# Estimating Millimeter Wave Channels Using Out-of-Band Measurements

Anum Ali[†], Nuria González-Prelcic[‡], and Robert W. Heath Jr.[†]
[†]The University of Texas at Austin, Austin, TX 78712.
[‡] Universidade de Vigo, Vigo, Spain.
{anumali,rheath}@utexas.edu, nuria@gts.uvigo.es

***Abstract*—Channel estimation and beam training can be a source of significant overhead in establishing millimeter wave (mmWave) communication links, especially in high mobility applications like connected vehicles. In this paper, we highlight the opportunities and challenges associated with leveraging channel state information acquired at a lower frequency as a form of side information on a higher frequency channel. We focus on the relationship between spatial correlation matrices of sub-6 GHz and mmWave channels. We provide a transform that can be used to relate the spatial correlation matrix derived at one frequency to another much different frequency. We derive an expression for the excess mean squared error and use it to evaluate the performance experienced by using the transformed correlation in mmWave channel estimation.**

## I. Introduction

A vast majority of today's commercial wireless communication systems operate below 6 GHz. There is significantly more bandwidth available in the millimeter-wave (mmWave) spectrum (i.e., $30 - 300$ GHz), making mmWave suitable for gigabit-per-second data rates [1]–[4]. One distinguishing feature of mmWave as envisioned is the use of large antenna arrays at both the transmitter and receiver. Configuring the antenna arrays, which maybe done through channel estimation or beam training, is a significant source of overhead [5].

MmWave systems will likely be deployed in conjunction with lower frequency systems to provide wide area control signals [6], [7]. In this paper, we expose what might be learned about the mmWave channel from much lower frequency channel measurements. Specifically, we investigate the possibility of using the *spatial correlation matrix* of a sub-6 GHz system as out-of-band side information about the mmWave channel. The spatial correlation matrix can be used to reduce training overhead [8] and may help in procoder and combiner design [9]. Estimating the channel or the spatial correlation directly using mmWave measurements requires overhead due to hardware constraints, though sparsity can reduce the measurements required [5].

Translating statistical information from one frequency to another appears in work on beamforming reciprocity in frequency division duplex systems [10]–[15]. These studies highlight that although the propagation channels in uplink (UL) and downlink (DL) are not reciprocal, the spatial information is consistent [10]. For example, in measurements (for 1935 MHz UL and 2125 MHz DL), it was observed that the deviation in angle-of-arrivals (AoA) of dominant paths of UL and DL is small with high probability [13]. This similarity can be exploited to use UL measurements to deduce DL channel information. Several strategies have been proposed to translate spatial correlation, based on least squares [10], [12], [14] and minimum variance distortion-less response [11]. In [15], a spatio-temporal correlation translation strategy was proposed based on two-dimensional interpolation. The translation strategies [10]–[12], [14], [15] work under a key assumption about the congruency (or agreement) of spatial information at the two frequencies under consideration, which is true when the duplex gap is small (e.g., 190 MHz in [13]). High spatial congruency, however, is unlikely when the frequency separation is large. Furthermore, sub-6 GHz to mmWave correlation translation involves antenna arrays with different number of elements, requiring a transformation from a smaller to larger spatial correlation matrix.

In this paper, we propose spatial correlation translation from sub-6 GHz to mmWave frequencies. First, we highlight that spatial congruency between sub-6 GHz and mmWave systems is reduced as the propagation environment has a different effect on low wavelength signals [16]. With the help of ray-tracing, we show that many but not all paths at sub-6 GHz and mmWave overlap. Second, we propose two translation strategies for the specific case of a narrowband single-input multiple-output (SIMO) system model with comparable apertures at sub 6-GHz and mmWave frequencies. We propose a translation strategy that extends [15] to the case where the known and translated correlation matrices differ in size. The second translation approach is based on the parametric estimation of the mean AoA and angle spread. We derive a metric called the excess mean squared error, and use it to compare the performance of a minimum mean square error estimator (MMSE) that uses a mismatched spatial correlation matrix. Simulations compare performance of the two proposed approaches with model mismatch.

Out-of-band measurements have been proposed and validated for mmWave beam steering in indoor 60 GHz WiFi [17]. That work proposes to retrieve directional information from legacy WiFi to reduce the beam steering overhead for the 60 GHz WiFi, and confirms the value of out-of-band information. In our paper, we focus on the congruency of the spatial correlation whereas [17] considers primarily the line-of-sight channel directions. We also quantify the impact of estimation error based on translated covariances.

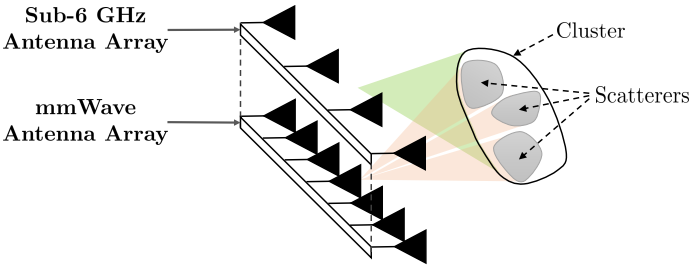


Fig. 1: Array size and spatial resolution: Three scatterers get resolved with larger array, whereas they are seen as a cluster with smaller array.

*Notation:* bold lowercase $\mathbf{x}$ is used for column vector, bold uppercase $\mathbf{X}$ is used to denote matrices, non-bold letters $x$, $X$ are used to denote scalar values. Superscript $\mathsf{T}$ and $*$ represent the transpose and conjugate transpose. $\mathbf{0}$ and $\mathbf{I}$ denote the zero vector and identity matrix respectively. $\mathcal{CN}(\mathbf{x}, \mathbf{X})$ denotes a complex circularly symmetric Gaussian random vector with mean $\mathbf{x}$ and correlation $\mathbf{X}$. We use $\mathbb{E}[\cdot]$ to denote expectation and $\|\cdot\|_{\mathrm{F}}$ to denote Frobenius norm. $\mathbf{X} \otimes \mathbf{Y}$ is the Kronecker product of $\mathbf{X}$ and $\mathbf{Y}$. Finally, $\mathrm{vec}(\cdot)$ yields a vector for a matrix argument and $\mathrm{tr}(\cdot)$ denotes the trace operator.

## II. Spatial properties: sub-6 GHz vs mmWave

In this section, we argue that there may be enough spatial congruence to infer spatial information about the mmWave channel from the lower frequency channel. Hence it is worthwhile to use spatial information extracted at a low frequency as out-of-band side information at high frequency.

There are several challenges associated with translating information from one frequency band to another. While the mechanisms for RF propagation are same, their effects are different at mmWave: diffraction is not as important, scattering can be higher, transmission has higher losses, and blockages are more significant. In this paper, we focus on spatial properties of the channel. The larger arrays used at mmWave imply increased spatial resolution. Hence, it is possible that a collection of scatterers that is seen as a cluster in sub-6 GHz (with smaller arrays) gets resolved into multiple scatterers at mmWave as shown in Fig. 1. For similar reasons, the higher bandwidths used for mmWave imply that there will be more resolvability of the taps in the impulse response. Higher spatial and temporal resolvability may lead to channel sparsity, which can be exploited by receiver algorithms. Paths due to transmission through objects, higher order reflections or diffraction may disappear at mmWave, though scattering may lead to more paths.

We useo 3D ray-tracing to demonstrate the congruency of spatial information in sub-6 GHz and mmWave. Ray-tracing has been reasonably successful at predicting site-specific mmWave propagation in the past [18]. We use Wireless Insite® software [19], and setup user equipment (UE) and base station (BS) nodes as shown in Fig. 2. All the nodes use half-wave dipole antennas that are excited with varying frequencies. We have a UE locations grid of $5 \times 7 = 35$ points at a height of 2m, with 5m separation along $x$ and $y$ directions. The UE locations grid shows the same UE placed at the grid points to generate sufficient data for experiments. Three BS are placed on rooftops of buildings as shown. We capture the AoA of dominant paths from UE to BS at different frequencies. As an example, in Fig. 3 we show the azimuth AoAs of the dominant paths from the first UE location (shown with an arrow in Fig. 2) to BS 1. The AoA of the dominant paths are calculated at 900 MHz and 72 GHz, and the powers of received paths are normalized for easier visualization. A majority of paths are common at two frequencies, with a few distinct paths.

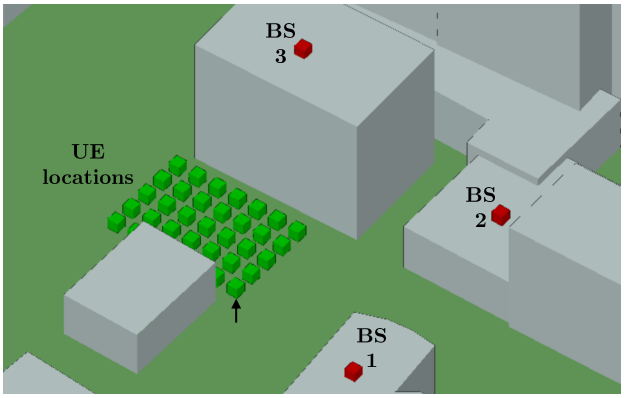


Fig. 2: Ray-tracing setup: An orthographic 3D view of UE and BS nodes placed outdoor in Rosslyn, Virginia. The arrow represents the first transmitter in the UE grid.

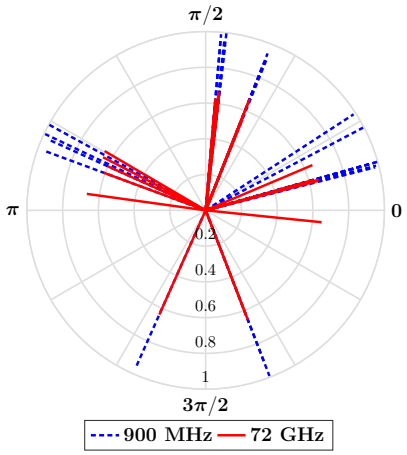


Fig. 3: AoA of dominant paths from first UE to BS 1, at 900 MHz and 72 GHz.

In Fig. 4, we show the fraction of paths with common AoA as we increase the frequency separation. We consider a total of 25 dominant paths (from each UE location to every BS) at each frequency. We use 5 logarithmically spaced test frequencies from 900 MHz to 90 GHz, with 900 MHz as the base case (i.e., the commonality of angles at different frequencies is tested against angles at 900 MHz). We average the fraction of common paths across all UE locations. From the results, we see that as expected, the fraction of common paths decreases

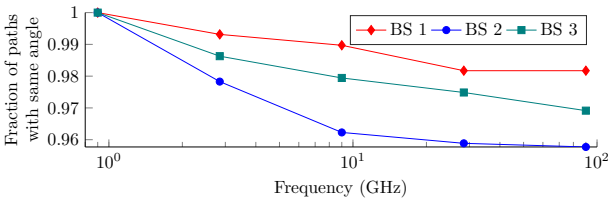


Fig. 4: Fraction of paths with common AoA vs frequency (base case 900 MHz).

with increased frequency separation. That said, the percentage of paths with common angles at frequencies far apart is still well over 90 percent, implying strong congruency in spatial information. We propose to exploit this congruency through the spatial correlation.

## III. Spatial correlation at sub-6 GHz and mmWave

In this section, we present a simple uplink system model. Subsequently, we highlight the structure in correlation matrices and outline the proposed correlation translation strategies.

### A. SIMO system model

Consider a communication link from a single-antenna UE to a BS equipped with $M$ antennas. Under a narrowband channel model, the $M$-dimensional received signal at the BS can be written as

$$\mathbf{y} = \mathbf{h}s + \mathbf{v}, \tag{1}$$

where $\mathbf{h}$ is the $M \times 1$ channel vector, $s$ denotes the transmitted symbol, and $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathbf{v}}^2 \mathbf{I}_M)$ is the additive white Gaussian noise. We use subscript $(\cdot)_{\mathrm{L}}$ to denote the parameters of low frequency system and $(\cdot)_{\mathrm{H}}$ for high frequency system, when the need arises.

For the channel $\mathbf{h}$, we assume that waves arriving from similar directions are grouped into a cluster. Using this modeling approach, the cluster is associated with a mean AoA $\bar{\theta}$, an angle spread $\sigma_\theta$, and a power azimuth spectrum (PAS) with characteristic function $\mathsf{B}(x)$ corresponding to $\sigma_\theta = 1$.

The spatial correlation matrix for the the channel vector $\mathbf{h}$ is defined as $\mathbf{R} = \mathbb{E}[\mathbf{h}\mathbf{h}^*]$. Under a small angle spread assumption [20], the spatial correlation matrix for a uniform linear array (ULA) can be represented as

$$[\mathbf{R}]_{i,j} = e^{\mathrm{j}(i-j)2\pi \frac{d}{\lambda} \sin(\bar{\theta})} \mathsf{B}\left( (i-j) 2\pi \frac{d}{\lambda} \sin(\bar{\theta}) \sigma_\theta \right). \tag{2}$$

For convenience, we use $\rho(\frac{d}{\lambda}(i-j)) = [\mathbf{R}]_{i,j}$. This correlation function will be used for spatial correlation translation assuming the same mean AoA and angle spread. The arguments highlight the dependence of correlation function on (i) interelement spacing $d$, (ii) frequency $f$ (via wavelength $\lambda$), and (iii) the relative distance of antenna elements $i$ and $j$ via index difference $m = i - j$. Furthermore, the correlation function is symmetric $\rho(\frac{d}{\lambda}m) = \rho^*(-\frac{d}{\lambda}m)$.

Now, consider a sub-6 GHz communication system operating at $f_{\mathrm{L}}$, and an mmWave communication system operating at $f_{\mathrm{H}}$. Assume that the two systems have co-located horizontally aligned receiver antenna arrays, and further that the physical size is same for both arrays, as shown in Fig. 1. The same physical size for both arrays is well motivated as due to smaller wavelengths at mmWave, more antennas can be packed into the same physical space.

### B. Structure in $\mathbf{R}_{\mathrm{L}}$ and $\mathbf{R}_{\mathrm{H}}$

To provide a concrete example, in this section we consider the differences between $\mathbf{R}_{\mathrm{L}}$ and $\mathbf{R}_{\mathrm{H}}$ assuming the same $\bar{\theta}$, $\sigma_\theta$, and PAS given by the truncated Laplacian distribution

$$P_\theta(\theta) = \begin{cases} \dfrac{\beta}{\sqrt{2}\sigma_\theta} e^{-|\sqrt{2}\theta/\sigma_\theta|} & \text{if } \theta \in [-\pi, \pi] \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $\beta = 1/(1 - e^{-\sqrt{2}\pi/\sigma_\theta})$ is a normalization factor needed to make the function integrate to one. With the distribution in (3), the correlation function for system operating at $f_{\mathrm{L}}$ is [21]

$$\rho_{\mathrm{L}}\left(\frac{d_{\mathrm{L}}}{\lambda_{\mathrm{L}}} m_{\mathrm{L}}\right) = \frac{\beta e^{\mathrm{j}\frac{2\pi}{\lambda_{\mathrm{L}}} d_{\mathrm{L}} m_{\mathrm{L}} \sin(\bar{\theta})}}{1 + \frac{\sigma_\theta^2}{2}[\frac{2\pi}{\lambda_{\mathrm{L}}} d_{\mathrm{L}} m_{\mathrm{L}} \cos(\bar{\theta})]^2}. \tag{4}$$

The correlation function $\rho_{\mathrm{H}}(\frac{d_{\mathrm{H}}}{\lambda_{\mathrm{H}}} m_{\mathrm{H}})$ is defined analogously. For a given value of $m$, a dilation factor $\kappa = \frac{d_{\mathrm{H}}}{d_{\mathrm{L}}} \frac{\lambda_{\mathrm{L}}}{\lambda_{\mathrm{H}}}$ captures the difference in the correlation due to the differences in interelement spacing and frequencies. If the relative antenna spacing is fixed at $d/\lambda = 1/2$ for both frequencies, then $\kappa = 1$.

Now suppose that $\kappa = 1$. For a receiver with $M_{\mathrm{L}}$ elements, we have $m_{\mathrm{L}} \in [-M_{\mathrm{L}} + 1, -m_{\mathrm{L}} + 2, \cdots, M_{\mathrm{L}} - 1]$. Similarly, $m_{\mathrm{H}} \in [-M_{\mathrm{H}} + 1, -m_{\mathrm{H}} + 2, \cdots, M_{\mathrm{H}} - 1]$. As the number of antennas at mmWave is higher than sub-6 GHz, the index difference $m_{\mathrm{H}}$ has a wider range. Hence, the correlation function $\rho_{\mathrm{L}}$ needs to be extrapolated to obtain $\rho_{\mathrm{H}}$. In terms of correlation matrices, a larger index difference for mmWave implies that a bigger correlation matrix needs to be constructed from a correlation matrix with smaller dimensions. For general values of $\kappa$, both interpolation and extrapolation are required.

### C. Correlation translation

We now propose two approaches for translating known $\mathbf{R}_{\mathrm{L}}$ to get an estimate of $\mathbf{R}_{\mathrm{H}}$. In prior work [10]–[15], the same array was used for both carriers, thus the only source of discrepancy was the slight differences in $d/\lambda$. In our case, the differences come from both $d/\lambda$ and differences in the number of elements. Along these lines, the first translation approach is a generalization of [15], and we call it non-parametric approach.

Interpolation and extrapolation of $\rho_{\mathrm{L}}$ is required to estimate the desired correlation $\hat{\mathbf{R}}_{\mathrm{H}}$ from $\mathbf{R}_{\mathrm{L}}$. We propose to interpolate the magnitude and phase functions separately, motivated by the natural separation of terms into magnitude and phase [15], [20], [21] as shown in (2).

The correlation magnitude for a clustered channel is smooth. For example, a single cluster channel with truncated Laplacian

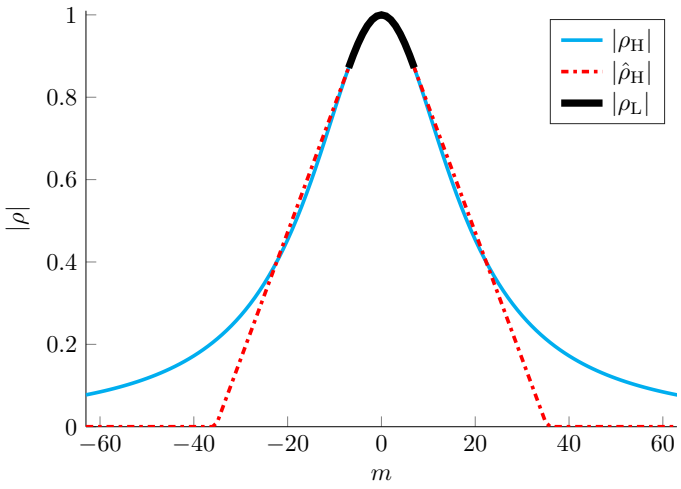


Fig. 5: Correlation magnitude extrapolation: The true and estimated magnitudes of correlation function, when $\bar{\theta} = 45°$ and $\sigma_\theta = 2°$.

AoA distribution has correlation magnitude [21]

$$\left|\rho\Big(\frac{d}{\lambda}m\Big)\right| = \frac{\beta}{1 + \frac{\sigma_\theta^2}{2}[\frac{2\pi}{\lambda}d(i-j)\cos(\bar{\theta})]^2}. \tag{5}$$

As such we expect that *spline* interpolation with the right choice of parameters will provide good performance. Good quality extrapolation, however, is more challenging. As an example, consider extrapolation from $M_{\mathrm{L}} = 8$ antenna system to a $M_{\mathrm{H}} = 64$ antenna system, with $\kappa = 1$. The correlation magnitudes for this setting are shown in Fig. 5. We have found that, for settings where long-range extrapolation is necessary, it is best to resort to *linear* extrapolation truncated at 0.

Note that the phase, $\psi(\rho) = \arg(\rho)$, can be unambiguously determined only in the interval $(-\pi, \pi]$. That said, the phase of $\rho$, $\psi(\rho)$, increases linearly in the argument of $g$ and hence the jumps of $2\pi$ can be observed when the phase leaves this interval. For interpolation of phase, the actual linearly increasing phase needs to be reconstructed. Hence, the observed phase is first *unwrapped* and then linearly interpolated (or extrapolated) before it is re-sampled for correlation translation. Based on the phase structure of (2), the phase interpolation should have little error if done correctly.

After the magnitude and phase are translated, they are combined to generate the translated correlation matrix. Due to the interpolation (and especially extrapolation) operations, the resulting matrix may not be positive definite. To enforce positive definiteness, we replace any negative eigenvalue with the least positive eigenvalue of the translated correlation.

As a second strategy, we suggest a parametric approach. Prior work has considered the specific problem of estimating both the AoA and the angle spread jointly from an empirically estimated spatial correlation matrix. We propose to use spread Root-MUSIC [20] with the parameters set for a single cluster. Then we use the estimated $\bar{\theta}$ and $\sigma_\theta$ in (2) to generate $\hat{\mathbf{R}}_{\mathrm{H}}$. A disadvantage of the parametric approach is reliance on a particular AoA distribution and the corresponding correlation expressions for construction of mmWave correlation matrix.

## IV. Channel Estimation with Translated Correlation and Excess Error

The translated correlation $\hat{\mathbf{R}}_{\mathrm{H}}$ can be used for many different purposes in an mmWave system. In this section, we consider the specific application of MMSE channel estimation. In particular, the derived $\hat{\mathbf{R}}_{\mathrm{H}}$ based on $\mathbf{R}_{\mathrm{L}}$ is used in place of the true $\mathbf{R}_{\mathrm{H}}$ in an MMSE estimator. Note that hardware constraints like the use of hybrid precoding and combining make it difficult to estimate an mmWave channel directly as considered here. We defer consideration of different receive hardware architectures to future work.

Assume that the mmWave channel, $\mathbf{h}$, is estimated using a block of $K$ symbols and collectively denote the $K$ training symbols as $\mathbf{s}^{\mathsf{T}} = [s_1, s_2, \cdots, s_K]$. Now, we can write the received signals in the following manner

$$\begin{aligned}\mathbf{Y} &= [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_K] \\ &= \mathbf{h}[s_1, s_2, \cdots, s_K] + [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_K] \\ &= \mathbf{h}\mathbf{s}^{\mathsf{T}} + \mathbf{V}. \end{aligned} \tag{6}$$

Stacking the columns of $\mathbf{Y}$, we obtain

$$\tilde{\mathbf{y}} = \mathrm{vec}[\mathbf{Y}] = (\mathbf{s} \otimes \mathbf{I}_M)\mathbf{h} + \mathrm{vec}[\mathbf{V}] = \tilde{\mathbf{S}}\mathbf{h} + \tilde{\mathbf{v}}, \tag{7}$$

where $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{v}}$ are $KM \times 1$ vectors, and we denote $KM \times M$ matrix $\tilde{\mathbf{S}} = \mathbf{s} \otimes \mathbf{I}_Q$. The linear minimum mean square error (LMMSE) estimate (which is identical to the MMSE in this case) with true correlation $\mathbf{R}_{\mathrm{H}}$, is [22]

$$\hat{\mathbf{h}} = \mathbf{R}_{\mathrm{H}}\tilde{\mathbf{S}}^*(\tilde{\mathbf{S}}\mathbf{R}_{\mathrm{H}}\tilde{\mathbf{S}}^* + \sigma_{\mathbf{v}}^2\mathbf{I}_{KM})^{-1}\tilde{\mathbf{y}}. \tag{8}$$

The minimum MSE of the LMMSE estimator is [22]

$$\mathrm{MMSE} = \mathrm{tr}\Big((\mathbf{R}_{\mathrm{H}}^{-1} + \frac{K}{\sigma_{\mathbf{v}}^2}\mathbf{I}_M\Big). \tag{9}$$

With imperfect correlation knowledge, however, the MSE has an additional component that we call *excess mean squared error* (EMSE). Hence, if we use translated correlation $\hat{\mathbf{R}}_{\mathrm{H}}$, in (8), the MSE is

$$\mathrm{MSE} = \mathrm{MMSE} + \mathrm{EMSE}. \tag{10}$$

The EMSE for additive perturbation in the real domain was derived in [23]. We provide its extension to the setup in (7), where the unknown is scaled (in our case by $\tilde{\mathbf{S}}$) and additively perturbed. We use the following theorem to quantify the EMSE incurred by the mismatch between true correlation $\mathbf{R}_{\mathrm{H}}$ and the translated correlation $\hat{\mathbf{R}}_{\mathrm{H}}$.

**Theorem 1.** *If the true correlation matrix, $\mathbf{R}_{\mathrm{H}}$, and the translated correlation matrix, $\hat{\mathbf{R}}_{\mathrm{H}}$, differ by an additive perturbation $\mathbf{\Gamma}$ (i.e., $\hat{\mathbf{R}}_{\mathrm{H}} = \mathbf{R}_{\mathrm{H}} + \mathbf{\Gamma}$) that ensures positive definiteness of $\hat{\mathbf{R}}_{\mathrm{H}}$, then the* EMSE *of the estimator is*

$$\mathrm{EMSE} = \mathrm{tr}\left(\mathbf{\Upsilon}\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}\mathbf{\Upsilon}^*\right), \tag{11}$$

*where* $\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}} = \mathbb{E}[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^*] = (\tilde{\mathbf{S}}\mathbf{R}_{\mathrm{H}}\tilde{\mathbf{S}}^* + \sigma_{\mathrm{v}}^2\mathbf{I}_{KM})$ *and*

$$\boldsymbol{\Upsilon} = \boldsymbol{\Gamma}\tilde{\mathbf{S}}^*(\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}} + \tilde{\mathbf{S}}\boldsymbol{\Gamma}\tilde{\mathbf{S}}^*)^{-1} - \mathbf{R}_{\mathrm{H}}\tilde{\mathbf{S}}^*\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}^{-1}}\tilde{\mathbf{S}}(\mathbf{I}_M + \boldsymbol{\Gamma}\tilde{\mathbf{S}}^*\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1}\tilde{\mathbf{S}})^{-1}\boldsymbol{\Gamma}\tilde{\mathbf{S}}^*\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1}. \quad (12)$$

*Proof:* See Appendix A. ∎

We use the results of Theorem 1 to evaluate the estimator performance using the mismatched correlation matrix.

## V. Simulation Results

In this section, we present simulation results of two experiments. In the first experiment, we test the correlation translation accuracy as a function of frequency separation. We simulate a system operating at 900 MHz with 4 antennas and half wavelength antenna spacing. For this sub-6 GHz system we assume a cluster with $\bar{\theta} = 45°$ and $\sigma_\theta = 15°$. We use the theoretically computed correlation matrix from (2) with a truncated Laplacian distribution. We then test the correlation translation accuracy at different frequencies. We assume a half wavelength spacing at each frequency, and hence as we increase frequency (and keep the physical array size constant), we increase the number of elements in the array. To simulate the effect of possible differences in spatial information across frequencies, we test the translation accuracy for three high frequency angle spreads $\sigma_\theta = 5°$, $10°$ and $15°$, keeping other parameters consistent with 900 MHz case. We use correlation matrix distance (CMD) to test the conversion accuracy [24]. The CMD between a matrix, $\mathbf{R}_{\mathrm{H}}$, and its estimate, $\hat{\mathbf{R}}_{\mathrm{H}}$, is defined as

$$d_{\mathrm{corr}}(\mathbf{R}_{\mathrm{H}}, \hat{\mathbf{R}}_{\mathrm{H}}) = 1 - \frac{\mathrm{tr}(\mathbf{R}_{\mathrm{H}}\hat{\mathbf{R}}_{\mathrm{H}})}{\|\mathbf{R}_{\mathrm{H}}\|_{\mathrm{F}}\|\hat{\mathbf{R}}_{\mathrm{H}}\|_{\mathrm{F}}} \quad \in (0, 1]. \quad (13)$$

The results for correlation translation are shown in Fig. 6. It can be observed that CMD increases with frequency separation. This behavior is more pronounced when the angle spread is $5°$, which is the case of greatest mismatch. Note that a smaller angle spread implies higher correlation. The channels for 900 MHz are mildly correlated (i.e., $\sigma_\theta = 15°$). When channels for other frequencies are highly correlated (i.e., $\sigma_\theta = 5°$), the translation (from mildly correlated to highly correlated channels) understandably yields a higher CMD. The parametric approach performs better compared with non-parametric approach. This is expected as the postulated AoA distribution used for constructing the translated correlation (i.e., truncated Laplacian) is the true AoA distribution. Further, in this simulation we use the true (not estimated) low frequency covariance matrix. Intuitively, if the postulated and actual AoA distributions did not match, the advantage of parametric approach over non-parametric approach will diminish.

In the second experiment, we test the EMSE (11) of the LMMSE estimator with translated correlation matrices. We estimate the channel of an mmWave system operating at 30 GHz with $M_{\mathrm{H}} = 32$. The number of pilots $K$ for LMMSE is equal to the number of antennas in mmWave system $M_{\mathrm{H}}$. The

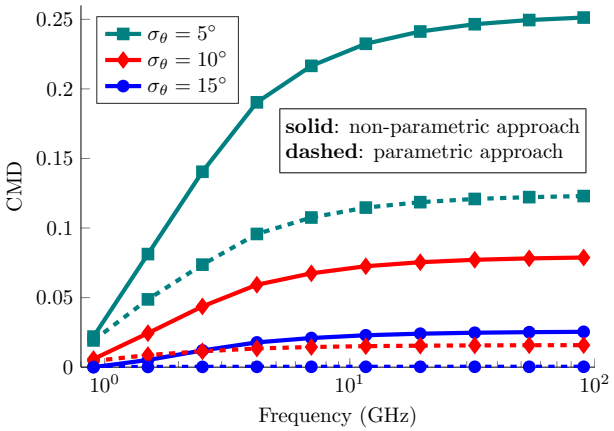


Fig. 6: Correlation conversion accuracy: Correlation matrix distance vs frequency separation with varying angle spreads.

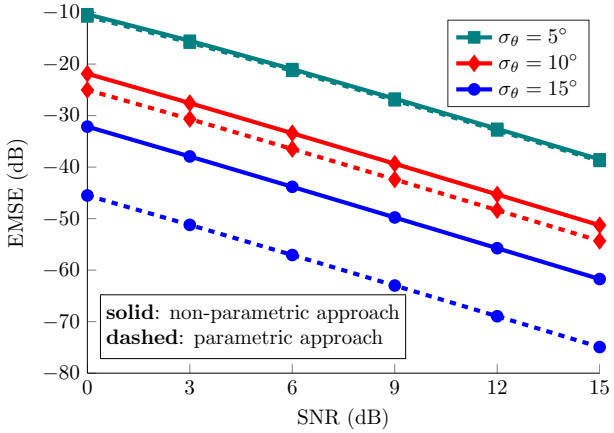


Fig. 7: Excess mean squared error: EMSE (dB) vs SNR (dB) for translated correlations with varying angle spreads.

translated correlation matrix is obtained from a sub-6 GHz system operating at 900 MHz with 4 antenna elements spaced at half wavelength. The mean AoA and angle spreads for sub-6 GHz and mmWave frequency are consistent with the first experiment. The EMSE results, as a function of signal-to-noise ratio (SNR), are shown in Fig. 7. We have the worst EMSE when high frequency angle spread is $5°$ by the same reasoning as in previous experiment. Finally, as the postulated and true AoA distributions match, the parametric approach expectedly yields lower EMSE compared to the non-parametric approach.

## VI. Conclusion

In this paper we argued that spatial correlation information is related at sub-6 GHz and mmWave bands. We proposed two methods to translate the lower frequency spatial correlation to the higher frequency: a non-parametric approach (based on interpolation/extrapolation) and a parametric approach (based on estimates of AoA and angle spreads). To evaluate their performance, we derived an expression for the excess mean squared error of the LMMSE estimator that uses the mismatched spatial correlation. We found that the performance

gap between translated and true correlation increased as a function of frequency, decreased with SNR, and increased with the angle spread gap. There are many directions for future work that involve expanding the model to include other array geometries, more complicated channels, broadband channels, multiple transmit and receive antennas, and hardware constraints. There are other ways to exploit spatial correlation information besides LMMSE estimation including initialization for beam training or statistical precoding followed by further channel estimation. Finally, there are other kinds of congruence that could be explored, notably sparsity which can be used in compressed channel estimation and compressive covariance estimation.

## APPENDIX A
## PROOF OF THEOREM 1

The channel estimate, (8), can be written as

$$\begin{aligned}\hat{\mathbf{h}} &= (\mathbf{R}_{\mathrm{H}} + \boldsymbol{\Gamma})\tilde{\mathbf{S}}^*(\tilde{\mathbf{S}}(\mathbf{R}_{\mathrm{H}} + \boldsymbol{\Gamma})\tilde{\mathbf{S}}^* + \sigma_{\mathbf{v}}^2\mathbf{I}_{KM})^{-1}\tilde{\mathbf{y}}, \\ &= (\mathbf{R}_{\mathrm{H}} + \boldsymbol{\Gamma})\tilde{\mathbf{S}}^*(\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}} + \tilde{\mathbf{S}}\boldsymbol{\Gamma}\tilde{\mathbf{S}}^*)^{-1}\tilde{\mathbf{y}}. \end{aligned} \quad (14)$$

By applying the matrix inversion lemma, we have that

$$\begin{aligned}&(\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}} + \tilde{\mathbf{S}}\boldsymbol{\Gamma}\tilde{\mathbf{S}}^*)^{-1} = \\ &\quad \mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1} - \mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1}\tilde{\mathbf{S}}(\mathbf{I}_M + \boldsymbol{\Gamma}\tilde{\mathbf{S}}^*\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1}\tilde{\mathbf{S}})^{-1}\boldsymbol{\Gamma}\tilde{\mathbf{S}}^*\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1}, \end{aligned} \quad (15)$$

which combined with (14) yields

$$\hat{\mathbf{h}} = (\mathbf{R}_{\mathrm{H}}\tilde{\mathbf{S}}^*\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}^{-1} + \boldsymbol{\Upsilon})\tilde{\mathbf{y}}, \quad (16)$$

where $\boldsymbol{\Upsilon}$ is defined in (12). The error correlation matrix $\mathbf{C}$ is

$$\mathbf{C} = \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}})(\mathbf{h} - \hat{\mathbf{h}})^*], \quad (17)$$

which after simple algebraic manipulation becomes

$$\mathbf{C} = \big(\mathbf{R}_{\mathrm{H}}^{-1} + \frac{K}{\sigma_{\mathbf{v}}^2}\mathbf{I}\big)^{-1} + \boldsymbol{\Upsilon}\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}\boldsymbol{\Upsilon}^*. \quad (18)$$

The $\mathrm{MSE}$ is the trace of error correlation matrix, i.e.,

$$\begin{aligned}\mathrm{MSE} &= \mathrm{tr}\Big(\big(\mathbf{R}_{\mathrm{H}}^{-1} + \frac{K}{\sigma_{\mathbf{v}}^2}\mathbf{I}\big)^{-1} + \boldsymbol{\Upsilon}\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}\boldsymbol{\Upsilon}^*\Big) \\ &\overset{(a)}{=} \mathrm{tr}\Big(\big(\mathbf{R}_{\mathrm{H}}^{-1} + \frac{K}{\sigma_{\mathbf{v}}^2}\mathbf{I}\big)^{-1}\Big) + \mathrm{tr}\Big(\boldsymbol{\Upsilon}\mathbf{R}_{\tilde{\mathbf{y}},\mathrm{H}}\boldsymbol{\Upsilon}^*\Big) \\ &= \mathrm{MMSE} + \mathrm{EMSE}. \end{aligned} \quad (19)$$

where the first term in $(a)$ is $\mathrm{MMSE}$ from (9) and the second term is $\mathrm{EMSE}$ as claimed in Theorem 1.

## VII. ACKNOWLEDGEMENT

This research was partially supported by the U.S. Department of Transportation through the Data-Supported Transportation Operations and Planning (D-STOP) Tier 1 University Transportation Center and by the Texas Department of Transportation under Project 0-6877 entitled "Communications and Radar-Supported Transportation Operations and Planning (CAR-STOP)".

## REFERENCES

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, 2011.
[2] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.
[3] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
[4] T. Bai and R. W. Heath Jr., "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, 2015.
[5] R. W. Heath Jr., N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, to appear.
[6] Y. Kishiyama, A. Benjebbour, T. Nakamura, and H. Ishii, "Future steps of LTE-A: evolution toward integration of local area and wide area systems," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 12–18, 2013.
[7] R. C. Daniels and R. W. Heath Jr, "Multi-band Modulation, Coding, and Medium Access Control," IEEE 802.11-07/2780R1, pp. 1–18, November 2007.
[8] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 546–557, 2004.
[9] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath Jr, "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Proc. Inf. Theory Appl. (ITA) Wksp*, 2013, pp. 1–5.
[10] T. Asté, P. Forster, L. Féty, and S. Mayrargue, "Downlink beamforming avoiding DOA estimation for cellular mobile communications," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 1998, pp. 3313–3316.
[11] K. Hugl, J. Laurila, and E. Bonek, "Downlink beamforming for frequency division duplex systems," in *Proc. IEEE Glob. Telecom. Conf. (GLOBECOM)*, vol. 4, 1999, pp. 2097–2101.
[12] Y.-C. Liang and F. P. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 222–232, 2001.
[13] K. Hugl, K. Kalliola, and J. Laurila, "Spatial reciprocity of uplink and downlink radio channels in FDD systems," COST 273 Technical Document TD(02) 066, 2002.
[14] B. K. Chalise, L. Haering, and A. Czylwik, "Robust uplink to downlink spatial covariance matrix transformation for downlink beamforming," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 5, 2004, pp. 3010–3014.
[15] M. Jordan, X. Gong, and G. Ascheid, "Conversion of the spatio-temporal correlation from uplink to downlink in FDD systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2009, pp. 1–6.
[16] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Prentice Hall, 2014.
[17] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: mm-wave beam steering without in-band measurement," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 2015, pp. 2416–2424.
[18] B. Neekzad, K. Sayrafian-Pour, J. Perez, and J. S. Baras, "Comparison of ray tracing simulations and millimeter wave channel sounding measurements," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, 2007, pp. 1–5.
[19] "Wireless Insite," http://www.remcom.com/wireless-insite.
[20] M. Bengtsson and B. Ottersten, "Low-complexity estimators for distributed sources," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2185–2194, 2000.
[21] A. Forenza, D. J. Love, and R. W. Heath Jr, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1924–1934, 2007.
[22] A. H. Sayed, *Fundamentals of adaptive filtering*, 2003: Wiley.
[23] I. Esnaola, A. M. Tulino, and H. V. Poor, "Mismatched MMSE estimation of multivariate Gaussian sources," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2012, pp. 716–720.
[24] M. Herdin, N. Czink, H. Özcelik, and E. Bonek, "Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, 2005, pp. 136–140.