Optimal windowing in MIMO OFDM for network interference suppression

Optimal Windowing in MIMO OFDM for Network Interference Suppression

Andrew D. Copeland, Daniel W. Bliss, and Andrew L. McKellips
MIT Lincoln Laboratory, Lexington, MA 02420
{andrew.copeland, bliss, mckellips}@ll.mit.edu

***Abstract*—The combined use of orthogonal frequency-division multiplexing (OFDM) and wireless multiple-input multiple-output (MIMO) links is becoming common. For many applications, such as ad hoc wireless networks, interference is a significant issue. MIMO links allow for the spatial mitigation of interference; however, spatial interference mitigation when using OFDM modulation performs somewhat poorly because of the relatively limited interference spectral confinement of the Discrete Fourier Transform. Traditional windowing approaches that minimize leakage of the interference into neighboring carriers induce intercarrier interference. In this paper, a technique is developed for determining optimal windows that trade off interference suppression with intercarrier and interframe interference. This technique assumes channel statistics are known. Optimized windows for various levels of interference, delay spreads, frame length, and cyclic prefix lengths can be produced.**

I. Network Introduction

With the widespread deployment of IEEE 802.11N [1], [2], multiple-input multiple-output (MIMO) links [3], [4], [5] that employ orthogonal frequency-division multiplexing (OFDM) modulation [6], [7] are becoming common. While current implementations of MIMO do not exploit the innate capability of MIMO to spatially mitigate interference, theoretically spatial mitigation is possible [5].

Interference is of concern for ad hoc wireless networks because by their nature multiple links tend to operate in proximity to each other. Historically, networks mitigated interference by employing time-division multiple access (TDMA) or frequency-division multiple access (FDMA) [6], [7]. These multiple access schemes reduce the potential network throughput. While the spatial interference mitigation theoretically enabled by MIMO cannot completely alleviate the need for access schemes such as TDMA, spatial interference mitigation can increase spectral utilization of wireless networks [8], [9].

The advantages of MIMO's potential spatial interference mitigation is complicated by interaction with OFDM modulation. The Discrete Fourier Transform (DFT) employed by OFDM systems provides relatively poor spectral confinement. Consequently, energy at one frequency is spread across many frequency bins. In the case of frequency-selective fading, a single transmitted signal will occupy multiple eigenvalues in the receive covariance matrix at a given frequency bin because contributions from the actual frequency array response will be present in combination with array responses from nearby frequencies. When spatial nulling of interference is attempted, this growth in signal rank can adversely affect mitigation performance. For OFDM signals this is typically not a significant issue because orthogonality of carriers is maintained with the use of a cyclic prefix [6], [7].

The traditional solution to minimizing leakage of energy into adjacent frequency bins is to employ windows [10], [11]. However, windows can break the orthogonality of OFDM, inducing intercarrier interference (ICI) between neighboring carriers. This effect is illustrated in Figure 1. Windowing

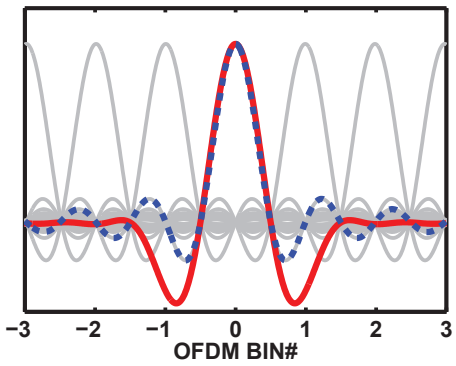

Fig. 1. **Windowing ICI** – Several unwindowed carriers depicted in gray. The center unwindowed carrier depicted in blue takes on a value of zero at the peak location of the other carriers. The center windowed carrier depicted in red is nonzero at the other carrier centers and thus introduces ICI.

using the optimum Nyquist approach of [12] maintains orthogonality; however, in principle there is a trade between interference mitigation performance and ICI. The window shape also impacts the amount of interframe interference (IFI) and the amount of thermal noise. In addition to the rectangular window, Figure 2 shows examples of a Taylor window and an optimum Nyquist window.

In this paper, a technique is developed for determining optimal windows that trade off interference suppression with ICI and IFI. This technique assumes that statistics of the link delay spread and interferer are known. By using this technique, windows for various levels of interference, delay spreads, frame length, and cyclic prefix lengths can be produced.

For the sake of simplifying the analysis in the paper, inter-

This work is sponsored by the DARPA under Air Force contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government. The views, opinions, and/or findings contained in this article are those of the author and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense. Approved for Public Release, Distribution Unlimited

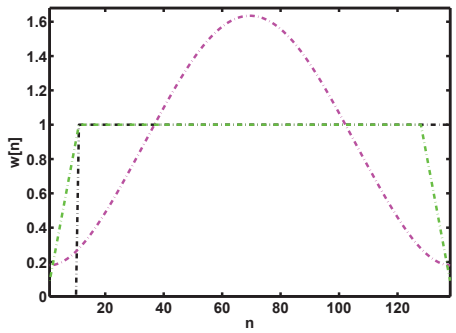

Fig. 2. **Traditional Windowing** – Depiction of a rectangular window shown in black, a Taylor window with parameter -40 dB shown in magenta, and a trapezoidal window shown in green. The trapezoidal window does not introduce ICI and satisfies the optimum Nyquist constraints specified in [12]. Here the FFT size N is 128 and the cyclic prefix is equal to 10.

ference is assumed to be rank-one. This rank-one interference could be due to simple external interference because of the use of single-input multiple-output (SIMO) links, or because of the use of a dominant-singular-vector informed-transmitter MIMO approach [5] in which the channel state information is known and employed by the transmitter of the interfering link. Here the interfering transmitter uses only the dominant (strongest) mode. Even if an uninformed-transmitter MIMO link is employed, the ICI-interference mitigation trade space is appropriate for each transmitting antenna.

II. Channel

In networks with SIMO links or dominant-singular-vector informed-transmitter MIMO, the input-output relationship between the receiver and the rank-one transmitted signal can be viewed as

$$\mathbf{x}_p = \mathbf{g}_p * \mathbf{s} + \boldsymbol{\nu}_p, \tag{1}$$

where $\mathbf{x}_p \in \mathbb{C}^{\bar{N}\times 1}$ is the complex received signal at antenna p, $\bar{N}$ is the length of the received signal, $\mathbf{s} \in \mathbb{C}^{\bar{N}+M-1\times 1}$ is the transmitted time series, $\mathbf{g}_p \in \mathbb{C}^{M\times 1}$ the linear channel between the rank-one transmitter and receiver p, M is the length of the longest channel response, $*$ denotes linear convolution defined as

$$\mathbf{g}_p * \mathbf{s} = \sum_{m=0}^{M-1} g_p[m]s[n-m], \tag{2}$$

and $\boldsymbol{\nu}_p \in \mathbb{C}^{\bar{N}\times 1}$ is Gaussian zero mean thermal noise with variance σ_ν^2. Vectors are written in lowercase bold ($\mathbf{a}$), while matrices are written in uppercase bold ($\mathbf{A}$). The individual elements of the vector $\mathbf{x}_p$ can be accessed using the notation $x_p[n]$ for values of $n \in [0, \bar{N}-1]$. Under this notation, the functions are defined to be zero for values of n outside the specified interval. The individual elements of the vector $\mathbf{s}$ can be accessed by using the notation $s[n]$ for values of $n \in [1-M, \bar{N}-1]$. Similarly, the values of $\mathbf{g}_p$ can be accessed by using the notation $g_p[n]$ for $n \in [0, M-1]$.

A. Delay Spread

For this analysis, the channel $\mathbf{g}$ is defined to be an exponential channel satisfying the relationship

$$E[|g[n]|^2] = E[|g[0]|^2] \exp^{-\frac{nT_s}{\tau}}, \tag{3}$$

with delay spread τ and sampling interval T_s. These techniques can be generalized to other channel models.

III. Interference

Interference can be added to the model in Equation (1)

$$\mathbf{x}_p = \mathbf{g}_p * \mathbf{s} + \mathbf{h}_p * \mathbf{j} + \boldsymbol{\nu}_p, \tag{4}$$

where $\mathbf{j} \in \mathbb{C}^{\bar{N}+M-1\times 1}$ is the transmitted interfering time series and $\mathbf{h_p} \in \mathbb{C}^{M\times 1}$ is the linear channel between the rank-one interferer and receiver p. In this setting, the interferer is treated as if it is not synchronized with the signals of interest. The unsynchronized signal of interest leads to a lack of cyclic prefix that eliminates the orthogonality condition.

IV. Optimal Approach

In this section, a method is introduced for the determination of windows that optimally trade interference mitigation with the combined effects of intercarrier interference, interframe interference, and thermal noise. The method solves for windows that maximize the expression

$$\Phi(\mathbf{w}) = Int(\mathbf{w}) - \beta(ICI(\mathbf{w}) + IFI(\mathbf{w}) + TN(\mathbf{w})), \tag{5}$$

where $\mathbf{w} \in \mathbb{C}^{\bar{N}\times 1}$ is the choice of window with $w[n]$ nonzero over the interval $[0, \bar{N}-1]$, β is a coefficient proportional to the expected signal-to-interference ratio (SIR), $Int(\mathbf{w})$ is a measure of interference mitigation, $ICI(\mathbf{w})$ is the intercarrier interference, $IFI(\mathbf{w})$ is the interframe interference, and $TN(\mathbf{w})$ is the thermal noise. Gradient descent is used to solve the Euler-Lagrange equation,

$$\partial_{\mathbf{w}}\Phi(\mathbf{w}_{opt}) = 0, \tag{6}$$

for the best window $\mathbf{w}_{opt}$. The actual gradient used in the descent is provided in the Appendix. Figure 3 provides a family of windows that is parameterized by β.

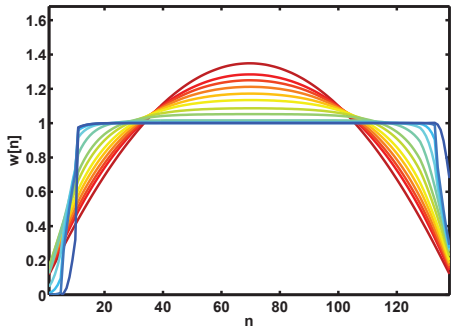

Fig. 3. **Family of Windows** – A set of optimal windows parameterized by the value of β. A large value of β corresponds to a flat window, while small values of β correspond to a more peaked window. Here the FFT size N is 128, $\bar{N}$ is 138, the delay spread is 600 nanoseconds, and the sampling interval is 100 nanoseconds.

The remainder of this section provides expressions for each of the contributing terms in the objective function in Equation (5).

A. Interference Mitigation

For a particular receiver p, the windowed version of the received interference is

$$\mathbf{z}_p = \mathbf{w} \odot (\mathbf{h}_p * \mathbf{j}), \tag{7}$$

where $\odot$ denotes the Hadamard or pointwise product between two vectors. Taking the DFT, the received OFDM bins are written as

$$\hat{\mathbf{z}}_p = \mathcal{F}\{\mathbf{w} \odot (\mathbf{h}_p * \mathbf{j})\}, \tag{8}$$

where $\mathcal{F}$ is the N point DFT of a vector of length $\bar{N}$ defined as

$$\mathcal{F}\{\mathbf{x}\} = \sum_{n=0}^{\bar{N}-1} e^{\frac{2\pi i n k}{N}} x[n]. \tag{9}$$

This definition of the DFT follows the treatment in [13].

The spatial covariance between receive antenna p and q for the k^{th} carrier frequency can be written as

$$R_{p,q}[k] = E[\hat{\mathbf{z}}_p^\dagger \hat{\mathbf{z}}_q] \tag{10}$$

$$= \sigma_j^2 (\mathbf{h}_p^k)^\dagger \mathbf{A} \mathbf{h}_q^k, \tag{11}$$

where $(\cdot)^\dagger$ denotes the Hermitian transpose, the interfering signal has an autocorrelation of a unit impulse times σ_j^2, $A_{m,\tilde{m}}$ is the cross correlation of the window

$$A_{m,\tilde{m}} = \sum_{n=0}^{\bar{N}-1} w[n] w[n + \tilde{m} - m], \tag{12}$$

and the channel for a particular carrier k is

$$\mathbf{h}_p^k = \mathbf{h}_p \odot \mathbf{v}_N^{-k}, \tag{13}$$

with the corresponding phase ramp

$$\mathbf{v}_N^{-k} = \begin{bmatrix} e^{(\frac{2\pi i k}{N}) \cdot 0} \\ e^{(\frac{2\pi i k}{N}) \cdot 1} \\ \vdots \\ e^{(\frac{2\pi i k}{N}) \cdot (\bar{N}-1)} \end{bmatrix}. \tag{14}$$

By stacking the P channel impulse responses $\mathbf{h}_p^k$ into the matrix $\mathbf{H}^k$, the final covariance matrix can be written as

$$\mathbf{R}[k] = \sigma_j^2 (\mathbf{H}^k)^\dagger \mathbf{A} \mathbf{H}^k. \tag{15}$$

B. Channel Whitening

In the analysis in the previous section, it was implicitly assumed that the channels were white, i.e., $E[(\mathbf{H}^k)^\dagger \mathbf{H}^k] = \mathbf{I}_P$, where $\mathbf{I}_P$ is the identity matrix of size $P \times P$. When the channels are in fact not white, the performance of the filters will be suboptimal. To combat this, $\mathbf{H}^k$ is replaced with $\tilde{\mathbf{H}}^k = \mathbf{\Lambda}_\mathbf{H}^{-1/2} \mathbf{H}^k$ in Equation (15),

$$\mathbf{R}[k] = \sigma_j^2 (\tilde{\mathbf{H}}^k)^\dagger \mathbf{\Lambda}_\mathbf{H}^{1/2} \mathbf{A} \mathbf{\Lambda}_\mathbf{H}^{1/2} \tilde{\mathbf{H}}^k, \tag{16}$$

where $\mathbf{\Lambda}_\mathbf{H} = E[\mathbf{H}^k (\mathbf{H}^k)^\dagger]$.

The matrix $\mathbf{R}$ is similar to the matrix $\mathbf{\Lambda}_\mathbf{H}^{1/2} \mathbf{A} \mathbf{\Lambda}_\mathbf{H}^{1/2}$ and therefore shares the same eigenvalues. In order to maximize interference suppression, the objective is to concentrate the energy in the largest eigenvalue. This is accomplished when the choice of $\mathbf{w}$ maximizes the ratio between the largest eigenvalue and the sum of the eigenvalues. This sum of the eigenvalues is equal to the trace of the matrix, which is simply proportional to a constant times $A_{0,0}$ or $\mathbf{w}^\dagger \mathbf{w} = 1$.

The optimal window can be found through the joint optimization of the functional

$$F(\mathbf{w}, \mathbf{x}) = \frac{\mathbf{x}^\dagger \mathbf{\Lambda}_\mathbf{H}^{1/2} \mathbf{A} \mathbf{\Lambda}_\mathbf{H}^{1/2} \mathbf{x}}{\mathbf{w}^\dagger \mathbf{w}} \tag{17}$$

subject to the constraint

$$\mathbf{x}^\dagger \mathbf{x} = 1, \tag{18}$$

where $\mathbf{x} \in \mathbb{C}^{\bar{M} \times 1}$ is an arbitrary vector with Euclidian norm one. Because this optimization is convex, the optimization can be performed for $\mathbf{w}$ and $\mathbf{x}$ separately while the other is held fixed. This process is repeated until $F(\mathbf{w}, \mathbf{x})$ converges to the maximum value.

The two optimizations can be written as

$$\mathbf{x}_{opt} = \underset{\mathbf{x}}{\operatorname{argmax}} F(\mathbf{w}_{opt}, \mathbf{x}) \tag{19}$$

and

$$\mathbf{w}_{opt} = \underset{\mathbf{w}}{\operatorname{argmax}} F(\mathbf{w}, \mathbf{x}_{opt}). \tag{20}$$

The value of $\mathbf{x}_{opt}$ in Equation (19) is equal to the eigenvector corresponding to the largest eigenvalue of the matrix $\mathbf{A}$. For a particular choice of the window $\mathbf{w}$, the metric for the interference mitigation is given by

$$Int(\mathbf{w}) = \max_{\mathbf{x}} F(\mathbf{w}, \mathbf{x}). \tag{21}$$

C. Intercarrier Interference

In addition to mitigating interference, the use of windows can degrade the signal of interest $\mathbf{s}$ through the introduction of ICI. The contribution of the signal of interest at a particular receiver can be written as

$$\mathbf{t} = \mathbf{w} \odot (\mathbf{g} * \mathbf{s}). \tag{22}$$

The power of ICI for carrier k is simply the sum of the received power contributed by all of the other subcarriers. In this section, $s[n]$ is periodic with period N, i.e., $s[n] = s[(n)_N]$, where $(\cdot)_N$ returns the argument modulo N. The expression for the expected power of ICI is

$$ICI(\mathbf{w}) = N^2 \sigma_s^2 \sigma_g^2 \sum_{n=0}^{N-1} \left(\sum_{\bar{n}=0}^{\lceil \frac{\bar{N}}{N} \rceil} w[n + \bar{n}N] - \frac{\mathbf{1}^\dagger \mathbf{w}}{N} \right)^2, \tag{23}$$

where $\lceil \cdot \rceil$ indicates the ceiling operation.

D. Interframe Interference

In cases where the cyclic prefix is not longer than the delay spread of the channel, IFI is introduced. The linear convolution can be broken up into a circular convolution plus an adjustment term that contributes to the IFI,

$$t[n] = w[n]\left((g\textcircled{N}s)[n] + \sum_{m=n+Cp+1}^{M-1} g[m](s[n-m] - s[(n-m)_N])\right), \quad (24)$$

where $n \in [0, \bar{N})$ and $\textcircled{N}$ denotes the periodic convolution defined as

$$(g\textcircled{N}s)[n] = \sum_{m=0}^{M-1} g[m]s[(n-m)_N]. \quad (25)$$

This can be time aliased to a vector of length N,

$$t[n] = \sum_{\tilde{n}=0}^{\lceil \frac{\bar{N}}{N} \rceil} w[n+\tilde{n}N]\left(\sum_{m=0}^{M-1} g[m]s[(n+\tilde{n}N-m)_N] + \sum_{m=n+Cp+1}^{M-1} g[m](s[n+\tilde{n}N-m] - s[(n+\tilde{n}N-m)_N])\right), \quad (26)$$

where $n \in [0, N)$. Simplifying the expression further,

$$t[n] = (g\textcircled{N}s)[n]\sum_{\tilde{n}=0}^{\lceil \frac{\bar{N}}{N} \rceil} w[n+\tilde{n}N] + w[n]\sum_{m=n+Cp+1}^{M-1} g[m]\left(s[n-m] - s[(n-m)_N]\right). \quad (27)$$

In the case with no ICI and a sufficiently long cyclic prefix, the received signal of interest can be written as

$$\hat{t}[n] = (g\textcircled{N}s)[n]. \quad (28)$$

The error term can then be written as $||t - \hat{t}||_2^2$. By using Parseval's theorem, the power of the ICI plus IFI penalty is

$$ICI(\mathbf{w}) + IFI(\mathbf{w}) = N^2\sigma_s^2\sigma_g^2\sum_{n=0}^{N-1}\left(\sum_{\tilde{n}=0}^{\lceil \frac{\bar{N}}{N} \rceil} w[n+\tilde{n}N] - 1\right)^2 + N\sigma_s^2\sum_{n=0}^{M-C_p-2} 2w[n]\left(1 - \sum_{\tilde{n}=1}^{\lceil \frac{\bar{N}}{N} \rceil} w[n+\tilde{n}N]\right)\sum_{m=n+Cp+1}^{M-1}\sigma_g^2[m], \quad (29)$$

where $\sigma_g^2 = \sum_{m=0}^{M-1}\sigma_g^2[m]/N$.

E. Thermal Noise

In addition to ICI and IFI, the shape of the window also impacts the effect of the thermal noise. Thermal noise can be added to the model in Equation (22)

$$\mathbf{y} = \mathbf{t} + \mathbf{w} \odot \boldsymbol{\nu}, \quad (30)$$

where $\nu[n]$ is zero mean Gaussian noise with variance σ_ν^2. The contribution of the $\nu[n]$ term to the overall noise level is uncorrelated to the other terms so it can be examined separately. The expected power contribution of the noise in the expression is

$$TN(\mathbf{w}) = N\sum_{n=0}^{N-1}\left(\sum_{\tilde{n}=0}^{\lceil \frac{\bar{N}}{N} \rceil} w[n+\tilde{n}N]\nu[n+\tilde{n}N]\right)^2 \quad (31)$$

$$= N\sigma_\nu^2\mathbf{w}^\dagger\mathbf{w}. \quad (32)$$

V. Performance

The ratio of the first eigenvalue to the second eigenvalue is used as a metric of interference suppression. It measures how well the interference is confined to a single degree of freedom for the carrier of interest. In Figure 4, the amount of interference suppression the windows of interest provide is plotted versus the combined effects of ICI, thermal noise, and IFI. The rectangular window has the least amount of interference suppression and the least contribution from the other three terms. The Nyquist window has similar amounts of interference mitagation to the third optimum window, but performs worse in the other metrics. The Taylor window does a good job at mitigating interference, but it does the worst on the combined term. Lastly, the family of windows provides the best performance for each level of interference mitigation.

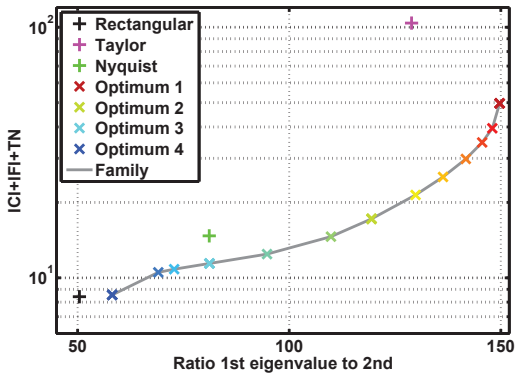

Fig. 4. **ICI Interference Trade-off** – The trade-off provided by the various windows between the interference mitigation (ratio of the 1st eigenvalue and the 2nd) with the combined effects of ICI, thermal noise, and IFI. The windows considered are those depicted in the family of windows shown in Figure 3 and the traditional windows of Figure 2.

The measures of the interference suppression along with the other terms are useful, but the ultimate goal is to maximize the channel capacity of the resulting windowed OFDM sequence. Figure 5 shows the measured capacity for the bank of windows for various SIR levels, while Figure 6 shows the ratio of the capacity curve for each window with the capacity curve for the rectangular window. Each data point was determined using a Monte Carlo simulation of ten thousand trials.

The optimum 1 window has the best performance for the low SIR case, the optimum 2 window is best in the SIR range from -22 dB to -11 dB, and the optimum 3 window is best in

the SIR range from -11dB to 6dB (not shown in figure), where the optimum 4 window takes over. The Taylor window is a strong performer for low SIR levels, but its performance versus the other windows drops off rapidly at higher SIR levels. The Nyquist window is similar to the optimum 3 window and is a strong performer at higher SIR levels. Finally, the rectangular window is dominated by the optimum 4 window. The optimum 4 window does well at high SIR levels and the worst at low SIR levels.

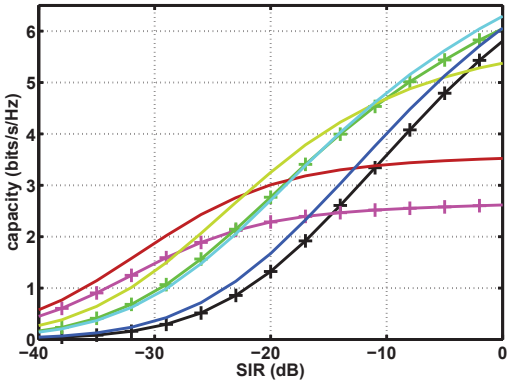

Fig. 5. **Capacity** – The capacity of the MIMO channel for each of the specified windows over a wide range of signal to interference ratios.

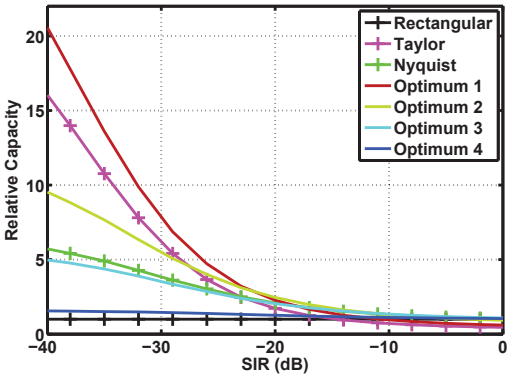

Fig. 6. **Relative Capacity** – The ratio of the capacity of the MIMO channel for each of the specified windows to the capacity obtained using the rectangular window over a wide range of signal-to-interference ratios.

VI. Conclusion

A new method for designing windows for various levels of interference, delay spreads, frame length, and cyclic prefix lengths is presented. These windows optimally trade off interference suppression with intercarrier and interframe interference. The windows achieve superior capacity to traditional windows in the presence of a wide range of interference levels.

The implementation of these windows is computationally simple. Furthermore, the windows can be precomputed for a range of SIR levels and delay spreads. In the example, only three windows are needed (optimum 1, 2, and 3) to achieve a hybrid window that performs well over a wide range of SIR levels. Because of these properties, the windows can be easily implemented in real systems.

VII. Acknowledgments

The authors would like to thank Jean-Francois Frigon of Ecole Polytechnique de Montreal for thoughtful discussions.

References

[1] E. Perahia and R. Stacey, *Next Generation Wireless LANs: Throughput, Robustness, and Reliability in 802.11n*. Cambridge: Cambridge University Press, 2007.

[2] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 2007.

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[5] D. W. Bliss, K. W. Forsythe, A. O. Hero, and A. F. Yegulalp, "Environmental issues for MIMO capacity," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2128–2142, Sept. 2002.

[6] A. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: Cambridge University Press, 2005.

[8] J. Ward and R. T. Compton, Jr., "High throughput slotted ALOHA packet radio networks with adaptive arrays," *IEEE Transactions on Communications*, vol. 41, no. 3, pp. 460–470, March 1993.

[9] S. Govindasamy, D. W. Bliss, and D. H. Staelin, "Spectral efficiency in single-hop ad-hoc wireless networks with interference using adaptive antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1358–1369, September 2007.

[10] T. T. Taylor, "Design of line-source antennas for narrow beamwidth and low side lobes," *IRE Transactions on Antennas and Propagation*, vol. AP-3, pp. 16–28, 1955.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice Hall, 1993.

[12] S. Muller-Weinfurtner and J. Huber, "Optimum Nyquist windowing for improved OFDM receivers," *IEEE Global Telecommunications Conference, 2000.*, vol. 2, pp. 711–715 vol.2, 2000.

[13] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, N.J: Prentice-Hall, 1983.

VIII. Appendix

The first variation of Equation (5) is

$$\partial_{\mathbf{w}}(\Phi) = \frac{\sum_{m=0}^{M-1}\sum_{\tilde{m}=0}^{M-1} x[m]x[\tilde{m}]\Lambda_{\mathbf{H}}^{\frac{1}{2}}[m]\Lambda_{\mathbf{H}}^{\frac{1}{2}}[\tilde{m}]w[n-(m-\tilde{m})]}{\sum_{\tilde{n}=0}^{\tilde{N}-1} w[\tilde{n}]^2} - \frac{\mathbf{x}^{\dagger}\Lambda_{\mathbf{H}}^{\frac{1}{2}}\mathbf{A}\Lambda_{\mathbf{H}}^{\frac{1}{2}}\mathbf{x}}{A[0]^2}w[n] - \frac{\beta}{\sigma_g^2 N^2 SNR}w[n] - \beta\sum_{m=n+Cp+1}^{M-1}\frac{\sigma_g^2[m]}{N\sigma_g^2}\frac{(1-\sum_{\tilde{n}=1}^{\lceil\frac{\tilde{N}}{N}\rceil}w[n+\tilde{n}N])}{N}(1-u[n-N]) + \beta\frac{w[(n)_N]}{N}\sum_{m=(n)_N+Cp+1}^{M-1}\frac{\sigma_g^2[m]}{N\sigma_g^2}u[n-N], \quad (33)$$

where $\beta = MN^3\sigma_g^2 SIR$, $SIR = \frac{\sigma_s^2}{\sigma_j^2}$, and $SNR = \frac{\sigma_s^2}{\sigma_\nu^2}$.