Beamspace Local LMMSE: An Efficient Digital Backend for mmWave Massive MIMO

Mohammed Abdelghany*, Upamanyu Madhow* and Antti Tölli†
* *Department of ECE, University of California at Santa Barbara, Santa Barbara, CA 93106 USA*
† *Centre for Wireless Communications, University of Oulu, P.O. Box 4500, 90014, Finland*
e-mail: mabdelghany@ucsb.edu, madhow@ece.ucsb.edu, antti.tolli@oulu.fi

***Abstract*—We explore an all-digital architecture for a mmWave massive MIMO cellular uplink in which the number of users scales with the number of antenna elements at the base station. We consider the design of multiuser detection strategies after a spatial DFT, which concentrates the energy of each user onto a few DFT bins in "beamspace." In this paper, we propose and investigate a local LMMSE receiver that exploits this property, using a small window in beamspace to demodulate each user. The proposed architecture is computationally efficient: the required window size depends on load factor (the number of users divided by the number of antenna elements) and does not scale with the number of elements. We also show that adaptive implementations of such local LMMSE receivers naturally extend to provide implicit channel estimation.**

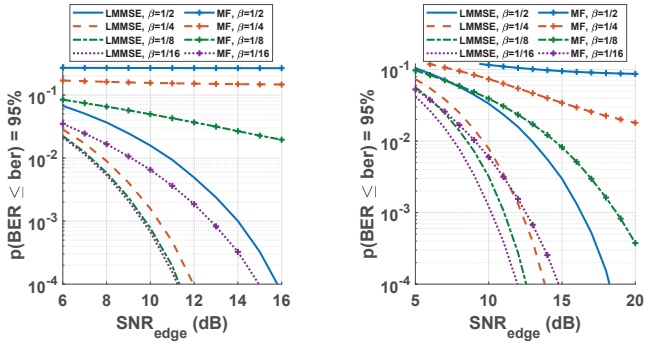
***Index Terms*—Beamspace, LMMSE, Local LMMSE, low-complexity Multiuser detection.**

I. Introduction

All-digital architectures enable taking full advantage of the large number of antennas that can be integrated in mmWave transceivers, with fully flexible beamforming that enables the number of simultaneous users K sharing the band to scale with the number of antennas N, with scaling ratio, or *load factor*, $\beta = \frac{K}{N}$. Standard criteria for beamforming include spatial matched filtering (MF), as well as linear interference suppression using the zero forcing (ZF) or linear minimum mean square error (LMMSE) criteria. Fig. 1(a) depicts the raw bit error rate (BER) achieved by 95% of the mobiles for the picocellular uplink considered in this paper. Clearly, interference suppression becomes necessary for moderate load factors (e.g., $\beta > 1/16$), where MF performance is far inferior to that of LMMSE, with the gap persisting even if power control is employed, as shown in Fig. 1(b). However, the computational complexity of LMMSE detection becomes prohibitive for large K and N.

Recent efforts at complexity reduction, for both uplink and downlink, include two-stage beamforming strategies [1]–[4]. In [1], a statistical outer beamformer based on grouping mobiles based on similar correlation matrices reduces the effective spatial dimension of the equivalent channels [1], [2]. This is followed by an inner beamformer that suppresses both intra- and inter-group interference, resulting in significant reduction in computation [3], [4].

In the present paper, we propose a Beamspace Local LMMSE algorithm which leverages the sparsity of the spatial channel in mmWave bands. A spatial discrete Fourier transform (DFT) is employed to concentrate the energy of each mobile into a smaller number of DFT bins, i.e., in "beamspace." We show that performance close to that of standard LMMSE can be obtained by a local LMMSE detector operating on a beamspace window of a size that does not scale with N. We provide analytical rules of thumb for choosing window size as a function of load factor β and target outage rate. We also show how our architecture provides a low-complexity solution for implicit channel estimation via an efficient adaptive implementation.



(a) BER without power control (b) BER with power control

Figure 1: Massive MIMO uplink performance using MF and LMMSE receivers for $\beta = K/N = \{1/16, 1/8, 1/4, 1/2\}$ and $N = 256$.

II. System Model

The massive MIMO system model is depicted in Fig. 2: a sector covered by a base station which can scan horizontally with a 1D half-wavelength spaced N-element array, serving K mobiles (assumed to each have a single antenna for simplicity).

The linear multiuser detector comprises a beamformer weights acquisition module, and a beamformer module. Using the received training matrix $\mathbf{Y}$, defined later, and the training sequences for each user $\{\mathbf{t}_i\}_{i=1}^K$, the weight acquisition block generates the beamformer weights $\{\mathbf{w}_i\}_{i=1}^K$. These weights are used by the beamformer module to estimate the users' data vector $\mathbf{x}$ out of the received vector $\mathbf{y}$.

We assume a line-of-sight (LoS) channel from each mobile to the base station. The channel vector for the k^{th} user can be written as follows:

$$\mathbf{h}_k = A_k \left[1\, e^{j\Omega_k}\, e^{j2\Omega_k}\, \ldots\, e^{j(N-1)\Omega_k}\right]^\intercal, \qquad (1)$$

where $A_k^2 = \left(\frac{\lambda}{4\pi R_k}\right)^2$ depends on the radial location R_k of

978-1-5386-6528-2/19/$31.00 ©2019 IEEE

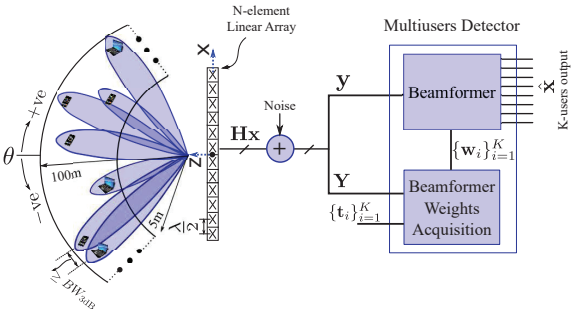


Figure 2: System model for the beamspace massive MIMO.

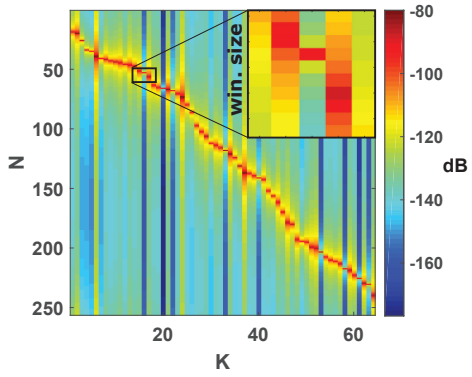


Figure 3: Sparse LoS channel in the beamspace.

user k and wavelength λ, using the Friis formula for path loss. A spatial frequency $\Omega_k = 2\pi \frac{d_x}{\lambda} \sin\theta_k$ defines the angular location of the k^{th} user, where d_x is the inter-distance between antenna elements and chosen to be $\lambda/2$.

The received signal vector $\mathbf{y}$ in the base station is given by

$$\mathbf{y} = \underbrace{\mathbf{H}}_{N\times K}\mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{H} = [\mathbf{h}_1\mathbf{h}_2\dots\mathbf{h}_K]$ is the channel matrix, $\mathbf{x}$ is the users symbols vector, $\mathbb{E}(x_k) = 1$, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ is the AWGN vector. As shown in Fig. 3, the LoS channel is sparse in beamspace.

A. Linear MMSE receiver

The LMMSE receiver is given by $\hat{\mathbf{x}} = \mathbf{W}\mathbf{y}$, such that

$$\mathbf{W} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I})^{-1} = (\mathbf{H}^H\mathbf{H} + \sigma^2\mathbf{I})^{-1}\mathbf{H}^H, \tag{3}$$

where the second equality is a direct result of Sherman–Morrison–Woodbury matrix identity.

B. Implicit channel estimation

Adaptive implementations of the LMMSE receiver implicitly estimate the channel using training sequences for the users of interest. Let $\mathbf{T} = [\mathbf{t}_1\mathbf{t}_2\dots\mathbf{t}_K]^\intercal$ be a matrix that hold the training sequences $\mathbf{t}_k, k = 1,\dots,K$ of length L, such that $||\mathbf{t}_k||_2^2 = L$. We assume that sequences are orthogonal across users, i.e., $\mathbf{t}_i^H\mathbf{t}_j = 0\,\forall i \neq j$. Hence the received training sequence can be written as,

$$\underbrace{\mathbf{Y}}_{N\times L} = \mathbf{H}\mathbf{T} + \mathbf{N}, \tag{4}$$

where $\mathbf{N} = [\mathbf{n}_1\mathbf{n}_2\dots\mathbf{n}_L]$ is the noise matrix. Therefore, the covariance of the received signal and the channel matrix can be computed empirically and approximated at high SNR as (assuming $L > N$)

$$\mathbf{C} = \frac{1}{L}\mathbf{Y}\mathbf{Y}^H = \mathbf{H}\mathbf{H}^H + \frac{1}{L}\mathbf{N}\mathbf{N}^H \approx \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}, \tag{5}$$

and

$$\mathbf{U} = \frac{1}{L}\mathbf{T}\mathbf{Y}^H = \mathbf{H}^H + \frac{1}{L}\mathbf{T}\mathbf{N}^H \approx \mathbf{H}^H, \tag{6}$$

while the computational complexity of the previous two steps are $O(N^2L)$ and $O(\beta N^2 L)$, respectively. The LMMSE solution can be formed as,

$$\hat{\mathbf{x}} = \mathbf{U}\mathbf{C}^{-1}\mathbf{y}. \tag{7}$$

If we used the Cholesky decomposition [5], then the computation complexity of inverting $\mathbf{C}$ is $O(N^3)$, while the complexity of computing $\hat{\mathbf{x}}$ is $O(\beta N^2)$ per received signal vector $\mathbf{y}$.

III. Beamspace Local LMMSE

In general, the observation at each antenna is a superposition of signals received from all the users. For a specific user, the information about its symbol is distributed equally between all the antennas. Hence, the base station must engage all the antenna observations to extract the user's symbol with high precision. In beam domain processing, on the other hand, the base station compresses the information of a user into a smaller number of spatial observations. As a result, the computational complexity can be greatly reduced.

In this section, we first describe the method to attain the weights of beamspace local LMMSE in the absence of explicitly estimated CSI using the received training sequence $\mathbf{Y}$. Then, we explain the beamforming process for the received data vector $\mathbf{y}$.

A. Beamspace Local LMMSE Weight Acquisition

Fig. 4 summarizes the steps of the local LMMSE weights acquisition. Starting by the received training sequence matrix $\mathbf{Y}$, the acquisition steps are given as follows.

1) Calculate the DFT of the received training sequence

The DFT is used to transform the antenna domain signal to beam domain by applying it on each column of the received training sequence matrix $\mathbf{Y}$ to get $\mathbf{Y}'$ as

$$y'_{p,\ell} = \sum_{n=1}^{N} y_{n,\ell} e^{-j2\pi(n-1)(p-1)/N}, \tag{8}$$

where $\mathbf{Y} = [y_{n,\ell}]$ and $\mathbf{Y}' = [y'_{p,\ell}]$, $\forall p = 1,\dots,N$. The fast Fourier transform (FFT) algorithm [6] is applied to realize the DFT operation with a complexity of $O(LN\log N)$ for the whole training sequence matrix.

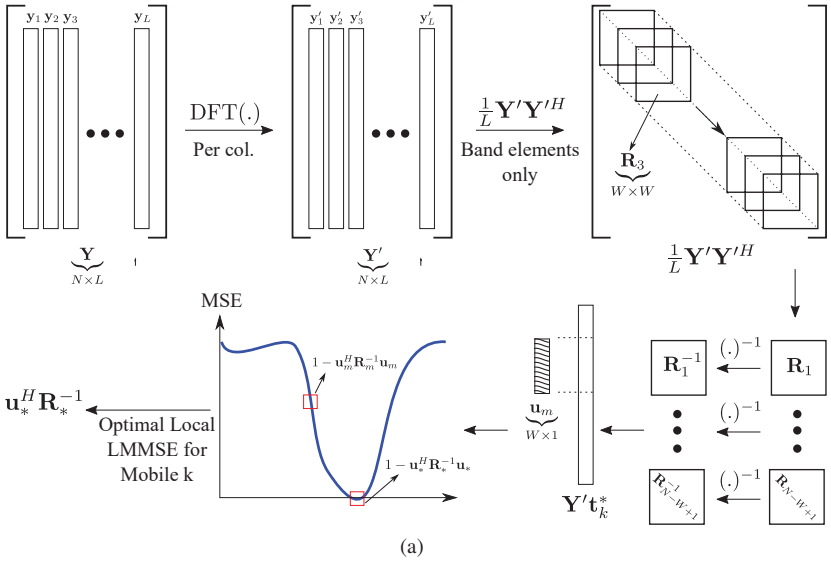
2) Generate the band matrix $\tilde{\mathbf{C}}$

The band matrix $\tilde{\mathbf{C}}$ is generated as

$$\tilde{c}_{n,p} = \begin{cases} \frac{1}{L}\mathbf{Y}'_{[n,*]}\mathbf{Y}'^H_{[p,*]}, & \text{if } |n-p| < W \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

where $\mathbf{Y}'_{[n,*]}$ is the n^{th} row in the matrix $\mathbf{Y}'$. Instead of computing the entire $N \times N$ beam domain sample covariance matrix ($\frac{1}{L}\mathbf{Y}'\mathbf{Y}'^H$), only the dominant elements around the diagonal within a window size W are computed. Hence, the complexity of this step is $O(WNL)$. We define the block matrices on the diagonal $\mathbf{R}_m$ as

$$\underbrace{\mathbf{R}_m}_{W\times W} = \tilde{\mathbf{C}}_{[m:m+W-1,m:m+W-1]}. \tag{10}$$



(a)

Figure 4: Local LMMSE weights acquisition in beamspace.

3) Invert each $\mathbf{R}_m$

One matrix inversion $\mathbf{R}_m^{-1}$ has a complexity of $O(W^3)$. Normally, the matrix inversion should be applied $N-W+1$ times for all possible m, and the resulting total complexity would be $O(NW^3)$. In the following, we propose a method that greatly reduces the number of operations required. The beam domain sample covariance matrix $\tilde{\mathbf{C}}$ has the following structure

$$\tilde{\mathbf{C}} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \dots & c_1 & \mathbf{b}_1^H & \cdot & \dots \\ \dots & \mathbf{b}_1 & \mathbf{A} & \mathbf{b}_2 & \dots \\ \dots & \cdot & \mathbf{b}_2^H & c_2 & \dots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{11}$$

Let us define

$$\mathbf{R}_m = \begin{bmatrix} c_1 & \mathbf{b}_1^H \\ \mathbf{b}_1 & \mathbf{A} \end{bmatrix} \text{ and } \mathbf{R}_{m+1} = \begin{bmatrix} \mathbf{A} & \mathbf{b}_2 \\ \mathbf{b}_2^H & c_2 \end{bmatrix}. \tag{12}$$

Using the elements in (12), $\mathbf{R}_m^{-1}$ can be defined as

$$\begin{aligned} \mathbf{R}_m^{-1} &= \begin{bmatrix} \frac{1}{s_1} & -\frac{1}{s_1}\mathbf{b}_1^H\mathbf{A}^{-1} \\ -\frac{1}{s_1}\mathbf{A}^{-1}\mathbf{b}_1 & \mathbf{A}^{-1} + \frac{1}{s_1}\mathbf{A}^{-1}\mathbf{b}_1\mathbf{b}_1^H\mathbf{A}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} x_{11} & \mathbf{x}_{12}^H \\ \mathbf{x}_{12} & \mathbf{X}_{22} \end{bmatrix}, \end{aligned} \tag{13}$$

where $s_1 = c_1 - \mathbf{b}_1^H\mathbf{A}^{-1}\mathbf{b}_1$ is the Schur complement of block $\mathbf{A}$ of matrix $\mathbf{R}_m$. Given $\mathbf{A}^{-1}$, the inverse $\mathbf{R}_{m+1}^{-1}$ for the block $m+1$ can be computed as

$$\mathbf{R}_{m+1}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \frac{1}{s_2}\mathbf{A}^{-1}\mathbf{b}_2\mathbf{b}_2^H\mathbf{A}^{-1} & -\frac{1}{s_2}\mathbf{A}^{-1}\mathbf{b}_2 \\ -\frac{1}{s_2}\mathbf{b}_2^H\mathbf{A}^{-1} & \frac{1}{s_2} \end{bmatrix}, \tag{14}$$

where $s_2 = c_2 - \mathbf{b}_2^H\mathbf{A}^{-1}\mathbf{b}_2$ is the Schur complement of block $\mathbf{A}$ of matrix $\mathbf{R}_{m+1}$. Finally, $\mathbf{A}^{-1}$ can be computed from the entries of $\mathbf{R}_m^{-1}$ as

$$\mathbf{A}^{-1} = \mathbf{X}_{22} - \frac{1}{x_{11}}\mathbf{x}_{12}\mathbf{x}_{12}^H. \tag{15}$$

The matrix inversion is computed only once for a particular m, while the rest of the matrix inverses are generated using the above recursive steps. Therefore, the total complexity is reduced to $O(NW^2)$.

4) Compute $\mathbf{h}'_k$ *for each user* k

The beam domain channel $\mathbf{h}'_k$ of user k can be computed empirically as

$$\mathbf{h}'_k \approx \frac{1}{L}\mathbf{Y}'\mathbf{t}_k^*, \tag{16}$$

while the approximation error vanishes for large L and SNR. The complexity of computing $\mathbf{h}'_k$ for all users is $O(NLK)$. Finally, the beam domain channels of size W for each sliding window m are constructed as

$$\mathbf{u}_{k,m} = \mathbf{h}'_k[m : m+W-1]. \tag{17}$$

5) Search the index m *with Minimum MSE*

In this step, the window m providing the minimum mean square error (MSE) is found for each user k. The MSE criterion for each index pair (k, m) is given as [7]

$$MSE_{k,m} = 1 - \mathbf{u}_{k,m}^H\mathbf{R}_m^{-1}\mathbf{u}_{k,m}, \tag{18}$$

and the optimum window index m for each user k is found as

$$m_k^* = \arg\min_m MSE_{k,m}. \tag{19}$$

The complexity of the entire search process is $O(\beta N^2W^2)$. Finally, the optimum local LMMSE weights for each user k in the beamspace is given as

$$\mathbf{w}_k = \mathbf{u}_{k,m_k^*}^H\mathbf{R}_{m_k^*}^{-1}. \tag{20}$$

B. Beamspace Local LMMSE Beamforming process

Given the computed beamspace weights computed in the previous subsection, the detection of data symbols from (2) simply consists of two steps:

1) Calculate the DFT $\mathbf{y}'$ *for the received vector* $\mathbf{y}$

The complexity of this step is $O(N \log N)$.

2) Calculate the estimate $\hat{x}_k = \mathbf{w}_k\mathbf{y}'[m_k^* : m_k^* + W - 1] \; \forall \; k$

The complexity of this step is $O(\beta WN)$.

IV. Window Size W Does Not Scale with N

In this section, we sketch an argument showing that the required window size W does not scale with the number of antenna elements N. Under a simplified model of user's spatial distribution as being uniform across the N FFT bins, the number of users falling into a typical window is a binomial random variable $X \sim Bin(K, W/N)$. For $\beta < 1$, the mean number of users falling in a window, which is $\nu = KW/N = \beta W$, is smaller than the available dimension W, which implies that linear interference suppression is expected to be successful for a window size W, where the choice of W depends only on β, and does not scale with N.

We can now obtain rules of thumb on the choice of W as a function of β and a target outage probability. For large N, K and fixed β, X tends to a Poisson with mean ν, with Chernoff bound on tail probabilities [8]

$$P(X \geq x) \leq \frac{e^{-\nu}(e\nu)^x}{x^x}, \tag{21}$$

Assuming that outage in a given window occurs if and only if the number of users $X > W$, the probability of outage in a window is given by $P[X > W]$. Assuming that all users falling in the window are in outage when this happens, the expected number of users in outage, after direct mathematical manipulation, is $E[XI_{X>W}] = \nu P[X \geq W] = \beta W P[X \geq W]$. Summing over the $N - W + 1$ distinct windows and dividing by the number of users K, we obtain

$$P(outage) = \frac{N - W + 1}{K} \beta W P[X \geq W] \leq W P[X \geq W]$$

Plugging in (21), we obtain upon simplification that

$$P(outage) \leq W \left(\beta e^{1-\beta}\right)^W \tag{22}$$

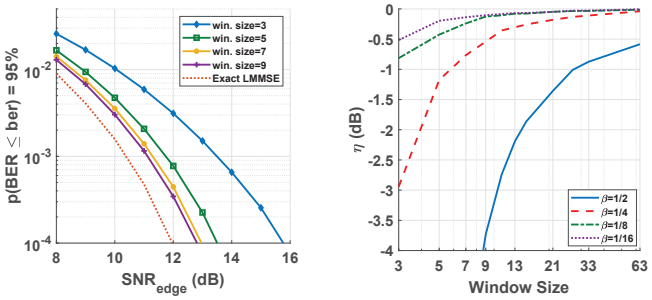
Given a 5% outage target, the window size computed using the above formula would be 2, 4, 8, 34 for load factors 1/16, 1/8, 1/4, 1/2, respectively, which are remarkably close to those obtained by simulations in Section V.

V. Numerical Results

We consider the system setup in Fig. 2 with number of antennas fixed at $N = 256$ for all numerical experiments. The field of view for the sector is restricted to $-\pi/3 \leq \theta \leq \pi/3$. The users are uniformly distributed inside a region bordered by a minimum and a maximum distance away from the base station, $R_{\text{min}} = 5$ m and $R_{\text{max}} = 100$ m, respectively. While the user terminals are placed randomly in our simulations, we enforce a minimum separation in spatial frequency between any two users in order not to incur excessive interference, arbitrarily choosing it as half the 3 dB beamwidth: $\Delta\Omega_{\text{min}} = \frac{2.783}{N}$ [9]. BW_{3dB} in Fig. 2 stands for the 3 dB beamwidth. We assume that users with similar spatial frequency can be served in different time or frequency resource blocks.

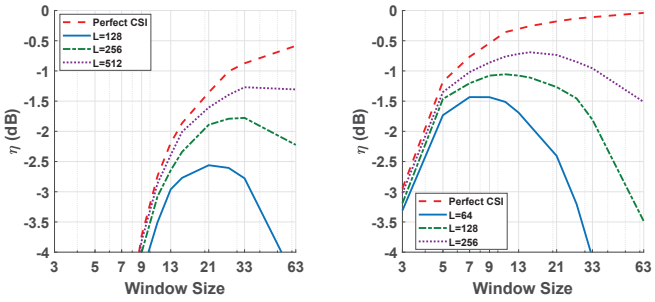
We measure link quality by the outage probability at a target uncoded BER of 10^{-3} for QPSK which requires SNR of 9.7 dB for a SISO AWGN link. This becomes the target SINR at the output of the multiuser detector for an edge user. No power control is deployed. The efficiency of the proposed beam space local LMMSE is defined as the ratio between the SNR_{edge} required to attain the link quality using standard LMMSE, relative to that required with local LMMSE:

$$\eta = \frac{SNR_{edge}(LMMSE)}{SNR_{edge}(local)} \tag{23}$$



(a) $\beta = 1/4$ (b) Target BER 10^{-3}

Figure 5: (a) BER achieved by at least 95% of the users for different W. (b) Edge user η with $\beta = \{1/2, 1/4, 1/8, 1/16\}$



(a) $\beta = 1/2$ (b) $\beta = 1/4$

Figure 6: Local LMMSE with implicit channel estimation.

Assuming perfect CSI, Fig. 5(a) shows the BER achieved by at least 95% of the users for different window sizes W and with load factor $\beta = 1/4$. Fig. 5(b) illustrates the efficiency η of the beamspace local LMMSE where the edge user SNR is adjusted such that at least 95% of the users achieve BER of 10^{-3}. In order to incur loss of only 1 dB or less in performance, a window size of 2, 3, 7, and 31 should be applied for load factors 1/16, 1/8, 1/4, and 1/2, respectively.

Figs. 6(a) and 6(b) plot the performance of the local LMMSE with implicit channel estimation as a function of training samples L, for $\beta = 1/2$ and $\beta = 1/4$, respectively. The training sequences are constructed from the Hadamard matrix such that user sequences are mutually orthogonal. There is a clear trade-off between the window size W and performance of the beamspace local LMMSE receiver (20). The optimum window size depends on both L and β. A larger window W provides an advantage in terms of the degrees of freedom available for suppressing inter-user interference in $\mathbf{R}_{m_k^*}$; see analysis in Section IV. However, at the same time, the estimation quality of outermost elements in $\mathbf{u}_{k,m_k^*}$ quickly decays as W is increased due to the sinc-like shape of the DFT beams. Thus, for finite L, the performance of (20) begins to deteriorate if W is made too large.

For a single user, local LMMSE approximates spatial

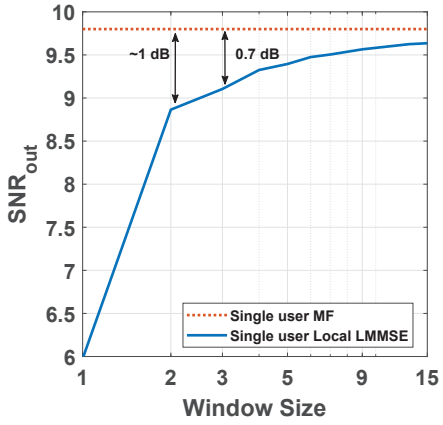
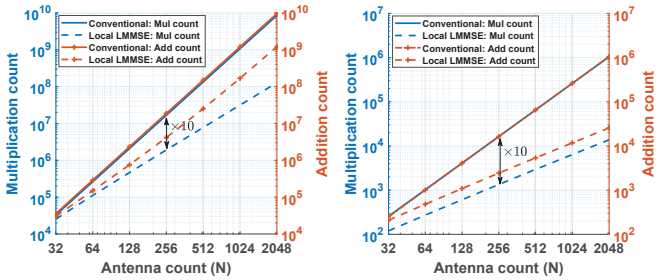


Figure 7: Local LMMSE vs. spatial MF for a single user.



(a) Acquisition complexity (b) Beamforming complexity

Figure 8: Complexity comparison of beamspace local LMMSE and conventional LMMSE.

matched filtering. Fig. 7 shows the worst-case loss in the output SNR of local LMMSE compared to the MF for a single user, which arises due to off-grid effects: with probability one, a user's location in beamspace is not aligned with the DFT bins. As shown, $W = 3$ is enough to collect most of the energy.

Fig. 8 compares, for $\beta = 1/4$, the complexity of conventional LMMSE and beamspace local LMMSE, resulting in 10-fold complexity reduction for $N = 256$ with local LMMSE for both weights acquisition (Fig. 8 (a)) and beamforming (Fig. 8 (b)).

VI. Conclusion

We have shown that, for the sparse spatial channels typical of mmWave bands, we can scale up the number of antennas and users without scaling the dimension of the signal subspace required to demodulate a given user. Thus, once we incur the $O(N \log_2 N)$ complexity of performing a spatial DFT, we can significantly reduce the complexity of multiuser detection: for example, the beamspace local LMMSE approach studied here achieves a ten-fold reduction in complexity compared to conventional LMMSE for the range of system parameters considered here. We have also shown how an adaptive implementation of this approach can be extended to provide implicit channel estimation. It is worth mentioning complementary research [10], [11] that indicates that analog frontend issues such as RF/baseband nonlinearities, analog-to-digital conversion and phase noise, are not fundamental bottlenecks for all-digital architectures. Taken together with the promising results in this paper, this motivates a concerted effort in designing scalable digital backends for mmWave massive MIMO.

Acknowledgment

This work was supported in part by the Semiconductor Research Corporation (SRC) under the JUMP program (2018-JU-2778) and by DARPA (HR0011-18-3-0004). Use was made of the computational facilities administered by the Center for Scientific Computing at the CNSI and MRL (an NSF MRSEC; DMR-1720256) and purchased through NSF CNS-1725797.

References

[1] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct 2013.

[2] J. Nam, A. Adhikary, J. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[3] A. Padmanabhan and A. Tölli, "An iterative approach for inter-group interference management in two-stage precoder design," in *Proc. IEEE Globecom 2018, Abu Dhabi, UAE*, Dec. 2018.

[4] T. Takahashi, A. Tölli, S. Ibi, and S. Sampei, "Layered belief propagation for low-complexity large MIMO detection based on statistical beams," in *Proc. IEEE International Conference on Communications 2019, Shanghai, China*, May. 2019.

[5] A. Krishnamoorthy and D. Menon, "Matrix inversion using Cholesky decomposition," in *2013 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. IEEE, 2013, pp. 70–72.

[6] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.

[7] U. Madhow, "MMSE interference suppression for timing acquisition and demodulation in direct-sequence CDMA systems," *IEEE Transactions on Communications*, vol. 46, no. 8, pp. 1065–1075, 1998.

[8] M. Mitzenmacher and E. Upfal, *Probability and computing: Randomized algorithms and probabilistic analysis*. Cambridge university press, 2005.

[9] C. A. Balanis, *Antenna Theory: Analysis and Design*. New York, NY, USA: Wiley-Interscience, 2005.

[10] M. Abdelghany, A. A. Farid, U. Madhow, and M. J. W. Rodwell, "Towards all-digital mmWave massive MIMO: Designing around non-linearities," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, Oct 2018, pp. 1552–1557.

[11] M. E. Rasekh, M. Abdelghany, U. Madhow, and M. J. W. Rodwell, "Phase noise analysis for mmWave massive MIMO: a design framework for scaling via tiled architectures," in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, Mar 2019.