EXCESS CYCLIC PREFIX WINDOW OPTIMIZATION FOR HIGH DOPPLER MIMO OFDM CAPACITY

Kalyana Gopala , Dirk Slock

EURECOM, Sophia-Antipolis, France
Email: gopala,slock@eurecom.fr

ABSTRACT

In this work, the authors extend their previous work [1] on the design of sum rate maximizing beamformers for a MIMO OFDM scenario under high Doppler. We first justify the linear approximation for the channel variation over an orthogonal frequency division multiplexing (OFDM) symbol. Then, a modified power update for the optimized spatio-frequential precoder is presented. Of great interest is the exploitation of the Excess Cyclic Prefix (ExCP) at the receiver through windowing. We prove that the sum rate metric is a convex function of the window parameters and present a novel majorization based algorithm to design an optimal window. The convergence of the iterative approach is proved and the theory is validated via numerical simulations.

Index Terms— MIMO, ICI, Excess Cyclic Prefix, Beamforming, OFDM

1. INTRODUCTION

High Doppler encountered in high-speed train (HST) environments violates the orthogonality requirement for orthogonal frequency division multiplexing (OFDM), resulting in inter-carrier interference (ICI). The signal to interference plus noise ratio (SINR) analysis due to ICI can be found in ([2], [3]). It is known that multiple receive antennas in a single input multiple output (SIMO) scenario is very effective in canceling out the ICI (for example, see [4]). Another important tool in the mitigation of ICI is the exploitation of excess cyclic prefix (CP). With an appropriate window function, the excess CP at the receiver may be exploited to reduce the ICI. This is particularly relevant for HST scenarios where due to the close proximity of the Base station towers to the railway tracks, the delay spread expected is very minimal. The significance of using the excess CP and the Nyquist criterion can be found in [5]. For a single input single output (SISO) scenario, optimal window coefficients were derived to minimize the combined ICI and noise power in [6].

In [1], the authors of this paper considered a multiple input multiple output (MIMO) scenario at 2.4GHz and derived the channel sum rate in the presence of ICI caused by channel variation. The novel contributions in this paper are as follows:

- We directly optimize the stream powers at all subcarriers via an ICI-interference aware spatio-frequential water filling (as opposed to introducing intermediate power constraints at subcarriers in [1]).
- We prove the convexity of the sum rate metric with respect to the window parameters.
- We introduce a quadratic minorizer for the Gaussian sum rate that can be optimized iteratively with respect to (w.r.t.) the window coefficients to exploit the excess CP.
- The convergence of the entire beamformer design follows by design.

In the following discussions, a bold notation in small letters indicates a vector and bold notation with capital letters indicates a matrix. Unless otherwise specified, vector refers to a column vector. The operation "diag" has an interpretation identical to that in Matlab.

2. SYSTEM MODEL

Consider a multiple input multiple output (MIMO) system with N_t transmit antennas and N_r receive antennas. An OFDM framework is chosen with N subcarriers and sampling rate f_s. Thus, for every combination of transmit (Tx) and receive (Rx) antenna, the time domain channel at sample n of an OFDM symbol may be represented as

$$\mathbf{h}(n) = \mathbf{h}_0 + \mathbf{h}^{'}(n) \tag{1}$$

where $\mathbf{h}_0$ is of dimension $L \times 1$ and represents the average channel across the OFDM symbol. $\mathbf{h}^{'}$ is also of dimension $L \times 1$ and captures the time variation, has an average value of zero and is orthogonal to $\mathbf{h}_0$. With this formulation, the ICI contribution comes entirely from $\mathbf{h}^{'}(n)$ (see [4] for example). The length of the CP is considered to be greater than the channel delay spread by N_e samples. In what follows, without loss of generality, we take the CP length to be the same as that of the excess CP length. The total length of the OFDM symbol including the excess CP length is taken as $N_s = N + N_e$. It is also assumed that the receiver would take advantage of this excess CP through windowing. Let w_i be the window weights satisfying the Nyquist criterion, ([5]),

$$w_i + w_{N+i} = 1, \qquad i \in \{-N_e \cdots -1\}. \tag{2}$$

EURECOM's research is partially supported by its industrial members: ORANGE, BMW, ST Microelectronics, Symantec, SAP, Monaco Telecom, iABG, and by the projects DUPLEX (French ANR) and MASS-START (French FUI).

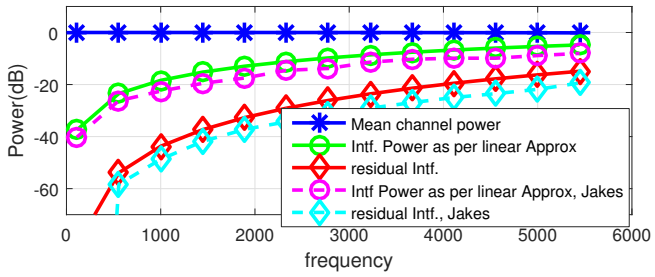

Fig. 1. Plot showing the residual error resulting from approximating the channel variation as linear at different Doppler frequencies.

The Nyquist criterion ensures that no new ICI components are introduced as a result of the windowing operation. To continue the analysis, we can approximate $\mathbf{h}^{'}(n)$ by a polynomial function. Fig. 1 shows the goodness of the assumption of the linear variation in the channel at different Doppler frequencies. In the plot, the Doppler is generated in two ways - as a single sinusoid or using the Jakes model [7]. The curve "mean channel power" plots the average channel power. The curves "interference power" plot the interference power as predicted by the linear model. Finally, the curves "residual interference" plot the error left un-modeled due to the linearity assumption. As expected, the single sinusoid model shows slightly more interference power than the Jakes model which considers all frequencies up to the max Doppler frequency. At 2.4GHz and 450kmph, the corresponding Doppler frequency is 1KHz. At this frequency, the residual error is around 40dB below the mean signal level implying that this approximation would hold good up to operating SNRs of 30dB. I.e, at this operating SNR, the error due to the approximation in channel variation is significantly lower than that of the AWGN noise floor. Indeed, if the operating SNR is lower, Fig. 1 shows that channel variation may be safely treated as linear across the OFDM symbol for even higher levels of Doppler.

Thus, for the duration of an OFDM symbol including the excess CP, (1) may be rewritten in terms of orthogonal basis functions for every Tx-Rx antenna pair as,

$$\begin{bmatrix} \mathbf{h}^T(-N_e) \\ \vdots \\ \mathbf{h}^T(N-1) \end{bmatrix} = \begin{bmatrix} 1 & (-N_e - \frac{N_s-1}{2}) \\ \vdots & \vdots \\ 1 & (N-1-\frac{N_s-1}{2}) \end{bmatrix} \begin{bmatrix} \mathbf{h}_0^T \\ \mathbf{h}_1^T \end{bmatrix} \quad (3)$$

where $\mathbf{h}_1$ is a constant across the OFDM symbol and captures the time variation per sample. The receiver output across all the receive antennas and subcarriers after the windowing and N-point fast Fourier transform (FFT) may be expressed as a column vector of length $N_r N$ (N_r received elements for each subcarrier). Proceeding similar to [1], we can arrive at the expression at each subcarrier k,

$$\mathbf{y}_k = \underbrace{\mathbf{H}_{0k}\mathbf{d}_k}_{\text{Signal term}} + \underbrace{\sum_{l=0,l\neq k}^{N-1} \mathbf{H}_{1l}\mathbf{d}_l \xi((l-k)_N) + \mathbf{v}_k}_{\text{ICI and noise terms}}. \quad (4)$$

$\mathbf{H}_{0k}$ (dimension $N_r \times N_t$) is the mean frequency domain channel observed at subcarrier k. The second term in equation (4) represents the ICI caused by linear time variation due to Doppler. $\mathbf{d}_k = [\mathbf{s}(kN_t+1)\cdots\mathbf{s}(kN_t+N_t-1)]^T$ is the $N_t \times 1$ vector of transmitted data symbols on the carrier k. $\mathbf{v}_k$ is the $N_r \times 1$ vector of AWGN (additive white Gaussian noise) noise observed at carrier index k, with variance $\mathbf{R}_{\mathbf{v}_k} = (e_k^T \mathbf{F}_N \mathbf{T}_{cp}^T \mathbf{D}_w \mathbf{D}_w^H \mathbf{T}_{cp} \mathbf{F}_N^H e_k)\mathbf{I}_{N_r}$. e_k is a column vector with 1 at the k^{th} element. $\mathbf{D}_b$ is a diagonal matrix representing the linear basis function of the time variation of the channel.

$$\mathbf{D}_w = \underbrace{\texttt{diag}([w_{-N_e}\ldots w_{-1}\ 1\ldots 1\ w_1\ldots w_{N_e}]^T)}_{\texttt{diag}(\mathbf{w})}$$

$$\mathbf{D}_b = \texttt{diag}\left(-(N_e - \frac{N_s-1}{2})\ldots(N-1-\frac{N_s-1}{2})\right).$$

$$\mathbf{T}_{cp} = \begin{bmatrix} \mathbf{0}_{N_e \times N-N_e} & \mathbf{I}_{N_e} \\ & \mathbf{I}_N \end{bmatrix}.$$

$\mathbf{\Xi} = \mathbf{F}_N \mathbf{T}_{cp}^T \mathbf{D}_w \mathbf{D}_b \mathbf{T}_{cp} \mathbf{F}_N^{-1}$ is a circulant matrix of dimension $N \times N$. As Ξ is circular, any element k,l of the matrix may be expressed as $\xi((l-k)_N)$ where the notation $()_N$ refers to modulo operation with respect to N. ξ would be the first row vector of $\mathbf{\Xi}$ and hence of length N. We constrain the window in such a manner that $\xi(0) = 0$ and this is factored in later during the window design in section 3.2.

Let P be the maximum sum power requirement across all the subcarriers and let P_i be the individual power at any subcarrier i such that $\sum_{i=0}^{N-1} P_i = P$. Let the transmit covariance matrix of subcarrier k be $\mathbf{Q}_k = \texttt{E}(\mathbf{d}_k \mathbf{d}_k^H)$ where $\texttt{E}(\cdot)$ is the expectation operator. Let $\mathbf{R}_{\bar{k}} = \mathbf{R}_{\mathbf{v}_k} + \sum_{l=0,l\neq k}^{N-1} |\xi((l-k)_N)|^2 \mathbf{H}_{1l}\mathbf{Q}_l\mathbf{H}_{1l}^H$. We are interested in determining the optimal $\mathbf{Q}_k$ and the window weights w_i such that the sum rate of the link is maximized under constraints,

$$\begin{aligned} \texttt{f}_0 : \max_{\mathbf{Q}_k,\mathbf{w}} \mathbf{C} &= \max_{\mathbf{Q}_k} \sum_{k=0}^{N-1} \log|\mathbf{I} + \mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H\mathbf{R}_{\bar{k}}^{-1}| \\ \texttt{subject to} \quad & \sum_{k=0}^{N-1} \texttt{tr}\{\mathbf{Q}_k\} \leq P, \quad \mathcal{V}\mathbf{w} = \mathbf{b} \end{aligned} \quad (5)$$

Here, matrix $\mathcal{V}$ captures the constraints on $\mathbf{w}$ and $\mathbf{b}$ is a column vector.

3. BEAMFORMER DESIGN

We employ the alternating (cyclic) minimization approach to alternately optimize the precoder design and window design. At the beginning of the iteration for the subcarrier i, let P_i be the power constraint, $\bar{\mathbf{Q}}_i$ be the current values of the precoder and w_i be the window values.

3.1. Transmit covariance matrix update

Our iterative optimization algorithm operates one subcarrier at a time. With a focus on subcarrier i, on the same lines as [8], the objective function $\texttt{f}_0$ may be rewritten as

$$\max_{\mathbf{Q}_i}\{\log|\mathbf{I} + \mathbf{H}_{0i}\mathbf{Q}_i\mathbf{H}_{0i}^H\mathbf{R}_{\bar{i}}^{-1}| + f_i(\mathbf{Q}_i, \mathbf{Q}_{-i})\} \quad (6)$$

where $f_i(\mathbf{Q}_i, \mathbf{Q}_{-i}) = \sum_{l \neq i} \log |\mathbf{I} + \mathbf{H}_{0l}\mathbf{Q}_l\mathbf{H}_{0l}^H\mathbf{R}_{\bar{l}}^{-1}|$. $\mathbf{Q}_{-i}$ refers to the transmit covariances of all the subcarriers except the i^{th}. It is shown in [8] ($Lemma$ 1) that $f_i(\mathbf{Q}_i, \mathbf{Q}_{-i})$ is convex in $\mathbf{Q}_i$. Thus, equation (6) is the sum of a concave and convex function and hence the overall sum rate is a non-convex function. As in [9], we replace the non-concave function above by its minorization which is concave.

$$\begin{aligned} \mathtt{f}_1 : \log|\mathbf{I} + \mathbf{H}_{0i}\mathbf{Q}_i\mathbf{H}_{0i}^H\mathbf{R}_{\bar{i}}^{-1}| - \mathtt{tr}\left\{\mathbf{B}_i(\mathbf{Q}_i - \bar{\mathbf{Q}}_i)\right\} + \\ f_i(\bar{\mathbf{Q}}_i, \bar{\mathbf{Q}}_{-i}) \quad \mathtt{subject\ \ to\ \ tr}\left\{\mathbf{Q}_i\right\} \leq \bar{P}_i \end{aligned} \tag{7}$$

$\bar{P}_i$ indicates the current value of P_i at any given stage of the algorithm. $\mathbf{B}_i = -\left[\frac{\partial f_i(\mathbf{Q}_i, \mathbf{Q}_{-i})}{\partial \mathbf{Q}_i}\right]^H$,

$$= \sum_{l \neq i} |\xi((l-i)_N)|^2 \mathbf{H}_{1i}\left\{\mathbf{R}_{\bar{l}}^{-1} - (\mathbf{R}_{\bar{l}} + \mathbf{H}_{0l}\mathbf{Q}_l\mathbf{H}_{0l}^H)^{-1}\right\}\mathbf{H}_{1i}^H$$

Let $\mathbf{A}_i = \mathbf{H}_{0i}^H\mathbf{R}_{\bar{i}}^{-1}\mathbf{H}_{0i}$, $\mathbf{Q}_i = \mathbf{V}_i\mathbf{\Lambda}_i\mathbf{V}_i^H$, and λ_{ij} be the j^{th} diagonal element of $\mathbf{\Lambda}_i$. The optimal solution to this sub-problem are the (normalized) generalized eigenvectors $\mathbf{V}_i$ (see [9])

$$\mathbf{A}_i\mathbf{V}_i = (\mathbf{B}_i + \mu\mathbf{I})\,\mathbf{V}_i\mathbf{\Sigma}_i \tag{8}$$

Let $\mathbf{V}_i^H\mathbf{A}_i\mathbf{V}_i = \mathbf{D}_{1i}$ and $\mathbf{V}_i^H\mathbf{B}_i\mathbf{V}_i = \mathbf{D}_{2i}$, where $\mathbf{D}_{1i}$, $\mathbf{D}_{2i}$ are diagonal matrices as $\mathbf{V}_i$ is the generalized eigenmatrix of $\mathbf{A}_i$, $\mathbf{B}_i + \mu\mathbf{I}$. Now optimizing the sum of equation (7) over all subcarriers after adding the power constraint, w.r.t. the powers Λ_i yields

$$\lambda_{ij} = \left[\frac{1}{\mathbf{D}_{2i}(j,j) + \mu} - \frac{1}{\mathbf{D}_{1i}(j,j)}\right]^+ \tag{9}$$

where $[x]^+$ indicates $\max(x, 0)$. The optimal Lagrange multiplier μ for the power constraint can now be determined using a bisection search as the λ_{ij} are monotonic in μ. Thus, the concave objective functions $\mathtt{f}_1$ can be solved alternatingly till the $\mathbf{Q}_i$ converge.

3.2. Optimization of the window parameters

Once Tx covariance matrices $\mathbf{Q}_i$ have been computed for all the subcarriers, we propose an iterative algorithm based on majorization [10] to determine the window parameters that optimize the metric,

$$\max_{\mathbf{w}} C = \sum_{k=0}^{N-1} \log|\mathbf{I} + \mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H\mathbf{R}_{\bar{k}}^{-1}|, \tag{10}$$

The window parameters, of course, are constrained to satisfy the Nyquist criterion. In addition, we impose that $\xi(0) = 0$.

Theorem 1. Sum rate in (10) is a convex function in $\Delta_w = \mathbf{w}\mathbf{w}^H$.

Proof. As $\xi(0) = 0$, $\mathbf{w}$ influences only the term $\mathbf{R}_{\bar{k}}^{-1}$. We make use of the following matrix rearrangements,

$$\begin{aligned} \mathbf{R}_{\mathbf{v}_k} &= (e_k^T\mathbf{F}_N\mathbf{T}_{cp}^T\mathbf{D}_w^H\mathbf{D}_w\mathbf{T}_{cp}\mathbf{F}_N^H e_k)\mathbf{I}_{N_r} \\ &= \mathbf{w}^H\mathbf{X}_1(k)\mathbf{X}_1^H(k)\mathbf{w}\mathbf{I}_{N_r} \\ &= \mathtt{tr}\{\mathbf{X}_1(k)\mathbf{X}_1^H(k)\mathbf{w}\mathbf{w}^H\}\mathbf{I}_{N_r}. \\ \xi((l-k)_N) &= \mathbf{e}_k^T\mathbf{F}_N\mathbf{T}_{cp}^T\partial\mathbf{D}_w\mathbf{D}_b\mathbf{T}_{cp}\mathbf{F}_N^{-1}\mathbf{e}_l \\ &= \mathbf{x}_2^T(k,l)\mathbf{w}. \end{aligned} \tag{11}$$

Here, $\mathbf{X}_1(k) = \mathtt{diag}(e_k^T\mathbf{F}_N\mathbf{T}_{cp}^T)$ and the column vector $\mathbf{x}_2(k,l) = \mathtt{diag}(\mathbf{D}_b\mathbf{T}_{cp}\mathbf{F}_N^{-1}\mathbf{e}_l\mathbf{e}_k^T\mathbf{F}_N\mathbf{T}_{cp}^T)$. ()* refers to taking the complex conjugate. Further, we have used the properties of the trace operator [11]. Thus, (10) may be rewritten as,

$$\begin{aligned} \mathbf{R}_{\bar{k}} = \mathtt{tr}\{\mathbf{X}_1(k)\mathbf{X}_1^H(k)\Delta_w^H\}\mathbf{I}_{N_r} + \\ \sum_{l=0, l\neq k}^{N-1} \mathtt{tr}\{\mathbf{x}_2^*(k,l)\mathbf{x}_2^T(k,l)\Delta_w^H\}\mathbf{H}_{1l}\mathbf{Q}_l\mathbf{H}_{1l}^H. \end{aligned} \tag{12}$$

Now, to show convexity, it is sufficient to show that $f(t) \triangleq C(\Delta_w^a + t\Delta_w^b)$ is convex w.r.t. $t \in [0,1]$ (see [8], $Lemma$ 1).

$$f(t) = \sum_{k=0}^{N-1} \log|\mathbf{I} + \mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}|, \tag{13}$$

Let $\mathbf{A}_k = \mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H$. Following the steps in [12] (see the three-step procedure therein),

$$\begin{aligned} \partial f(t) = -\mathtt{tr}\Big\{ \sum_{k=0}^{N-1} \left(\mathbf{I} + \mathbf{A}_k(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)\right)^{-1} \\ \mathbf{A}_k(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\partial t\mathbf{R}_{\bar{k}}^b(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\Big\}. \end{aligned} \tag{14}$$

$$\begin{aligned} \frac{\partial f(t)}{\partial t} &= -\mathtt{tr}\Big\{ \sum_{k=0}^{N-1} \left(\mathbf{I} + \mathbf{A}_k(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)\right)^{-1} \\ &\qquad \mathbf{A}_k(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\mathbf{R}_{\bar{k}}^b(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\Big\} \\ &= -\mathtt{tr}\Big\{ \sum_{k=0}^{N-1} \left(\mathbf{A}_k + \mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b\right)^{-1}\mathbf{A}_k(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\mathbf{R}_{\bar{k}}^b\Big\} \end{aligned} \tag{15}$$

Proceeding similarly to take the second derivative,

$$\begin{aligned} \frac{\partial^2 f(t)}{\partial t^2} = \mathtt{tr}\Big\{ \sum_{k=0}^{N-1} \Big(\left(\mathbf{A}_k + \mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b\right)^{-1} + \left(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b\right)^{-1}\Big) \\ \mathbf{R}_{\bar{k}}^b(\mathbf{A}_k + \mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\mathbf{A}_k(\mathbf{R}_{\bar{k}}^a + t\mathbf{R}_{\bar{k}}^b)^{-1}\mathbf{R}_{\bar{k}}^b\Big\} \geq 0. \end{aligned} \tag{16}$$

The second derivative involves the sum of the trace of positive semi definite matrices and hence results in a non-negative number indicating that the metric C is convex in Δ_w. □

To arrive at a stationary point of $\mathbf{w}$, we follow the iterative steps in minorization [10]. Let $\bar{\mathbf{w}}$ be the value of $\mathbf{w}$ at any given iteration. Construct a minorizer for the sum rate as follows,

$$\bar{C} + \mathtt{tr}\Big\{\frac{\partial C}{\partial \Delta_w}(\Delta_w - \Delta_{\bar{w}})\Big\}. \tag{17}$$

$\bar{C}$ refers to the value of C at $\Delta_w = \Delta_{\bar{w}}$. As the objective function has been proved to be convex already, (17) forms a touching tangent to this convex function at the point $\Delta_w = \Delta_{\bar{w}}$ and is hence a minorizer. At every step, the maximization of the minorizer results in an updated Δ_w that causes a non-decreasing change in C.

$$\partial C = -\sum_{k=0}^{N-1} \mathtt{tr}\{\mathbf{X}_1(k)\mathbf{X}_1^H(k)\Delta_w\}\mathtt{tr}\{\mathbf{X}_3(k)\}$$
$$-\sum_{k=0}^{N-1}\sum_{l\neq k} \mathtt{tr}\{\mathbf{x}_2^*(k,l)\mathbf{x}_2^T(k,l)\Delta_w\}\mathtt{tr}\{\mathbf{X}_4(k,l)\}.$$
$$\mathbf{X}_3(k) = \mathbf{R}_{\bar{k}}^{-1}(\mathbf{I}+\mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H\mathbf{R}_{\bar{k}}^{-1})^{-1}\mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H\mathbf{R}_{\bar{k}}^{-1}.$$
$$\mathbf{X}_4(k,l) = \mathbf{H}_{1l}\mathbf{Q}_l\mathbf{H}_{1l}^H\mathbf{R}_{\bar{k}}^{-1}(\mathbf{I}+\mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H\mathbf{R}_{\bar{k}}^{-1})^{-1}$$
$$\mathbf{H}_{0k}\mathbf{Q}_k\mathbf{H}_{0k}^H\mathbf{R}_{\bar{k}}^{-1},$$
$$\frac{\partial C}{\partial \Delta_w} = \mathbf{U} = -\sum_{k=0}^{N-1} \mathbf{X}_1(k)\mathbf{X}_1^H(k)\mathtt{tr}\{\mathbf{X}_3(k)\}$$
$$-\sum_{k=0}^{N-1}\sum_{l\neq k} \mathbf{x}_2^*(k,l)\mathbf{x}_2^T(k,l)\mathtt{tr}\{\mathbf{X}_4(k,l)\}$$

Note that $\mathbf{U}$ is a Hermitian matrix. The Nyquist conditions on the window together with the (single) additional constraint $\xi(0) = 0$ lead to a set of linear constraints of the form $\mathcal{V}\mathbf{w} = \mathbf{b}$. $\mathbf{b}$ is a column vector of all ones except the last entry which is zero. To impose the constraint that $\xi(0) = 0$, note from equation (11) that $\xi(0) = \mathbf{x}_2^T(k,k)\mathbf{w}$. The solution for the optimization of $\mathbf{w}$ is then given by,

$$\mathbf{w} = \mathbf{U}^{-1}\mathcal{V}^H(\mathcal{V}\mathbf{U}^{-1}\mathcal{V}^H)^{-1}\mathbf{b}. \quad (18)$$

3.3. Overall Algorithm and Convergence

The overall algorithm alternates between the transmit beamformer optimization and the window coefficient optimization. At every iteration of the transmit beamformer, a concave subproblem $\mathtt{f}_1$ is created and optimized based on the updated value of $\mathbf{Q}_i, \mathbf{Q}_{-i}$ from the last optimization in the alternating process. The window optimization is based on another minorizer. Hence, the overall algorithm is such that the sum rate metric update is non-decreasing at each step, and is upper bounded due to the noise, which ensures convergence.

4. NUMERICAL RESULTS

Consider a single user MIMO fading channel based on equation (3) and $N_t = N_r = 3$ with a signal to AWGN noise ratio of 20dB. For every Tx-Rx pair, FIR Rayleigh fading channels are generated independently with the power delay profile (PDP) as [0 -5 -5] in dB for $\mathbf{h}_0$ and $\mathbf{h}_1$. An OFDM system operating at 2.4GHz band is considered with 15KHz of subcarrier spacing. The entries of $\mathbf{h}_1$ are scaled such that the overall ICI power experienced at any receive antenna corresponds to a Doppler frequency shift of 450kmph. We consider $N = 64$ and $N_e = 16$ corresponding to a 25% excess CP. Fig. 2 plots the achievable sum rates for different configurations. The curve ("WF, No ExCP") gives the water-filling performance in the absence of cyclic prefix exploitation. This is the standard water-filling algorithm where $\mathbf{H}_1$ is naively taken as all zeros and no excess CP is exploited. Also given is a curve ("ICI aware WF,no ExCP") that does the iterative optimization of the transmit beamformer with the knowledge of

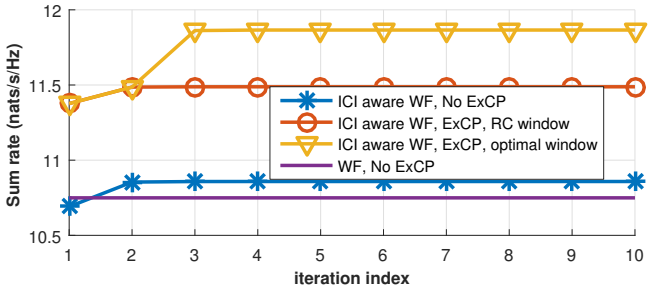

Fig. 2. Sum rate comparison with $N_t = 3$, $Nr = 3$, $N = 64$, $N_e = 16$ at a Doppler of 450Kmph and 20dB SNR.

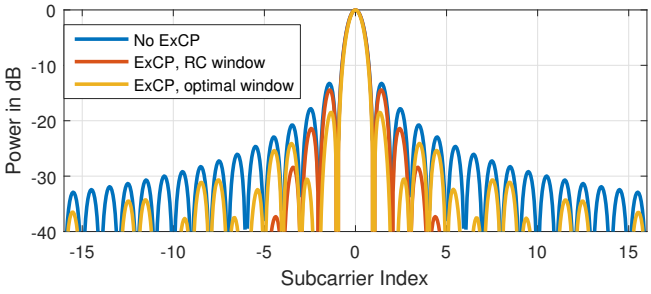

Fig. 3. Comparison of windows used to exploit excess CP

ICI, again in the absence of excess CP. The curve "ICI aware WF, ExCP RC window" shows the performance of the ICI aware transmit beamforming optimization that uses a raised cosine window (as in [13]) to exploit the excess CP. Finally, the curve ("ICI aware WF, ExCP optimal window") shows the performance with ICI aware transmit beamformer optimization and optimized window coefficients for excess CP. Fig. 3 gives the roll-off obtained for the optimized window in comparison with other windows. It is very clearly seen that the optimal window does a good side lobe reduction for the closest side lobes and does not over attenuate the farther side lobes (as done by the raised cosine window). Thus, the optimal window can be observed to strike a better balance in side lobe reduction compared to the raised cosine window. This is quite intuitive too and explains why the optimal window performs superior to the raised cosine window. In the simulations, we see that the iterations always exhibit a non-decreasing behavior in the sum rate as is predicted by the theory (section 3.3).

5. CONCLUSION

In this paper, we have extended our previous work on the water-filling problem for an OFDM system in the presence of ICI. We justified the linearity assumption on channel variation. Then, we proved that the sum rate metric is convex in terms of the window coefficients. Minorizers of the sum rate were constructed to iteratively derive the window weights. The ICI roll off for the optimally derived window is compared with other windows and the observations are intuitively appealing. While the window parameter optimization has been developed in the general context of MIMO Tx beamformer design, it is optimal in the context of ICI reduction for a single input multiple output (SIMO) scenario as well.

6. REFERENCES

[1] K. Gopala and D. Slock, "High Doppler MIMO OFDM capacity maximizing spatial transceivers exploiting excess cyclic prefix," in *Intern'l Symp. on Wireless Communication Systems (ISWCS)*, Sept 2016.

[2] M Faulkner, L.R Wilhelmsson, and J. Svensson, "Low-Complex ICI Cancellation for Improving Doppler Performance in OFDM Systems," in *IEEE Vehicular Technology Conference*, Sept 2006.

[3] Yuexing Peng, Wenbo Wang, and Young Il Kim, "Performance Analysis of OFDM System Over Time-Selective Fading Channels," in *IEEE Wireless Comm. and Networking Conference (WCNC)*, April 2009.

[4] Kalyana Gopala and Dirk Slock, "Doppler compensation and Beamforming for High Mobility OFDM transmissions in multipath," in *EAI International Conf. on Cognitive Radio Oriented Wireless Networks*, June 2016.

[5] S. H. Muller-Weinfurtner, "Optimum Nyquist windowing in OFDM receivers," *IEEE Trans. on Communications*, Mar 2001.

[6] C. Y. Ma, S. W. Liu, and C. C. Huang, "On Optimum Segment Combining Weight for ICI Self-Cancellation in OFDM Systems under Doubly Selective Fading Channels," in *IEEE Vehicular Technology Conference (VTC Spring)*, May 2012.

[7] William C. Jakes and Donald C. Cox, Eds., *Microwave Mobile Communications*, Wiley-IEEE Press, 1994.

[8] Seung-Jun Kim and G.B. Giannakis, "Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks," *IEEE Trans. on Info. Theory*, May 2011.

[9] K.Gopala and D.Slock, "MIMO OFDM Capacity Maximizing Beamforming for Large Doppler Scenarios," in *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2016.

[10] Petre Stoica and Y. Selen, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher," *IEEE Signal Proc. Magazine*, Jan 2004.

[11] K. B. Petersen and M. S. Pedersen, "The Matrix Cookbook," Nov 2012.

[12] A. Hjorungnes and D. Gesbert, "Complex-Valued Matrix Differentiation: Techniques and Key Results," *IEEE Trans. on Signal Processing*, June 2007.

[13] A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Low Complexity CFO Compensation in Uplink OFDMA Systems With Receiver Windowing," *IEEE Trans. on Signal Processing*, May 2015.