Frequency-Domain Interpolation of the Zero-Forcing Matrix in Massive MIMO-OFDM

Salil Kashyap, Christopher Mollén, Emil Björnson and Erik G Larsson

Frequency-Domain Interpolation of the Zero-Forcing Matrix in Massive MIMO-OFDM

Salil Kashyap, Christopher Mollén, Emil Björnson, and Erik G. Larsson
Dept. of Electrical Eng. (ISY), Linköping University, 581 83 Linköping, Sweden
Email: {salil.kashyap, christopher.mollen, emil.bjornson, erik.g.larsson}@liu.se

***Abstract*—We consider massive multiple input multiple output (MIMO) systems with orthogonal frequency division multiplexing (OFDM) that use zero-forcing (ZF) to combat interference. To perform ZF, large dimensional pseudo-inverses have to be computed. In this paper, we propose a discrete Fourier transform (DFT)-interpolation-based technique where substantially fewer ZF matrix computations have to be done with very little deterioration in data rate compared to computing an exact ZF matrix for every subcarrier. We claim that it is enough to compute the ZF matrix at $L(\ll N)$ selected subcarriers where L is the number of resolvable multipaths and N is the total number of subcarriers and then interpolate. The proposed technique exploits the fact that in the massive MIMO regime, the ZF impulse response consists of L dominant components. We benchmark the proposed method against full inversion, piecewise constant and linear interpolation methods and show that the proposed method achieves a good tradeoff between performance and complexity.**

***Index Terms*—Massive MIMO, interpolation, zero-forcing**

I. Introduction

Massive MIMO systems have emerged as a leading 5G wireless communications technology where the base station (BS) uses an antenna array with a few hundred antenna elements to communicate with tens of users over the same time-frequency resource [1]. Orders of magnitude higher data rates and energy efficiency than contemporary wireless systems can be delivered. In this paper, we focus on techniques to reduce the computational complexity of detection and precoding in massive MIMO-OFDM systems. We consider systems that suppress interference using ZF where large-dimensional pseudo-inverses need to be computed and we ask the following question: *How often do we need to compute the ZF pseudo-inverse across subcarriers and then interpolate to obtain ZF matrices over all the subcarriers without incurring a noticeable loss in the ergodic rate?* Note that the same ZF matrix can be used for uplink detection and downlink precoding. However, for notational convenience, we discuss the former case in this paper.

A. Contributions

1) We propose a DFT-interpolation based low complexity ZF matrix computation technique.[1] We claim and show numerically that in the massive MIMO regime it is enough to compute the ZF matrix at $L(\ll N)$ equally spaced subcarriers, with L being the number of resolvable multipaths, N the total number of subcarriers, and then DFT-interpolate to obtain the detection/precoding matrices at all the N subcarriers. This is because in the massive MIMO regime, the channel of the desired user is approximately orthogonal to the space spanned by the channels of the interfering users. Thus, in this regime, ZF has an impulse response of length L. Furthermore, the empirical distribution of the singular values of the ZF matrix converges to the same deterministic limiting distribution across all subcarriers.
2) We derive a new expression for the achievable uplink ergodic rate with imperfect channel state information (CSI) for the proposed technique.
3) We compare the performance and complexity of the proposed technique against different ZF implementations, namely full inversion, piecewise constant, and linear interpolation.

B. Related Literature

In [2], the authors considered MIMO-OFDM systems with equal number of transmit and receive antennas and presented algorithms to compute inversion for square matrices by separately interpolating the adjoint and the determinant. The authors in [3], [4] proposed algorithms to compute $\mathbf{QR}$ decomposition at only a few select subcarriers and then determining the $\mathbf{Q}$ and $\mathbf{R}$ matrices for the remaining subcarriers by interpolation. In [5], the authors considered the interpolation of the inverse of square matrices and claimed that the power of the polynomial coefficients of the adjoint of a channel transfer function matrix can be well approximated by a Gaussian function. They developed methods to estimate the parameters of this Gaussian approximation function.

In contrast, we are particularly interested in non-square channel matrices and focus on direct interpolation of the pseudo-inverse itself by exploiting the fact that in the massive MIMO regime, the ZF impulse response is of length L.

II. System Model

We consider the uplink of a single-cell massive MIMO-OFDM system, where the bandwidth is divided into N orthogonal subcarriers. The BS is equipped with an array of

This research is funded by the European Union Seventh Framework Programme under grant agreement number ICT-619086 (MAMMOET).

[1] We do not consider interpolation of a maximum ratio (MR) filter, as interpolating an MR filter is the same as interpolating the channel matrix and does not reduce computational complexity.

978-1-5090-1749-2/16/$31.00

M antennas and there are K single-antenna users in the cell. The channel from the k^{th} user to the m^{th} antenna at the BS is denoted by $\tilde{\boldsymbol{g}}_k^m = [\tilde{g}_k^m[0] \ \tilde{g}_k^m[1] \ \cdots \ \tilde{g}_k^m[L-1]]^T$, where L is the number of resolvable multipaths, $\tilde{g}_k^m[i]$ consists of both small scale fading and distance-dependent path loss of the k^{th} user. We assume that the path loss from a user is the same to all the antennas at the BS. This assumption is justified because the size of a co-located antenna array is much smaller than the distance between the users and the BS. Furthermore, we assume Rayleigh fading. Therefore, $\tilde{\boldsymbol{g}}_k^m \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_k)$, where $\mathbf{\Lambda}_k$ is a diagonal matrix with the diagonal representing the channel power delay profile (PDP) of the k^{th} user that includes the path loss as well.

A. Uplink Pilot Signaling and Channel Estimation

The frequency-domain signal $\boldsymbol{y}_m \in \mathbb{C}^{N_p \times 1}$ received at the m^{th} antenna of the BS during uplink pilot signaling is

$$\boldsymbol{y}_m = \sum_{i=1}^{K} \sqrt{p_i^t} \mathbf{\Upsilon}_i^t \mathbf{\Omega}_r \tilde{\boldsymbol{g}}_i^m + \boldsymbol{w}_m, \tag{1}$$

where p_i^t is the pilot training power per subcarrier of the i^{th} user, $\mathbf{\Upsilon}_i^t \in \mathbb{C}^{N_p \times N_p}$ is a diagonal matrix with the N_p-length pilot sequence $\boldsymbol{x}_i^t$ corresponding to user i, $\mathbf{\Omega}_r \in \mathbb{C}^{N_p \times L}$ consists of the first L columns and N_p rows of the N-point discrete Fourier transform (DFT) matrix $\mathbf{\Omega}$ where $[\mathbf{\Omega}]_{m,n} = e^{-j2\pi(m-1)(n-1)/N}$. These rows correspond to the set of subcarriers on which the N_p pilots are sent. The noise vector at the m^{th} antenna of the BS is denoted by $\boldsymbol{w}_m$. Furthermore, $\boldsymbol{w}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_p})$. If the pilot sequences are chosen such that[2] $\mathbf{\Omega}_r^{\text{H}} \mathbf{\Upsilon}_k^{t^{\text{H}}} \mathbf{\Upsilon}_i^t \mathbf{\Omega}_r = N_p \mathbf{I}_L \delta_{ki}$, where $\delta_{ki} = 1$ if $k = i$, then a sufficient statistic for estimating $\tilde{\boldsymbol{g}}_k^m$ is

$$\tilde{\boldsymbol{y}}_m = \frac{1}{\sqrt{N_p}} \mathbf{\Omega}_r^{\text{H}} \mathbf{\Upsilon}_k^{t^{\text{H}}} \boldsymbol{y}_m = \sqrt{p_k^t N_p} \tilde{\boldsymbol{g}}_k^m + \tilde{\boldsymbol{w}}_m, \tag{2}$$

where $\tilde{\boldsymbol{w}}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$. Therefore, based on $\tilde{\boldsymbol{y}}_m$, the minimum mean square error (MMSE) estimate of the time-domain channel $\tilde{\boldsymbol{g}}_k^m$ from the k^{th} user to the m^{th} antenna at the BS is

$$\hat{\tilde{\boldsymbol{g}}}_k^m = \sqrt{p_k^t N_p} \mathbf{\Lambda}_k \left(p_k^t N_p \mathbf{\Lambda}_k + I_L\right)^{-1} \tilde{\boldsymbol{y}}_m. \tag{3}$$

B. Uplink Data Transmission

The data signal $\boldsymbol{y}(s) \in \mathbb{C}^{M \times 1}$ received on the uplink over the s^{th} subcarrier is given by

$$\boldsymbol{y}(s) = \mathbf{G}(s) \mathbf{\Phi}_d^{1/2}(s) \boldsymbol{x}(s) + \boldsymbol{w}(s), \tag{4}$$

where $\mathbf{G}(s) \in \mathbb{C}^{M \times K}$ denotes the frequency-domain channel matrix over the s^{th} subcarrier such that $\mathbf{G}(s) = [\boldsymbol{g}_1(s) \ldots \boldsymbol{g}_K(s)]$ and $\boldsymbol{g}_k(s) \in \mathbb{C}^{M \times 1}$ is the frequency-domain channel vector of the k^{th} user over the s^{th} subcarrier. Furthermore, $[\mathbf{G}(s)]_{m,k} = G_k^m(s) = \boldsymbol{\omega}_s^H \tilde{\boldsymbol{g}}_k^m$, where $\boldsymbol{\omega}_s^H$ denotes the s^{th} row consisting of only the first L columns of the N-point DFT matrix $\mathbf{\Omega}$. Also, $\mathbf{\Phi}_d(s)$ is a $K \times K$ diagonal matrix of the data power per subcarrier of the K users such that $[\mathbf{\Phi}_d(s)]_{k,k} = p_k^d$. The data vector of the K users over the s^{th} subcarrier is denoted by $\boldsymbol{x}(s)$ and the noise vector at the BS over the s^{th} subcarrier is denoted by $\boldsymbol{w}(s)$. Furthermore, $\boldsymbol{x}(s) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$ and $\boldsymbol{w}(s) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$.

[2]To ensure orthogonality among pilot sequences of different users, it is necessary to have $N_p \geq KL$.

III. Uplink Ergodic Rate Analysis

We let the detector matrix $\hat{\mathbf{A}}(s)$ be an $M \times K$ matrix which depends on the estimated frequency-domain channel matrix and on the choice of the detection method. The received vector on the s^{th} subcarrier after using the detector is given by

$$\boldsymbol{r}(s) = \hat{\mathbf{A}}^{\text{H}}(s) \boldsymbol{y}(s) = \hat{\mathbf{A}}^{\text{H}}(s) \mathbf{G}(s) \mathbf{\Phi}_d^{1/2}(s) \boldsymbol{x}(s) + \hat{\mathbf{A}}^{\text{H}}(s) \boldsymbol{w}(s). \tag{5}$$

Thus, the k^{th} element of $\boldsymbol{r}(s)$ is

$$r_k(s) = \sqrt{p_k^d} \hat{\boldsymbol{a}}_k^H(s) \boldsymbol{g}_k(s) x_k(s) + \sum_{i=1, i \neq k}^{K} \sqrt{p_i^d} \hat{\boldsymbol{a}}_k^H(s) \boldsymbol{g}_i(s) x_i(s) + \hat{\boldsymbol{a}}_k^H(s) \boldsymbol{w}(s), \tag{6}$$

where $\hat{\boldsymbol{a}}_k(s) \in \mathbb{C}^{M \times 1}$ is the column of $\hat{\mathbf{A}}(s)$ corresponding to the k^{th} user and is a function of the estimated channel.

Note that the MMSE estimate of $\boldsymbol{g}_k(s)$ is $\hat{\boldsymbol{g}}_k(s) = \boldsymbol{g}_k(s) - \boldsymbol{e}_k(s)$, where $\boldsymbol{e}_k(s) \in \mathbb{C}^{M \times 1}$ is the estimation error vector over the s^{th} subcarrier that is independent of $\hat{\boldsymbol{g}}_k(s)$. Furthermore, the m^{th} entry of $\boldsymbol{e}_k(s)$ is given by

$$e_k^m(s) = \boldsymbol{\omega}_s^H \tilde{\boldsymbol{g}}_k^m - \sqrt{p_k^t N_p} \boldsymbol{\omega}_s^H \mathbf{\Psi}_k \tilde{\boldsymbol{g}}_k^m - \boldsymbol{\omega}_s^H \mathbf{\Psi}_k \tilde{\boldsymbol{w}}_m, \tag{7}$$

for all $m = 1, \ldots, M$, where $\tilde{\boldsymbol{w}}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$ and is independent of $\tilde{\boldsymbol{g}}_k^m$ and $\mathbf{\Psi}_k = \sqrt{p_k^t N_p} \mathbf{\Lambda}_k \left(p_k^t N_p \mathbf{\Lambda}_k + \mathbf{I}_L\right)^{-1}$.

Thus, we can rewrite (6) as

$$r_k(s) = \sqrt{p_k^d} \hat{\boldsymbol{a}}_k^H(s) (\hat{\boldsymbol{g}}_k(s) + \boldsymbol{e}_k(s)) x_k(s) + \sum_{i=1, i \neq k}^{K} \sqrt{p_i^d} \hat{\boldsymbol{a}}_k^H(s) (\hat{\boldsymbol{g}}_i(s) + \boldsymbol{e}_i(s)) x_i(s) + \hat{\boldsymbol{a}}_k^H(s) \boldsymbol{w}(s). \tag{8}$$

Therefore, an achievable uplink ergodic rate for the k^{th} user over the s^{th} subcarrier with estimated CSI is given by (9). This is a lower bound on the ergodic capacity obtained using the methodology in [6] and holds for any choice of the detector matrix $\hat{\mathbf{A}}(s)$.

A. Proposed DFT Interpolation Based ZF Detector

The ZF detector over subcarrier s and with imperfect CSI is $\hat{\mathbf{G}}(s) \left(\hat{\mathbf{G}}(s)^H \hat{\mathbf{G}}(s)\right)^{-1}$ where $[\hat{\mathbf{G}}(s)]_{m,k} = \boldsymbol{\omega}_s^H \hat{\tilde{\boldsymbol{g}}}_k^m$. Let L_0 denote the number of subcarriers where the ZF matrix is computed. The proposed DFT-interpolation of ZF matrices involves the following steps:

1) L_0 equally spaced ZF matrices $\hat{\mathbf{G}}(s)(\hat{\mathbf{G}}(s)^H \hat{\mathbf{G}}(s))^{-1}$ of dimension $M \times K$ are computed at $s = 1, N/L_0 + 1, \ldots, (L-1)N/L_0 + 1$. For each m and k as illustrated in Fig. 1, an L_0-length vector $\boldsymbol{u}$ is obtained by picking the $(m,k)^{\text{th}}$ entry of each of these L_0 matrices.

$$R_k(s)=\mathbb{E}\left[\log_2\left(1+\frac{\frac{p_k^d\left|\mathbb{E}\left(\hat{\boldsymbol{a}}_k^H(s)(\hat{\boldsymbol{g}}_k(s)+\boldsymbol{e}_k(s))\middle|\hat{\boldsymbol{g}}_k(s)\ \forall\ k,s\right)\right|^2}{\|\hat{\boldsymbol{a}}_k^H(s)\|^2}}{\frac{\sum_{i=1}^{K}p_i^d\mathbb{E}\left(\left|\hat{\boldsymbol{a}}_k^H(s)(\hat{\boldsymbol{g}}_i(s)+\boldsymbol{e}_i(s))\right|^2\middle|\hat{\boldsymbol{g}}_k(s)\ \forall\ k,s\right)}{\|\hat{\boldsymbol{a}}_k^H(s)\|^2}-\frac{p_k^d\left|\mathbb{E}\left(\hat{\boldsymbol{a}}_k^H(s)(\hat{\boldsymbol{g}}_k(s)+\boldsymbol{e}_k(s))\middle|\hat{\boldsymbol{g}}_k(s)\ \forall\ k,s\right)\right|^2}{\|\hat{\boldsymbol{a}}_k^H(s)\|^2}+1}\right)\right] \quad (9)$$

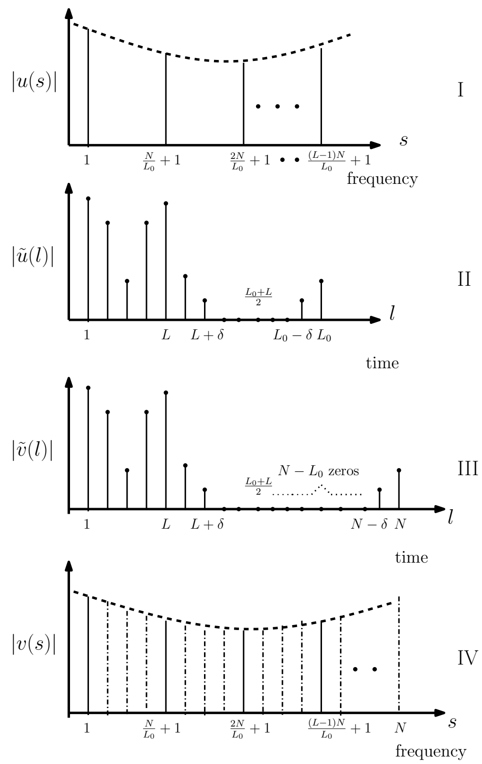


Fig. 1: DFT-interpolation: I. Compute L_0 equally spaced ZF matrices, II. Compute L_0-point IDFT, III. Pad $N-L_0$ zeros starting at $\frac{L_0+L}{2}$, IV. N-point DFT of the ZF impulse response.

2) An L_0 point inverse DFT (IDFT) of $\boldsymbol{u}$ is computed. Let $\tilde{\boldsymbol{u}}=\boldsymbol{\Omega}_{L0}^H\boldsymbol{u}$ denote the IDFT of $\boldsymbol{u}$, where $[\boldsymbol{\Omega}_{L0}^H]_{j,k}=\frac{1}{L_0}e^{j2\pi(j-1)(k-1)/L_0}$.
3) Next, $\tilde{\boldsymbol{u}}$ is padded with $N-L_0$ zeros starting at $(L_0+L)/2$ since the ZF power delay profile is symmetric around $L/2$. This helps recover the exact ZF impulse response $\tilde{\boldsymbol{v}}$, where $\tilde{\boldsymbol{v}}=\text{ZEROPAD}\{\tilde{\boldsymbol{u}}\}$.
4) The N-point DFT of $\tilde{\boldsymbol{v}}$ is computed which gives $\boldsymbol{v}=\boldsymbol{\Omega}\tilde{\boldsymbol{v}}$. This is repeated for each m and k to obtain N ZF matrices $\hat{\mathbf{G}}_{\text{DFT-intp}}(s)$ of dimension $M\times K$ each such that $[\hat{\mathbf{G}}_{\text{DFT-intp}}(s)]_{m,k}=\boldsymbol{v}(s)$.

Therefore, for this scheme and with imperfect CSI, the detector matrix is $\hat{\mathbf{A}}(s)=\hat{\mathbf{G}}_{\text{DFT-intp}}(s)$, where $\hat{\mathbf{G}}_{\text{DFT-intp}}(s)$ is the DFT-interpolated detector matrix corresponding to the s^{th} subcarrier. Note that for $L_0<L$, time-domain aliasing is severe and that results in a significant loss in performance. However, for $L_0\geq L$, DFT interpolation performs well because in the massive MIMO regime ($M,K\gg 1$, with $\frac{M}{K}>10$), the channel of the desired user is approximately orthogonal to the space spanned by the channels of the interfering users. Thus, in this regime MR and ZF are equivalent. Since, MR has an impulse response of length L, ZF will also have an impulse response of length L ($\delta\rightarrow 0$ in Fig. 1 as M increases).[3] Using (9) and applying standard results on Gaussian random vectors, the achievable uplink ergodic rate of the k^{th} user over the s^{th} subcarrier with estimated CSI and for the proposed ZF detector can be shown to simplify to (10).

B. Full Inversion Based ZF Detector

The ZF matrix is computed in a brute-force manner over each of the N subcarriers based on the estimated channel matrix, i.e., $L_0=N$. Therefore, the detector matrix $\hat{\mathbf{A}}(s)=\hat{\mathbf{G}}(s)(\hat{\mathbf{G}}(s)^H\hat{\mathbf{G}}(s))^{-1}$ and an expression for the achievable rate can be derived as in [6].

C. Piecewise Constant ZF Detector

The ZF matrices are computed at L_0 equally spaced subcarriers using the estimated channel matrix and the same detector is used to decode transmissions over a cluster of adjacent subcarriers. For example, the ZF detector computed over subcarrier $n=N/L_0+1$ is used to decode transmissions over some adjacent subcarrier s, where $s\in[n-N/(2L_0),n+N/(2L_0)]$. Therefore, for this scheme, the detector matrix to decode transmissions over the s^{th} subcarrier is $\hat{\mathbf{A}}(s)=\hat{\mathbf{G}}(n)(\hat{\mathbf{G}}(n)^H\hat{\mathbf{G}}(n))^{-1}$ and an expression for the achievable rate can be obtained by substituting $\hat{\mathbf{A}}(s)$ in (9).

D. Linear Interpolation Based ZF Detector

As before, L_0 ZF matrices are computed at equally spaced subcarriers. The linearly interpolated ZF matrix at any subcarrier s such that $1\leq s\leq\frac{N}{L_0}+1$ is given by $\hat{\mathbf{A}}(s)=\frac{L_0}{N}\left(\frac{N}{L_0}+1-s\right)\hat{\mathbf{A}}(1)+\frac{L_0(s-1)}{N}\hat{\mathbf{A}}\left(\frac{N}{L_0}+1\right)$, where $\hat{\mathbf{A}}(1)=\hat{\mathbf{G}}(1)(\hat{\mathbf{G}}(1)^H\hat{\mathbf{G}}(1))^{-1}$ and $\hat{\mathbf{A}}(N/L_0+1)=\hat{\mathbf{G}}(N/L_0+1)(\hat{\mathbf{G}}(N/L_0+1)^H\hat{\mathbf{G}}(N/L_0+1))^{-1}$ and an expression for achievable rate can be obtained by substituting $\hat{\mathbf{A}}(s)$ in (9).

Complexity Analysis: There are clearly multiple ways to reduce the number of pseudo-inverses that are computed, each attached with a certain additional interpolation complexity as given in Table I. We note that one complex multiplication involves 4 real multiplications and 2 real additions [7] and that a complex addition involves 2 real additions. We also note that a complex number can be represented by a real 2×2 matrix

[3]The ZF impulse response is a collection of N matrices of dimension $M\times K$ in the time-domain, and note that we interpolate on an element-by-element basis. Also, note that an arbitrary ZF impulse response can have N dominant taps, it is only in the massive MIMO regime that the ZF impulse response has just $L(\ll N)$ dominant taps.

$$R_k(s) = \mathbb{E}\left[\log_2\left(1+\frac{p_k^d|\hat{\boldsymbol{g}}_{k_{\text{DFT-intp}}}^H(s)\hat{\boldsymbol{g}}_k(s)|^2}{\sum\limits_{i=1,i\neq k}^{K} p_i^d|\hat{\boldsymbol{g}}_{k_{\text{DFT-intp}}}^H(s)\hat{\boldsymbol{g}}_i(s)|^2+\|\hat{\boldsymbol{g}}_{k_{\text{DFT-intp}}}^H(s)\|^2\left(1+\sum\limits_{i=1}^{K}p_i^d\sum_{l=1}^{L}\frac{[\boldsymbol{\Lambda}_i]_{l,l}}{1+p_i^tN_p[\boldsymbol{\Lambda}_i]_{l,l}}\right)}\right)\right] \quad (10)$$

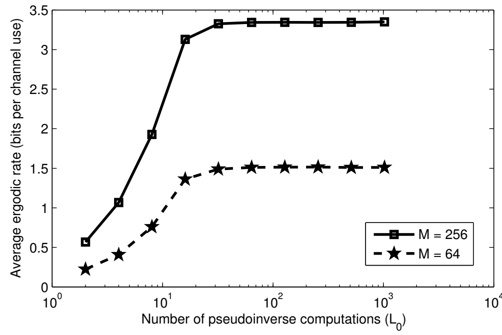
with a particular symmetric structure. We further exploit the result that the Cholesky factorization of a $2K\times 2K$ Hermitian matrix requires $\frac{8}{3}K^3$ real additions and multiplications and the forward and backward substitution methods to solve a triangular system of linear equations require $4MK^2$ operations each [8]. Thus, the cost of computing a pseudo-inverse at every subcarrier using Cholesky factorization of $\hat{\mathbf{G}}(s)^H\hat{\mathbf{G}}(s)$ followed by forward and backward substitution requires a total of $8MK^2+\frac{8}{3}K^3$ real additions and multiplications. We also know that for a length N complex data vector, its FFT using the split-radix algorithm requires $4N\log_2 N-6N+8$ real additions and multiplications [9]. Next, we compare the performance and complexity of these different interpolation schemes numerically.

IV. Numerical Results

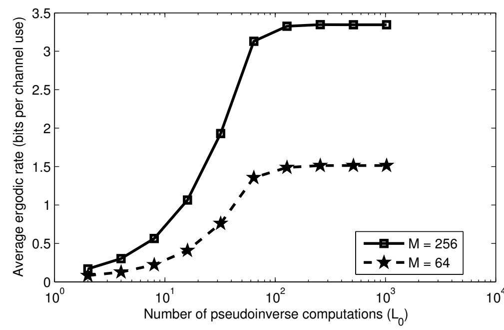
In this section, we present numerical results to investigate on how few subcarriers the ZF detector needs to be computed without incurring a significant rate loss compared to the full inversion method. For simplicity, we let the receive SNR $\rho = p_k^d\operatorname{Tr}(\boldsymbol{\Lambda}_k) = p_k^t\operatorname{Tr}(\boldsymbol{\Lambda}_k)$ be the same for all users, which for instance can be achieved by uplink power control. We consider a frequency-selective channel with uniform power delay profile and we take the number of pilot subcarriers $N_p = KL$.

Fig. 2a plots the average ergodic rate (sum rate of any user divided by the total number of subcarriers) for the proposed detector for $K = 16$ users and $L = 16$ channel taps as a function of the number of pseudo-inverse computations L_0 for two different values of M. It can be observed that there is a marginal loss in the average ergodic rate of about 9.8% for $M = 64$ and 6.6% for $M = 256$ when $L_0 = L$ compared to when $L_0 = N$. Fig. 2b plots the same for a more frequency-selective channel with $L = 64$. Similar conclusions are obtained from this case, thus illustrating the generality of the results. This is because in the massive MIMO regime ($M, K \gg 1$, with $\frac{M}{K} > 10$), the channel of the desired user is approximately orthogonal to the space spanned by the channels of the interfering users and the ZF impulse response is of length L, which is why it is enough to compute the pseudo-inverse at L selected subcarriers and then interpolate.

Fig. 3a plots the sum rate as a function of computational complexity for $L_0 = 16$, for different values of K and for all the four ZF detectors described in Section III. As expected, full inversion gives the highest sum rate, however, it also has the highest complexity. The DFT-interpolation based ZF detector gives a 12.5 % higher sum rate for $K = 16$ compared to piecewise constant at the cost of the increased complexity due to interpolation. It gives about 8 % lower sum rate for $K = 16$ compared to full inversion at significantly



(a) $L = 16$



(b) $L = 64$

Fig. 2: DFT-interpolation: Average ergodic rate vs. L_0 ($K = 16$, $N = 1024$, $\rho = -10$ dB)

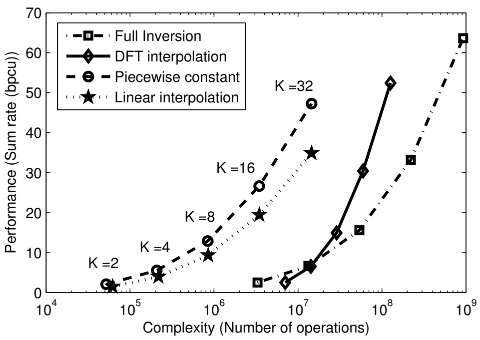
reduced complexity. It thus achieves a good tradeoff between complexity and performance for moderate values of K. Note that for $L_0 = L = 16$, linear interpolation performs poorer than piecewise constant because it is better to use the nearest ZF detector as in piecewise constant rather than taking a linear combination of two uncorrelated ZF detectors. Fig. 3b plots the same for $L_0 = 32$. It can be observed that the DFT-interpolation based ZF detector gives the same sum rate as one would obtain by full inversion at much reduced complexity.

Fig. 4a plots the ergodic rate of any user as a function of the subcarrier index for $L_0 = L = 4$. We observe that for the DFT-interpolation based ZF detector, the loss in ergodic rate is marginal when compared to full inversion. It also gives substantially better performance compared to piecewise constant and linear interpolation which fluctuate over the subcarriers.

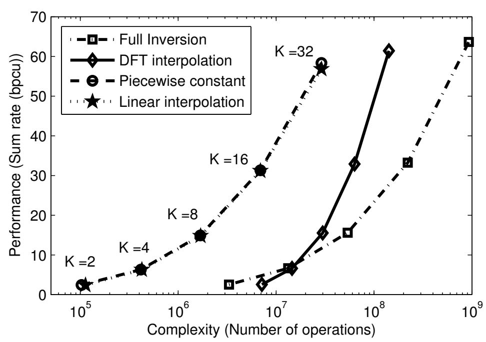
Fig. 4b plots the same for the case when $L_0 = 8 > L$. In this case, DFT-interpolation performs as well as full inversion. Also, linear interpolation based detector works better compared to piecewise constant. However, both of these give inferior performance as compared to DFT-interpolation.

TABLE I: Computational Complexity of Different ZF Detectors

Method	No. of pseudo-inverse computations	No. of computations in interpolation	Total no. of operations (Real additions + multiplications)
Full inversion	N	0	$(8MK^2 + \frac{8}{3}K^3)N$
DFT-interpolation (Proposed)	L_0	$\mathcal{O}(N \log N)$	$(8MK^2 + \frac{8}{3}K^3)L_0$ $+MK(4N \log_2 N - 6N + 8)$ $+MK(4L_0 \log_2 L_0 - 6L_0 + 8)$
Piecewise constant	L_0	0	$(8MK^2 + \frac{8}{3}K^3)L_0$
Linear interpolation	L_0	$N - L_0$ complex multiplications and $2(N - L_0)$ complex additions	$(8MK^2 + \frac{8}{3}K^3)L_0 + 10(N - L_0)$

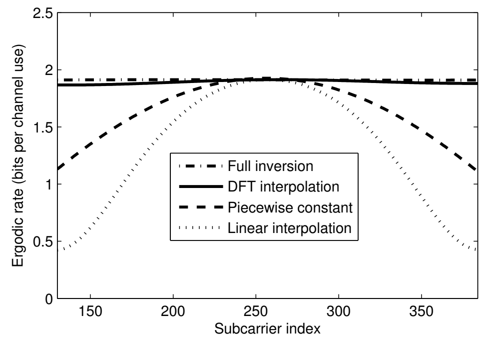


(a) $L_0 = 16$

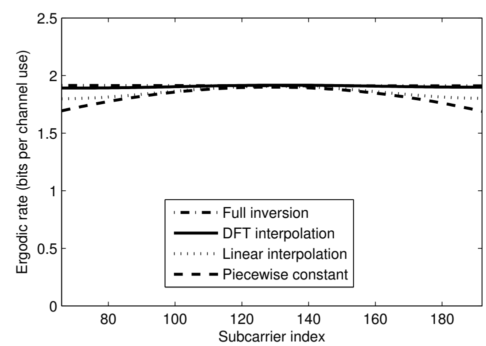


(b) $L_0 = 32$

Fig. 3: Benchmarking: Performance vs. complexity ($M = 100$, $L = 16$, $N = 1024$, $\rho = -10$ dB)



(a) $L_0 = 4$



(b) $L_0 = 8$

Fig. 4: Benchmarking: Ergodic rate vs. subcarrier index ($M = 128$, $K = 4$, $L = 4$, $N = 1024$, $\rho = -10$ dB)

V. Conclusions

We investigated on how few subcarriers do we need to compute the ZF matrix in a massive MIMO-OFDM system without incurring a visible rate loss compared to the full inversion scheme. We showed numerically that it is enough to compute the ZF matrix at $L(\ll N)$ equally spaced subcarriers and then DFT-interpolate to get the detector at all the N subcarriers. This is because in the massive MIMO regime, the ZF impulse response has L dominant components. We compared the proposed ZF implementation to full inversion, piecewise constant and linear interpolation and showed that it achieves a good tradeoff between complexity and performance.

References

[1] E. G. 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.

[2] M. Borgmann and H. Bölcskei, "Interpolation-based efficient matrix inversion for MIMO-OFDM receivers," in *Proc. Asilomar Conf. on Signals, Syst., and Comput.*, vol. 2, Nov 2004, pp. 1941–1947.

[3] D. Cescato and H. Bölcskei, "QR decomposition of Laurent polynomial matrices sampled on the unit circle," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4754–4761, Sep. 2010.

[4] ——, "Algorithms for interpolation-based QR decomposition in MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1719–1733, Apr. 2011.

[5] J. A. Zhang, X. Huang, H. Suzuki, and Z. Chen, "Gaussian approximation based interpolation for channel matrix inversion in MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1407–1417, Mar. 2013.

[6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[7] S. G. Johnson and M. Frigo, "A modified split-radix FFT with fewer arithmetic operations," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 111–119, Jan. 2007.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd Ed.)*. Baltimore, MD, USA: Johns Hopkins University Press, 1996.

[9] R. Yavne, "An economical method for calculating the discrete Fourier transform," in *Proc. AFIPS Fall Joint Computer Conf.*, vol. 33, 1968, pp. 115–125.