - \cos \vartheta_q)}.
\]

(27)

By denoting \(\mathbf{c}(\cos \theta, \cos \vartheta_q) = [1, e^{j2\pi \frac{\Delta f_1}{\lambda}(\cos \theta - \cos \vartheta_q)}, \ldots, e^{j2\pi \frac{\Delta f_M}{\lambda}(\cos \theta - \cos \vartheta_q)}]^T\), (27) could be rewritten as

\[
G_{AW}(\cos \theta, \cos \vartheta_q) = \frac{1}{M} \mathbf{c}^T(\cos \theta, \cos \vartheta_q) \mathbf{u}.
\]

(28)

Note that the only difference between the continuous-form channels in (26) and (4) is that \(G(\cos \theta, \cos \vartheta_q)\) in (4) is replaced by \(G_{AW}(\cos \theta, \cos \vartheta_q)\) in (27). Actually, by letting \(\mathbf{u} = 1_{M \times 1}\), \(G_{AW}(\cos \theta, \cos \vartheta_q)\) will reduce to \(G(\cos \theta, \cos \vartheta_q)\). Thus, the considered scenario in Section III can be categorized as a special case of equal antenna weighting. Following the similar approach as in the previous section, the channel PSD can be expressed as

\[
P_{AW}(\omega) = \frac{1}{\omega_d}|G_{AW}(\tilde{\omega})|^2W(\tilde{\omega}),
\]

(29)

where the window function remains exactly the same as (8), while the beam function can be redefined as

\[
|G_{AW}(\tilde{\omega})|^2 = \left|\frac{1}{M} \sum_{r=1}^{M} u_r e^{-j2\pi \frac{\Delta f_r}{\lambda} \tilde{\omega}}\right|^2 = \frac{1}{M} \mathbf{\varsigma}^T(\tilde{\omega}) \mathbf{u}^2.
\]

(30)

Here, \(\mathbf{\varsigma}(\tilde{\omega}) = [1, e^{-j2\pi \frac{\Delta f_1}{\lambda} \tilde{\omega}}, \ldots, e^{-j2\pi \frac{\Delta f_M}{\lambda} \tilde{\omega}}]^T\).

As a result, the Doppler spread with antenna weighting can be calculated as

\[
\sigma_{DS, AW} = \sqrt{\int_{-\mu(\Theta_1, \theta_R)\omega_d}^{\mu(\Theta_1, \theta_R)\omega_d} \omega^2 P_{AW}(\omega) d\omega} = \frac{\omega_d}{\omega_d} \sqrt{\int_{-\mu(\Theta_1, \theta_R)\omega_d}^{\mu(\Theta_1, \theta_R)\omega_d} |\mathbf{\varsigma}^T(\tilde{\omega}) \mathbf{u}|^2 W(\tilde{\omega}) d\tilde{\omega}},
\]

\[
\sigma_{DS, AW} = \frac{\omega_d}{\omega_d^2} \int_{-\mu(\Theta_1, \theta_R)\omega_d}^{\mu(\Theta_1, \theta_R)\omega_d} \omega^2 P_{AW}(\omega) d\omega = \omega_d \int_{-\mu(\Theta_1, \theta_R)\omega_d}^{\mu(\Theta_1, \theta_R)\omega_d} |\mathbf{\varsigma}^T(\tilde{\omega}) \mathbf{u}|^2 W(\tilde{\omega}) d\tilde{\omega},
\]

(31)
\[ \omega_d \left( \frac{u^H \left[ \int_{-\mu(\theta_L,\theta_R)}^{\mu(\theta_L,\theta_R)} \tilde{\omega}^2 W(\tilde{\omega}) \zeta^* (\tilde{\omega}) \zeta^T (\tilde{\omega}) d\tilde{\omega} \right] u}{u^H \left[ \int_{-\mu(\theta_L,\theta_R)}^{\mu(\theta_L,\theta_R)} W(\tilde{\omega}) \zeta^* (\tilde{\omega}) \zeta^T (\tilde{\omega}) d\tilde{\omega} \right] u} \right) = \omega_d \sqrt{\frac{u^H C_2 u}{u^H C_0 u}}, \] \tag{31}

where

\[ C_2 = \int_{-\mu(\theta_L,\theta_R)}^{\mu(\theta_L,\theta_R)} \tilde{\omega}^2 W(\tilde{\omega}) \zeta^* (\tilde{\omega}) \zeta^T (\tilde{\omega}) d\tilde{\omega}, \]

\[ C_0 = \int_{-\mu(\theta_L,\theta_R)}^{\mu(\theta_L,\theta_R)} W(\tilde{\omega}) \zeta^* (\tilde{\omega}) \zeta^T (\tilde{\omega}) d\tilde{\omega}. \]

Note that both \( C_0 \) and \( C_2 \) are real symmetric Toeplitz matrix.

The optimal antenna weights \( \hat{u} \) minimizing the Doppler spread can be acquired by solving the following optimization problem

\[ \hat{u} = \arg \min_u \sigma_{DS,\text{AW}} = \arg \min_u \omega_d^2 \frac{\hat{u}^H C_2 \hat{u}}{\hat{u}^H C_0 \hat{u}} = \arg \min_u \omega_d^2 \frac{\hat{u}^H C_2 \hat{u}}{\hat{u}^H C_0 \hat{u}}, \]

\[ s.t. \; \hat{u}^H C_0 \hat{u} = 1, \] \tag{32}

where \( \hat{u} \) denotes the trial antenna weights, and the constraint \( \hat{u}^H C_0 \hat{u} = 1 \) is added to eliminate the magnitude ambiguity of \( \hat{u} \) and also to avoid the degenerate solution of \( \hat{u} = 0 \). Let \( \alpha \) be an arbitrary real scalar. It is noted that \( \hat{u} \) and \( \alpha \hat{u} \) would yield the same Doppler spread.

Decomposing \( C_0 \) as \( C_0 = QQ^H \) and defining \( \tilde{\gamma} = Q^H \hat{u} \), we obtain \( \tilde{u} = Q^{-H} \tilde{\gamma} \) and \( \tilde{\gamma}^H \tilde{\gamma} = 1 \). Therefore, the optimization problem \( \text{(32)} \) can be transformed into

\[ \tilde{\gamma} = \arg \min_{\tilde{\gamma}} \tilde{\gamma}^H Q^{-1} C_2 Q^{-H} \tilde{\gamma}, \]

\[ s.t. \; \tilde{\gamma}^H \tilde{\gamma} = 1. \] \tag{33}

The minimization problem \( \text{(33)} \) can be readily solved with Lagrange multiplier method. Assuming \( \tilde{\mu} \) as the Lagrange multiplier, we arrive at the following Lagrange objective function

\[ J(\tilde{\gamma}, \tilde{\mu}) = \tilde{\gamma}^H Q^{-1} C_2 Q^{-H} \tilde{\gamma} + \tilde{\mu} \left( 1 - \tilde{\gamma}^H \tilde{\gamma} \right). \] \tag{34}

The first-order condition of \( J(\tilde{\gamma}, \tilde{\mu}) \) leads to \( \frac{\partial J(\tilde{\gamma}, \tilde{\mu})}{\partial \tilde{\gamma}^r} = Q^{-1} C_2 Q^{-H} \tilde{\gamma} - \tilde{\mu} \tilde{\gamma} = 0 \), i.e., \( Q^{-1} C_2 Q^{-H} \tilde{\gamma} = \tilde{\mu} \tilde{\gamma} \). Combining this equation with the constraint \( \tilde{\gamma}^H \tilde{\gamma} = 1 \), we know that the candidate solutions would be \( \tilde{\gamma}^r = v_r \left( Q^{-1} C_2 Q^{-H} \right), \tilde{\mu}^r = \lambda_r \left( Q^{-1} C_2 Q^{-H} \right), r = 1, 2, \ldots, M \). Considering that \( J(\tilde{\gamma}^r, \tilde{\mu}^r) = \tilde{\mu}^r \), the optimal antenna weights minimizing the Doppler spread can be obtained
Fig. 5: Comparison of the normalized beam functions $|\mathcal{G}(\omega)|^2$ (equal antenna weighting) and $|\mathcal{G}_{AW}(\omega)|^2$ (optimal antenna weighting) for $M=16$-element ULA, with antenna spacing $d = 0.45\lambda$ and ‘Equi-cos’ beamforming directions.

in closed-form as

$$\hat{\gamma} = v_{\min}\left(Q^{-1}C_2Q^{-H}\right), \quad \hat{u} = Q^{-H}v_{\min}\left(Q^{-1}C_2Q^{-H}\right). \quad (35)$$

Here, $v_{\min}(X)$ denotes the eigenvector corresponding to the minimum eigenvalue of matrix $X$.

Remark 1: It can be seen from (32) that the optimization of antenna weights does not depend on the maximum Doppler shift $f_d$ and thus is independent of the velocity of the HST. Besides, the optimal antenna weights (35) can be expressed in closed-form as a function of $C_0$ and $C_2$. Thus, for a given antenna array structure, the optimal antenna weights $\hat{u}$ are uniquely determined by the window function $\mathcal{W}(\bar{\omega})$, which depends on the AoD distribution. In other words, as long as the AoD distribution is obtained, the antenna weights can be optimized, and the obtained $\hat{u}$ remains valid irrespective of the HST velocity.

Since the optimal antenna weights $\hat{u}$ reduces the Doppler spread through changing the beam function, we compare in Fig. 5 the beam function $|\mathcal{G}_{AW}(\omega)|^2$ obtained after antenna weighting with $\hat{u}$ and $|\mathcal{G}(\omega)|^2$ with equal antenna weighting. The 16-element ULA with antenna spacing $d = 0.45\lambda$ is adopted, and the maximum Doppler shift is $f_d = 1000$ Hz. Note that the absolute
values of the beam functions have been scaled such that their maximums are all 1, as in Fig. 3. For each beam function, we define the ratio between the sidelobe levels and the maximum gain as side-to-main ratio (SMR) $\rho$. It can be seen that the average SMR $\rho$ is about $10^{-2}$ for $|G(\omega)|^2$ with equal antenna weighting, while the beam function $|G_{AW}(\omega)|^2$ with optimal antenna weighting yields an average SMR of $\rho_{AW} \approx 10^{-4}$, two orders of magnitude smaller than the former. Such low SMR is obtained at the cost of a slightly wider mainlobe. Nevertheless, the SMR has greater impact on the Doppler spread than the mainlobe width. Hence, the Doppler spread could be significantly reduced through the proposed optimal antenna weighting technique, which substantially attenuates the high Doppler frequencies.

V. SIMULATION RESULTS

In this section, we will first verify the accuracy of the PSD analysis and investigate the impact of some parameters on PSD through numerical examples, and then demonstrate the superiority of the optimal antenna weighting technique over equal antenna weighting case. Unless otherwise stated, the antenna spacing is taken as $\frac{d}{\lambda} = 0.45$, the maximum Doppler shift is set as $f_d = 1000$ Hz, the ULA consists of $M = 16$ antennas, and the Jakes’ channel model is adopted, i.e., the signal AoDs follow uniform distribution between $(0, \pi)$.

A. Verification of the PSD Analysis

![Fig. 6: Comparison of the PSD $P(\omega)$ between two circumstances of (a) $\theta_L = 0^\circ$, $\theta_R = 120^\circ$ with ‘Equi-cos’ beamforming directions and (b) Jakes’ channel model with ‘Equi-angle’ beamforming directions.](image-url)
In Fig. 6, we compare the channel PSD under different signal AoD regions and beamforming directions. The AoDs are constrained within $(\theta_L, \theta_R)$ with $\theta_L = 0^\circ$, $\theta_R = 120^\circ$ and the beamforming directions $\vartheta_q$ are configured such that $\cos \vartheta_q$ are evenly distributed between $(\cos \theta_R, \cos \theta_L)$ in Fig. 6(a), while the signal AoDs follow the Jakes’ channel model and the beamforming directions $\vartheta_q$ are uniformly chosen from $(0, \pi)$ in Fig. 6(b). That is to say, (16) and (20) should be employed to compute the window function, respectively.

In order to verify the correctness of the PSD derivation (6), we provide the numerical PSD obtained in the following way: We first calculate the channel autocorrelation $R_g(\tau)$ at $N$ discrete time points by averaging over sufficient number of channel realizations (4) and then apply an $N$-point discrete Fourier transform (DFT) to obtain the discretized PSD. Note that $N$ should be accordingly increased with the number of antennas $M$ to capture the faster fluctuation of the magnitude of the PSD. Fig. 6 reveals that whether the channel follows Jakes’ channel model or not, the analyzed PSD (6) perfectly coincides with its numerical counterpart, confirming the validity of the PSD analysis. Furthermore, we can find that the PSD in Fig. 6(a) is asymmetric about $\omega = 0$ while that in Fig. 6(b) is symmetric. In fact, we have pointed out in Fig. 2 that an average AoD $\bar{\theta}$ different from $\frac{\pi}{2}$ will result in asymmetric window function, which accounts for the asymmetry of the PSD in Fig. 6(a).

Next, we evaluate the impact of the number of antennas on the PSD in Fig. 7. The ULA with $M = 4, 16, 64$ antennas are considered. The Jakes’ channel model and ‘Equi-cos’ beamforming directions are adopted for all cases such that the window function remains the same. Therefore, the exclusive contributing factor to the difference of the PSDs is the beam function $|G(\omega)|^2$, which in fact corresponds to the radiation pattern obtained with the MF beamformer pointing to the normal direction of ULA, as mentioned earlier. When the number of antennas $M$ increases, the radiation pattern exhibits lower sidelobe levels and narrower main and side lobes. These features are all reflected by the beam function and thus by the PSDs depicted in Fig. 7. Since the sidelobes of the PSD cover the undesired high Doppler frequencies, a larger number of antennas can better reduce the sidelobe levels and would lead to smaller Doppler spread.
Fig. 7: Comparison of the PSD $P(\omega)$ under ULA with different numbers of antennas $M = 4, 16, 64$, the antenna spacing is $d = 0.45\lambda$ and the configuration of beamforming directions satisfies ‘Equi-cos’.

Fig. 8: Comparison of the Doppler spread $\sigma_{DS}$ calculated by (10), when the beamforming directions are configured in two different ways, with $M = 16, 64, 256$ and $\frac{d}{\lambda} = [0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.47, 0.49, 0.495, 0.5]$.

Then, the Doppler spreads computed by (10) are compared in Fig. 8 under a set of normalized antenna spacings $\frac{d}{\lambda} = [0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.47, 0.49, 0.495, 0.5]$, for different numbers of antennas $M = 16, 64, 256$ and different configurations of beamforming directions.
‘Equi-cos’ (even \( \cos \theta_q \)) and ‘Equi-angle’ (even \( \theta_q \)). The maximum Doppler shift is set as \( f_d = 5000 \) Hz. The following observations can be made from Fig. 8:

1) The Doppler spread decreases with the increasing number of antennas \( M \), which is within expectation. Actually, an enlarged antenna array provides higher spatial resolution and thereby the residual Doppler shifts tend to vanish, and the time variation of the equivalent channel after Doppler shifts compensation and transmit beamforming can be significantly alleviated.

2) There exists an optimal antenna spacing \( d_{\text{opt}} \) which yields the minimal Doppler spread, and an antenna spacing \( d \) either smaller or larger than \( d_{\text{opt}} \) would be detrimental to appeasing the residual time variation of the equivalent channel. A too small \( d \) cannot fully exploit the spatial resolution of the ULA, which is unfavorable for reducing the residual Doppler shifts. And as \( d \) increases to 0.5, the aliasing between 0° and 180° would considerably enhance the PSD at high Doppler frequencies. Both factors contribute to large Doppler spread.

3) The Doppler spread is not sensitive to how the beamforming directions are configured for \( \frac{4}{\lambda} \leq 0.45 \). However, different configurations of beamforming directions ‘Equi-cos’ and ‘Equi-angle’ have great impact on the Doppler spread for \( \frac{4}{\lambda} = 0.5 \). The significant divergence between the Doppler spread of ‘Equi-cos’ and that of ‘Equi-angle’ for \( \frac{4}{\lambda} = 0.5 \) can be explained as follows. As previously mentioned, the beam function \( |G(\omega)|^2 \) accomplishes a full period within \((-2\omega_d, 2\omega_d)\) under \( \frac{4}{\lambda} = 0.5 \), which implies \( G(\pm 2\omega_d) = G(0) \). Considering that the window function of ‘Equi-cos’ approaches 0 while that of ‘Equi-angle’ approaches 1 as \( \omega \) gets closer to \( \pm 2\omega_d \), the PSD of ‘Equi-cos’ at undesired high Doppler frequencies would be much smaller than ‘Equi-angle’. Thus, the former better attenuates the time variation of the equivalent channel, resulting in smaller Doppler spread.

B. Superiority of the Optimal Antenna Weighting Technique

In this subsection, we will demonstrate numerically the superiority of the proposed optimal antenna weighting technique, in terms of Doppler spread and uncoded symbol error rate (SER).

First, we assess in Fig. 9 the effect of the proposed antenna weighting technique on reducing the Doppler spread, under ULA with different numbers of antennas \( M = 8, 16, 32, 64, 128 \). The
maximum Doppler shift is set as $f_d = 5000$ Hz, and the beamforming directions satisfy ‘Equi-cos’ configuration. The acquired optimal antenna weights for different numbers of antennas $M = 8, 16, 32, 64$ are shown in Table II. Note that the maximum weight is normalized to 1. It can be seen from Fig. 9 that in contrast to the case with equal antenna weighting, the proposed antenna weighting technique with the optimal weights can substantially reduce the Doppler spread and thus suppress the residual time variation of the equivalent channel, regardless of the number of antennas. The reduction of the Doppler spread originates from the attenuation of the high Doppler frequencies, since the beam function $|G_{AW}(\omega)|^2$ obtained with the optimal antenna weights has much lower average SMR, as demonstrated in Fig. 5.

After Doppler shifts compensation, multi-branch transmit beamforming and equal or optimal antenna weighting, the received signal only suffers from a slight residual time variation. Thus, the conventional channel estimation and data detection for time-invariant channel can be directly performed. Fig. 10 compares the SER performance obtained with the received signals after equal and optimal antenna weighting, respectively. The receiver employs a 4-element ULA with antenna spacing $\frac{d}{\lambda} = 0.5$ and the maximum-ratio-combining (MRC) is used to detect the data symbols.
TABLE II: Antenna weights obtained for different numbers of antennas $M = 8, 16, 32, 64$

| $M$ | Normalized antenna weights |
|-----|-----------------------------|
| 8   | 0.384, 0.656, 0.876, 1.000, 1.000, 0.876, 0.656, 0.384 |
| 16  | 0.106, 0.221, 0.364, 0.525, 0.687, 0.832, 0.941, 1.000, 1.000, 0.941, 0.832, 0.687, 0.525, 0.364, 0.221, 0.106 |
| 32  | 0.060, 0.125, 0.207, 0.300, 0.399, 0.497, 0.591, 0.675, 0.748, 0.810, 0.863, 0.907, 0.943, 0.971, 0.990, 1.000, 1.000, 0.990, 0.943, 0.907, 0.863, 0.810, 0.748, 0.675, 0.591, 0.497, 0.399, 0.300, 0.207, 0.125, 0.060 |
| 64  | 0.030, 0.063, 0.104, 0.153, 0.206, 0.261, 0.314, 0.364, 0.410, 0.454, 0.494, 0.534, 0.573, 0.613, 0.652, 0.691, 0.727, 0.761, 0.792, 0.821, 0.847, 0.871, 0.893, 0.914, 0.934, 0.952, 0.967, 0.979, 0.988, 0.994, 0.998, 1.000, 1.000, 0.998, 0.994, 0.988, 0.979, 0.967, 0.952, 0.934, 0.914, 0.893, 0.871, 0.847, 0.821, 0.792, 0.761, 0.727, 0.691, 0.652, 0.613, 0.573, 0.534, 0.494, 0.454, 0.410, 0.364, 0.314, 0.261, 0.206, 0.153, 0.104, 0.063, 0.030 |

Fig. 10: Comparison of the SER obtained with optimal and equal antenna weighting, with 4-element receive ULA and ULA composed of $M = 32, 64, 128$ antennas at transmitter.

The transmitter is equipped with a large-scale ULA with $\frac{d}{\lambda} = 0.45$, and $M = 32, 64, 128$ transmit antennas are considered. Each OFDM frame consists of 5 blocks, with the first block serving as pilot block. The number of subcarriers is taken as $N = 128$, and both pilot and data symbols are randomly drawn from 16-QAM constellation. The maximum Doppler shift is set as $f_d = 1000$ Hz and the block duration is assumed to be $T_b = 0.1$ ms, which implies that the normalized Doppler shift is $f_dT_b = 0.1$. Moreover, the beamforming directions satisfy ‘Equi-cos’ configuration.

From Fig. 10, the superiority of the proposed optimal antenna weighting technique is evident. Even with $M = 128$ transmit antennas, the equal antenna weighting scheme suffers from severe
SER performance floor, which can be attributed to the residual time variation of the channel caused by the uncompensated Doppler shifts. In fact, the numerical results in [23] reveal that only when the transmit antennas are increased to $M = 1024$, would the residual time variation become negligible and the SER performance floor disappear. In contrast, the SER performance with optimal antenna weighting does not exhibit obvious floor even with $M = 64$ and 128 transmit antennas. This is due to the fact that the optimal antenna weighting technique can substantially reduce the Doppler spread, compared to equal antenna weighting. The reduction of the Doppler spread is reflected by the significantly improved SER performance. In other words, with the optimal antenna weighting technique, far fewer transmit antennas are required to attain the same detection performance as equal antenna weighting scheme.

VI. CONCLUSIONS

In this paper, we considered the angle-domain Doppler shifts compensation scheme for high-mobility uplink communication and derived the exact PSD and Doppler spread as a measure to assess the residual time variation of the equivalent channel. The analysis reveals that the PSD can be fully characterized by the beam function and window function, which depends on the antenna array configuration and the beamforming directions, respectively. Based on the delicately derived PSD with explicit expression, the impact of some essential parameters including the antenna spacing and beamforming directions on the channel PSD was discussed. Inspired by the PSD analysis, an antenna weighting technique was further proposed to reduce the Doppler spread through equivalently modifying the beam function. Numerical results were provided to corroborate the PSD analysis and the antenna weighting technique.

APPENDIX A

PROOF OF Lemma

According to the definition, the channel PSD can be expressed as

$$P(\omega) = \int_{-\infty}^{+\infty} R_g(\tau) e^{-j\omega \tau} d\tau,$$
\begin{align*}
&= \frac{1}{(\theta_R - \theta_L)Q} \sum_{q=1}^{Q} \int_{\theta_l}^{\theta_R} |G(\cos \theta, \cos \vartheta_q)|^2 \left[ \int_{-\infty}^{\infty} e^{-j\omega_d(\cos \theta - \cos \vartheta_q)\tau} e^{-j\omega \tau} d\tau \right] d\theta,
&= \frac{2\pi}{(\theta_R - \theta_L)Q} \sum_{q=1}^{Q} \int_{\theta_l}^{\theta_R} |G(\cos \theta, \cos \vartheta_q)|^2 \delta(\omega + \omega_d(\cos \theta - \cos \vartheta_q)) d\theta, \quad (36)
\end{align*}

where we have exploited \( \int_{-\infty}^{\infty} e^{-j\omega_d(\cos \theta - \cos \vartheta_q)\tau} e^{-j\omega \tau} d\tau = 2\pi \delta(\omega + \omega_d(\cos \theta - \cos \vartheta_q)) \).

Besides, there holds
\[
\int_{\theta_L}^{\theta_R} |G(\cos \theta, \cos \vartheta_q)|^2 \delta(\omega + \omega_d(\cos \theta - \cos \vartheta_q)) d\theta,
\]
\[
= \frac{1}{\omega_d} \int_{\omega_d \cos \theta_L}^{\omega_d \cos \theta_R} \left| G \left( \frac{y}{\omega_d}, \cos \vartheta_q \right) \right|^2 \frac{1}{\sqrt{1 - \left( \frac{y}{\omega_d} \right)^2}} \delta(y + \omega - \omega_d \cos \vartheta_q) dy,
\]
\[
= \begin{cases} 
\frac{1}{\omega_d} |G(\vartheta_q)|^2 \frac{1}{\sqrt{1 - (\omega - \cos \vartheta_q)^2}}, & -\cos \theta_L \leq \omega - \cos \vartheta_q \leq -\cos \theta_R \\
0, & \text{otherwise} 
\end{cases}
\]
\[
= \begin{cases} 
\frac{1}{\omega_d} |G(\tilde{\omega})|^2 \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta_q)^2}}, & -\cos \theta_L \leq \tilde{\omega} - \cos \vartheta_q \leq -\cos \theta_R \\
0, & \text{otherwise} 
\end{cases} \quad (37)
\]

Combining (36) and (37), we obtain
\[
P(\omega) = \frac{2\pi}{(\theta_R - \theta_L)Q} \sum_{q=1}^{Q} \frac{1}{\omega_d} |G(\tilde{\omega})|^2 \frac{1}{\sqrt{1 - (\tilde{\omega} - \cos \vartheta_q)^2}} T_q(\tilde{\omega}) = \frac{1}{\omega_d} |G(\tilde{\omega})|^2 W(\tilde{\omega}). \quad (38)
\]

Here, \( T_q(\tilde{\omega}) = \begin{cases} 
1, & q \in S(\tilde{\omega}) \\
0, & q \notin S(\tilde{\omega}) 
\end{cases} \), where \( S(\tilde{\omega}) \) is the set of beamforming branches contributing to the PSD at \( \tilde{\omega} \). From the derivation (37), \( S(\tilde{\omega}) \) can be given by \( S(\tilde{\omega}) = \{ q \mid \tilde{\omega} - \cos \theta_R \leq \cos \vartheta_q \leq \tilde{\omega} + \cos \theta_L \} \). However, there is also an implicit constraint about \( \vartheta_q \), i.e., \( \cos \theta_R \leq \cos \vartheta_q \leq \cos \theta_L \).

By making the implicit constraint explicit, \( S(\tilde{\omega}) \) can be re-expressed as (9).

This completes the proof.

**REFERENCES**

[1] J. Wu, and P. Fan, “A survey on high mobility wireless communications: Challenges, opportunities and solutions,” *IEEE Access*, vol. 4, no. 27, pp. 450-476, Jan. 2016.

[2] R. He, B. Ai, G. Wang, K. Guan, Z. Zhong, A. F. Molisch, C. Briso-Rodriguez, and C. P. Oestges, “High-speed railway communications: From GSM-R to LTE-R,” *IEEE Veh. Technol. Mag.*, vol. 11, no. 3, pp. 49-58, Sep. 2016.
[3] F. Hlawatsch, and G. Matz, *Wireless communications over rapidly time-varying channels*. Academic Press, 2011.

[4] D. Chizhik, “Slowing the time-fluctuating MIMO channel by beam forming,” *IEEE Trans. Wirel. Commun.*, vol. 3, no. 5, pp. 1554-1565, Sep. 2004.

[5] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, “OFDM and its wireless applications: A survey,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673-1694, May 2009.

[6] G. B. Giannakis and C. Tepedelenlioglu, “Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels,” *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969-1986, Oct. 1998.

[7] M. F. Rabbi, S. W. Hou, and C. C. Ko, “High mobility orthogonal frequency division multiple access channel estimation using basis expansion model,” *IET Commun.*, vol. 4, no. 3, pp. 353-367, Feb. 2010.

[8] F. Qu and L. Yang, “On the estimation of doubly-selective fading channels,” *IEEE Trans. Wirel. Commun.*, vol. 9, no. 4, pp. 1261-1265, Apr. 2010.

[9] X. Wang, G. Wang, R. Fan, and B. Ai, “Channel estimation with expectation maximization and historical information based expansion model for wireless communication systems on high speed railways,” *IEEE Access*, vol. 6, no. 9, pp. 72-80, Aug. 2018.

[10] J. Zhao, H. Xie, F. Gao, W. Jia, S. Jin, and H. Lin, “Time varying channel tracking with spatial and temporal BEM for massive MIMO systems,” 2018. [Online]. Available: [http://arxiv.org/abs/1802.10461](http://arxiv.org/abs/1802.10461)

[11] M. Souden, S. Affes, J. Benesty, and R. Bahroun, “Robust Doppler spread estimation in the presence of a residual carrier frequency offset,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4148-4153, Oct. 2009.

[12] F. Bellili and S. Affes, “A low-cost and robust maximum likelihood Doppler spread estimator,” in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 4325-4330.

[13] Y. R. Tsai and K. J. Yang, “Approximate ML Doppler spread estimation over flat Rayleigh fading channels,” *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 1007-1010, Nov. 2009.

[14] Y. Zhang, Q. Yin, P. Mu, and L. Bai, “Multiple Doppler shifts compensation and ICI elimination by beamforming in high-mobility OFDM systems,” in *Proc. Int. ICST Conf. Commun. and Netw. in China*, Aug. 2011, pp. 170-175.

[15] W. Guo, P. Mu, Q. Yin, and H. M. Wang, “Multiple Doppler frequency offsets compensation technique for high-mobility OFDM uplink,” in *Proc. IEEE ICSPCC*, Aug. 2013, pp. 1-5.

[16] T. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wirel. Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[17] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, “Scaling up MIMO: opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

[18] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186-195, Feb. 2014.

[19] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742-758, Oct. 2014.

[20] W. Zhang, F. Gao, S. Jin, and H. Lin, “Frequency synchronization for uplink massive MIMO systems,” *IEEE Trans. Wirel. Commun.*, vol. 17, no. 1, pp. 235-249, Jan. 2018.

[21] W. Guo, W. Zhang, P. Mu, and F. Gao, “High-mobility OFDM downlink transmission with large-scale antenna array,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8600-8604, Sep. 2017.
[22] Y. Ge, W. Zhang, and F. Gao, “High-mobility OFDM downlink transmission with partly calibrated subarray-based massive uniform linear array,” in *Proc. IEEE VTC-Spring*, Jun. 2017, pp. 1-6.

[23] W. Guo, W. Zhang, P. Mu, F. Gao, and B. Yao, “Angle-domain Doppler pre-compensation for high-mobility OFDM uplink with massive ULA,” in *Proc. IEEE GLOBECOM*, Dec. 2017, pp. 1-6.

[24] W. C. Jakes, and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE Press, 1994.

[25] Y. R. Zheng, and C. Xiao, “Simulation models with correct statistical properties for Rayleigh fading channels,” *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920-928, Jun. 2003.

[26] K. Liu, V. Raghavan, and A. M. Sayeed, “Capacity scaling and spectral efficiency in wide-band correlated MIMO channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2504-2526, Oct. 2003.

[27] G. Barriac and U. Madhow, “Characterizing outage rates for space-time communication over wideband channels,” *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2198-2208, Dec. 2004.

[28] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, “A stochastic model of the temporal and azimuthal dispersion seen at the base station in outdoor propagation environments,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 437-447, Mar. 2000.

[29] L. You, X. Gao, X. Xia, N. Ma, and Y. Peng, “Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels,” *IEEE Trans. Wirel. Commun.*, vol. 14, no. 6, pp. 3352-3366, Jun. 2015.

[30] K. E. Baddour and N. C. Beaulieu, “Robust Doppler spread estimation in nonisotropic fading channels,” *IEEE Trans. Wirel. Commun.*, vol. 4, no. 6, pp. 2677-2682, Nov. 2005.

[31] F. Bellili, Y. Selmi, S. Affes, and A. Ghrayeb, “A low-cost and robust maximum likelihood joint estimator for the Doppler spread and CFO parameters over flat-fading Rayleigh channels,” *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3467-3478, Aug. 2017.

[32] I. S. Gradshteyn, and I. M. Ryzhik, *Table of integrals, series, and products*. New York, NY, USA: Academic, 2007.