ASYMPTOTIC ANALYSIS OF REDUCED RANK WIENER FILTERS.

Philippe Loubaton, Walid Hachem***

*Institut Gaspard Monge et Laboratoire Traitement et Communication de l'Information (CNRS/ENST)
Université de Marne-La-Vallée, 5, Bld. Descartes, 77454 Marne-La-Vallée Cedex 2, France.
**Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France.

ABSTRACT

In this paper, we revisit recent papers of Honig-Xiao and Trichard *et al.* devoted to the asymptotic analysis of reduced rank Wiener filters. Appropriate connections between the asymptotic behavior of the Signal to Noise Ratios (SNRs) at the outputs of these filters and the theory of orthogonal polynomials for the power moment problem are established. Using some classical results of this theory, it can be established in particular that the reduced rank filter output SNR converges exponentially in the filter rank toward the full rank Wiener filter output SNR. The convergence rate is given. Interestingly, it depends only on the support of the limiting eigenvalue distribution of the observation covariance matrix, but not on its particular form.

1. INTRODUCTION

In multidimensional signal processing, it is often useful to approximate the Wiener filter by a reduced rank version of this filter. The latter acts on a projection of the received signal on a judiciously chosen small dimensional subspace. The use of a reduced rank filter can be motivated by complexity constraints or, in an adaptive setting, by fast convergence requirements. It is then of major interest to quantify the SNR loss at the output of this filter due to its non optimum character.
The *Krylov subspaces*, widely used as projection subspaces, will be considered in this paper. To fix our ideas, let us begin with the generic signal model

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

where $\mathbf{y}$ is the received $N \times 1$ signal, s is the unit-variance scalar signal to be estimated and $\mathbf{x}$ is a signal decorrelated with s. The $N \times N$ covariance matrix of $\mathbf{x}$ is denoted $\mathbf{R}_I$ and will be assumed invertible. Recall that the MMSE receiver is described by the equation $s_{\mathrm{MMSE}} = \mathbf{h}^H \mathbf{R}^{-1} \mathbf{y}$ where $\mathbf{R} = \mathbf{h}\mathbf{h}^H + \mathbf{R}_I$ is the received signal $\mathbf{y}$ covariance matrix. This receiver will be called in the sequel the full rank MMSE receiver. Its output SNR that we index by the number of dimensions of the received signal is given by the standard expression

$$\beta^{(N)} = \mathbf{h}^H \mathbf{R}_I^{-1} \mathbf{h} \, . \tag{2}$$

The n^{th} Krylov subspace associated to the pair $(\mathbf{R}_I, \mathbf{h})$ is the subspace of $\mathbb{C}^N$ spanned by the columns of $\mathbf{K}_n = [\mathbf{h}, \mathbf{R}_I\mathbf{h}, \ldots, \mathbf{R}_I^{n-1}\mathbf{h}]$. The n-stage reduced rank Wiener filter considered in this paper is the MMSE estimator of s operating on the transformed signal $\tilde{\mathbf{y}}_n = \mathbf{K}_n^H \mathbf{y}$.
The motivation behind choosing the Krylov subspaces and the implementation of the subsequent filters are discussed in a number of works (see [5] and [4]). In this paper, we focus on the convergence of their output SNR toward that of the full rank MMSE receiver.
The performance of these filters has been studied by Honig and Xiao ([5]) and Trichard *et. al.* ([9], [8]) in the context of a CDMA transmission. A signal model considered in these papers writes

$$\mathbf{y} = \mathbf{W}\mathbf{s} + \mathbf{v} \, . \tag{3}$$

$\mathbf{s} = [s_1, \ldots, s_K]^T$ is the $K \times 1$ symbol vector where K is the number of users, $\mathbf{W}$ is the $N \times K$ code matrix, and $\mathbf{v}$ is the classical noise with covariance matrix $\omega^2 \mathbf{I}_N$. The purpose is to estimate the symbol s_1, so this equation appears as a particular case of (1) : if we partition $\mathbf{W}$ and $\mathbf{s}$ as $\mathbf{W} = [\mathbf{w} \ \mathbf{U}]$ and $\mathbf{s} = [s_1 \ \mathbf{s}_I^T]^T$, then we replace $\mathbf{h}$ by $\mathbf{w}$ and $\mathbf{x}$ by $\mathbf{U}\mathbf{s}_I + \mathbf{v}$. These authors assumed that the code matrix $\mathbf{W}$ is a random matrix with centered i.i.d. elements having a variance of $1/N$, and studied the performance of the reduced rank filter in the "large system" regime where N tends to infinity in such a way that K/N converges toward a constant α. They established that in this case, the Signal to Interference plus Noise Ratio (SINR) at the output of the reduced rank Wiener filter presented above expresses as a continued fraction expansion in n and concluded for the rapid convergence of this SINR toward the full rank SINR.

The purpose of this paper is to precise this convergence in a broad number of situations where reduced rank filtering is used. We shall also consider transmission models with large dimensions and do our study in the asymptotic regime where $N \to \infty$. Getting back to the general model (1), the main assumption we need to formulate concerns the quantities $s_k^{(N)} = \mathbf{h}^H \mathbf{R}_I^k \mathbf{h}$. Precisely, these are required to converge toward the moments of some probability measure carried by a compact interval $[\delta_1; \delta_2]$ in $(0; \infty)$. Our results rely then on the theory of orthogonal polynomials for the so called power moment problem.
Section 2 recalls some mathematical results relative to these polynomials. In section 3, the SNR convergence is studied. Some examples are then given in section 4.

2. ORTHOGONAL POLYNOMIALS FOR THE POWER MOMENT PROBLEM

In the sequel, it will be assumed that $\mathbf{h}$ is a unit norm vector and that the Krylov matrix $\mathbf{K}_n$ has a full rank n for each n. It can be easily shown that the SNR $\beta_n^{(N)}$ at the output of the n-stage reduced rank Wiener filter is

$$\beta_n^{(N)} = \mathbf{h}^H \mathbf{K}_n \left(\mathbf{K}_n^H \mathbf{R}_I \mathbf{K}_n \right)^{-1} \mathbf{K}_n^H \mathbf{h} \, . \tag{4}$$

We denote by $\mathbf{C}_n^{(N)} = [\mathbf{c}_0^{(N)}, \ldots, \mathbf{c}_{n-1}^{(N)}]$ the orthonormal basis obtained by a Gram-Schmidt orthogonalization of the column space of $\mathbf{K}_n$. Then, $\beta_n^{(N)}$ can be also be written as

$$\beta_n^{(N)} = \left[\left({\mathbf{C}_n^{(N)}}^H \mathbf{R}_I \mathbf{C}_n^{(N)} \right)^{-1} \right]_{1,1} \tag{5}$$

where the notation $[\mathbf{X}]_{i,j}$ designates the entry (i, j) of a matrix $\mathbf{X}$. It is useful to recall that, due to the particular structure of $\mathbf{K}_n$, the positive matrix ${\mathbf{C}_n^{(N)}}^H \mathbf{R}_I \mathbf{C}_n^{(N)}$ is tridiagonal, i.e.

$${\mathbf{C}_n^{(N)}}^H \mathbf{R}_I \mathbf{C}_n^{(N)} = \begin{bmatrix} a_0^{(N)} & b_0^{(N)} & 0 & \ldots & 0 \\ b_0^{(N)} & a_1^{(N)} & b_1^{(N)} & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & b_{n-2}^{(N)} \\ 0 & \ldots & 0 & b_{n-2}^{(N)} & a_{n-1}^{(N)} \end{bmatrix} . \tag{6}$$

As remarked by Honig and Xiao, it is possible to have a better understanding of the significance of the right hand side of (5) if we assume that $N \to +\infty$. Recall that it will be assumed throughout the paper that each coefficient $\mathbf{h}^H \mathbf{R}_I^k \mathbf{h}$ converges towards a certain limit s_k when $N \to \infty$ (note that $s_0 = 1$ because we have assumed that $\|\mathbf{h}\| = 1$ for each N). We first mention that the positive definite $n \times n$ matrix $\mathbf{K}_n^H \mathbf{K}_n$ is a Hankel matrix with (k, l) entries given by $\mathbf{h}^H \mathbf{R}_I^{k+l-2} \mathbf{h}$. Because $\mathbf{K}_n^H \mathbf{K}_n$ is a positive definite Hankel matrix for each $n \le N$, the $n \times n$ Hankel matrix $\mathbf{S}_n$ defined by

$$[\mathbf{S}_n]_{k,l} = s_{k+l-2}, \;\; k, l = 1, \ldots, n$$

is positive. Similarly, the $(n-1) \times (n-1)$ Hankel matrix $\mathbf{K}_{n-1}^H \mathbf{R}_I \mathbf{K}_{n-1}$ is positive definite, and the Hankel matrix $\tilde{\mathbf{S}}_{n-1}$ defined by

$$\left[\tilde{\mathbf{S}}_{n-1} \right]_{k,l} = s_{k+l-1}, \;\; k, l = 1, \ldots, n-1$$

is also positive. Consequently (see e.g. [1, p. 76]), the sequence $(s_k)_{k \ge 0}$ coincides with the power moment sequence of a certain probability measure σ (recall that $s_0 = 1$) carried by $\mathbb{R}^+$. In other words, it exists a probability measure σ such that

$$s_k = \int_0^\infty \lambda^k d\sigma(\lambda)$$

for $k \ge 0$. In the following, we assume that this measure is unique (the moment problem is called determinate in this case, see [1]), that it is carried by a compact interval $[\delta_1, \delta_2]$ with $0 < \delta_1 < \delta_2 < \infty$, and that it is absolutely continuous with an almost everywhere strictly positive (Radon-Nykodym) derivative on $[\delta_1, \delta_2]$. When the measure is not absolutely continuous, weaker results can be obtained (see [6]). This case will be assessed in a forthcoming paper.

It is classical to associate to the sequence $(s_k)_{k \ge 0}$ (or equivalently to the measure σ) the scalar product on the space $\mathbb{L}^2(\sigma)$ defined by

$$< f(\lambda), g(\lambda) >= \int_{\delta_1}^{\delta_2} f(\lambda) g(\lambda) d\sigma(\lambda) \; .$$

By the very definition of sequence $(s_k)_{k \ge 0}$, this scalar product verifies

$$< \lambda^k, \lambda^l >= \lim_{N \to \infty} \mathbf{h}^H \mathbf{R}_I^{k+l} \mathbf{h} = s_{k+l} \; . \tag{7}$$

In the following, we denote by $(p_k(\lambda))_{k \ge 0}$ the orthonormal polynomials obtained by a Gram-Schmidt orthogonalization of the vector space generated by $\{1, \lambda, \ldots, \lambda^{n-1}, \ldots\}$. The properties of these polynomials are well established. We first recall some of their useful basic properties (see e.g. [1], [7])

Proposition 1 (The three terms recursion relation) *The family of polynomials $\{p_k\}$ satisfies the relation*

$$\lambda p_k(\lambda) = b_{k-1} p_{k-1}(\lambda) + a_k p_k(\lambda) + b_k p_{k+1}(\lambda) \tag{8}$$

where coefficients a_k and b_k, defined by $a_k =< \lambda p_k(\lambda), p_k(\lambda) >$ and $b_k =< \lambda p_k(\lambda), p_{k+1}(\lambda) >$ are positive. The reccurence formula is initialized by $p_0(\lambda) = 1$ and $p_1(\lambda) = \frac{\lambda - a_0}{b_0}$, with $a_0 = s_1$.

This relation implies that for each n, the $n \times n$ matrix which (k, l) entry is $< \lambda p_{k-1}(\lambda), p_{l-1}(\lambda) >$ coincides with the positive definite tri-diagonal matrix $\mathbf{T}_n(\sigma)$ given by

$$\mathbf{T}_n(\sigma) = \begin{pmatrix} a_0 & b_0 & 0 & \ldots & 0 \\ b_0 & a_1 & b_1 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & b_{n-2} \\ 0 & \ldots & 0 & b_{n-2} & a_{n-1} \end{pmatrix} . \tag{9}$$

Proposition 2 *For each k, the zeros $(\lambda_{n,k})_{k=1,\ldots,n}$ of $p_k(\lambda)$ are simple (and thus real) and belong to $[\delta_1, \delta_2]$. Moreover, the $(\lambda_{n,k})_{k=1,\ldots,n}$ coincide with the eigenvalues of matrix $\mathbf{T}_n(\sigma)$, the corresponding unit norm eigenvectors being vectors*

$$\frac{1}{[\sum_{l=0}^{n-1} (p_l(\lambda_{n,l}))^2]^{1/2}} [p_0(\lambda_{n,k}), \ldots, p_{n-1}(\lambda_{n,k})]^T \; .$$

We need to introduce the **second kind orthogonal polynomials**. The second kind orthogonal polynomials $(q_k(\lambda))_{k \ge 0}$ are defined from the $(p_k(\lambda))_{k \ge 0}$ by

$$q_k(\lambda) = \int_{\delta_1}^{\delta_2} \frac{p_k(\lambda) - p_k(u)}{\lambda - u} \, d\sigma(u) \tag{10}$$

It is easily seen that $\deg(q_k) = k - 1$, and that $q_0(\lambda) = 0$ and $q_1(\lambda) = \frac{1}{b_0}$. Moreover, the second kind polynomials verify formula (8) with the same coefficients.

Let $G_\sigma(z)$ the Stieltjes transform of measure σ :

$$G_\sigma(z) = \int_{\delta_1}^{\delta_2} \frac{1}{\lambda - z} d\sigma(\lambda)$$

G_σ is of course analytic in $\mathbb{C} - [\delta_1, \delta_2]$. Moreover, $\text{Im}(G_\sigma)(z) > 0$ if $\text{Im}(z) > 0$, and thus belongs to the so-called Nevanlinna class. It is clear that for $|z|$ large enough, $G_\sigma(z)$ can be written as

$$G_\sigma(z) = -\sum_{k=0}^{\infty} \frac{s_k}{z^{k+1}}$$

It is well known that $-\frac{q_n(z)}{p_n(z)}$ is a Padé approximant of $G_\sigma(z)$ in the sense that

$$-\frac{q_n(z)}{p_n(z)} = -\sum_{k=0}^{2n-1} \frac{s_k}{z^{k+1}} + o(z^{-(2n)})$$

Moreover, $-\frac{q_n(z)}{p_n(z)}$ converges uniformly towards $G_\sigma(z)$ on compact subsets of $\mathbb{C} - [\delta_1, \delta_2]$. We finally note that $G_\sigma(z)$ admits the continued fraction expansion

$$G_{\sigma(z)} = \frac{-1}{z - a_0 - \frac{b_0^2}{z - a_1 - \frac{b_1^2}{z - a_2 \cdots}}} \tag{11}$$

and that the rational approximant $-\frac{q_n(z)}{p_n(z)}$ coincides with the truncation of the above expansion up to order n.

3. APPLICATION TO THE CONVERGENCE OF REDUCED RANK WIENER FILTERS.

It is clear that $(p_k)_{k=0,\dots,n-1}$ are the polynomial counterparts of the limits when $N \to +\infty$ of the orthonormal vectors $\mathbf{c}_0^{(N)}, \dots, \mathbf{c}_{n-1}^{(N)}$. In particular, using relation (7), it is easily checked that the tridiagonal matrix $\mathbf{T}_n(\sigma)$ coincides with the limit of matrix $\mathbf{C}_n^{(N)H} \mathbf{R}_I \mathbf{C}_n^{(N)}$. This remark allows to precise the asymptotic behavior of the SINR $\beta_n^{(N)}$ when $N \to \infty$.

Proposition 3 *For each $n \geq 0$, the SINR $\beta_n^{(N)}$ converges when $N \to \infty$ toward β_n given by*

$$\beta_n = \left[\mathbf{T}_n(\sigma)^{-1}\right]_{1,1} \tag{12}$$

Using the eigenvalue/eigenvector decomposition of $\mathbf{T}_n(\sigma)$ in conjunction with certains properties of $p_n(\lambda)$ and $q_n(\lambda)$, it can be shown that

Proposition 4 *β_n coincides with $-\frac{q_n(0)}{p_n(0)}$.*

β_n thus converges to the term $\beta = G_\sigma(0)$ given by

$$\beta = \int_{\delta_1}^{\delta_2} \frac{1}{\lambda} d\sigma(\lambda)$$

which also represents the limit as $N \to \infty$ of $\beta^{(N)}$ given by (2). The study of the convergence of the reduced rank Wiener filters thus reduces, in the asymptotic regime, to the convergence of $-\frac{q_n(0)}{p_n(0)}$ toward $G_\sigma(0)$. The following theorem is the main result of this paper.

Theorem 1 *Let $\mu = \frac{\delta_2+\delta_1}{\delta_2-\delta_1}$ and denote by ϕ the term defined by*

$$\phi = \frac{1}{\mu + \sqrt{\mu^2 - 1}}$$

Then, the error $e_n = \beta - \beta_n$ satisfies for n large enough the inequality

$$A\phi^{2n} \leq |e_n| \leq B\phi^{2n} \tag{13}$$

where A and B are two strictly positive constants.

This result follows directly from the fact that (see e.g. [7] for the case of measures carried by [-1, 1])

$$|p_n(0)| \sim C\phi^{-n} \text{ if } n \to \infty \tag{14}$$

from some constant C, and that $\delta_2^{-1}|p_n(0)|^{-2} \leq |e_n| \leq \delta_1^{-1}|p_n(0)|^{-2}$ for each n (see e.g. [6] and [2]).

The theorem shows that the convergence of β_n towards β is of exponential type, and that the larger $\frac{\delta_2}{\delta_1}$ is, the smaller is the convergence rate. It is also quite interesting to remark that the convergence rate does not depend on the particular form of the measure σ but on its support only. One should however notice that in practice, the values of n for which the asymptotic regime (13) is reached certainly depends on σ: (13) holds if n is chosen in such a way that $\frac{|p_n(0)|}{\phi^{-n}}$ is close enough from the constant C. This point will be developed on an example below.

4. EXAMPLES

The CDMA Equal Power Case

As a first example, we treat the case presented in the introduction and considered by Honig and Xiao, and compute in closed form the corresponding orthogonal polynomials $p_n(x)$. The purpose is to assess the values on n for which the SNR loss $e_n(0) = G_\sigma(0) + (q_n(0)/p_n(0)) = \beta - \beta_n$ attains the asymptotic regime in n. Recall that this regime is described by equations (14) and (13).

We recall that in this context (see section 1) , $\mathbf{h}$ coincides with $\mathbf{w}$ and $\mathbf{R}_I = \mathbf{U}\mathbf{U}^H + \omega^2 \mathbf{I}_N$ where $\mathbf{W} = [\mathbf{w}\ \mathbf{U}]$ is a random $N \times K$ matrix with $1/N$ variance i.i.d. entries. The measure σ defined by $\int \lambda^k d\sigma(\lambda) = \lim_{N\to\infty, K/N\to\alpha} \mathbf{h}^H \mathbf{R}_I^k \mathbf{h}$ is the limit eigenvalue distribution of $\mathbf{R}_I$, *i.e.*, the so-called Marchenko-Pastur distribution. We only consider here the case $\alpha \geq 1$, because otherwise, σ is not absolutely continuous (σ has a mass at ω^2) so that the results of this paper cannot be applied. In this case, $\delta_1 = \omega^2 + (\sqrt{\alpha} - 1)^2$ and $\delta_2 = \omega^2 + (\sqrt{\alpha} + 1)^2$. We recall that the Stieltjes transform $G_\sigma(z)$ of this distribution is solution of the equation

$$G_\sigma(z) = \frac{-1}{z - \omega^2 - \frac{\alpha}{1 + G_\sigma(z)}} \,. \tag{15}$$

Moreover (see e.g. [5] and [9]), the sequence $(a_n)_{n\geq 0}$ is given by

$$a_0 = \omega^2 + \alpha, \ a_n = a = \omega^2 + \alpha + 1 \text{ for } n \geq 1 \tag{16}$$

while sequence $(b_n)_{n\geq 0}$ is reduced to $b_n = b = \sqrt{\alpha}$ for each n. It turns out that in this particular case, it is possible to express $p_n(0)$ in closed form, and thus to check for which values of n the inequality (13) holds.

In order to evaluate $p_n(0)$, we denote by $G_{\tilde\sigma}(z)$ the function defined by the continued fraction expansion

$$G_{\tilde\sigma(z)} = \frac{-1}{z - a - \frac{b^2}{z - a - \frac{b^2}{z - a \dots}}} \tag{17}$$

Using (11) and (16), it is easily seen that $G_\sigma(z)$ and $G_{\tilde\sigma}(z)$ are related by the relation

$$G_\sigma(z) = \frac{-1}{z - a_0 + b^2 G_{\tilde\sigma(z)}} \tag{18}$$

Moreover, $G_{\tilde\sigma}(z)$ satisfies the equation

$$G_{\tilde\sigma}(z) = \frac{-1}{z - a + b^2 G_{\tilde\sigma}(z)}$$

Solving this equation, and using the Stieljes inversion formula, we get that $\tilde\sigma$ coincides with the Wigner semi-circle law on $[\delta_1, \delta_2]$ with derivative

$$\tilde\sigma'(x) = \frac{2}{\pi} \times \frac{1}{(\delta_2 - \delta_1)/2} \times \sqrt{1 - \left(\frac{x - (\delta_2 + \delta_1)/2}{(\delta_2 - \delta_1)/2}\right)^2} \tag{19}$$

The associated orthogonal polynomials $\tilde{p}_n(x)$ are well known : $\tilde{p}_n(x) = \overline{p}_n(\frac{x-(\delta_2+\delta_1)/2}{(\delta_2-\delta_1)/2})$ where $\overline{p}_n(u)$ represents the second kind Tchebyschev polynomials given by

$$\overline{p}_n(u) = \frac{(z^{n+1} - z^{-(n+1)})}{(z - z^{-1})}$$

where z is defined as the solution of the equation $u = \frac{1}{2}(z + z^{-1})$ which satisfies $|z| \geq 1$. From this, we deduce that

$$\tilde{p}_n(0) = (-1)^n \frac{(\phi^{-(n+1)} - \phi^{(n+1)})}{(\phi^{-1} - \phi)} \tag{20}$$

Using the relation

$$-\frac{q_n(0)}{p_n(0)} = \frac{-1}{x - a_0 - b^2 \frac{\tilde{q}_{n-1}(0)}{\tilde{p}_{n-1}(0)}} \tag{21}$$

we get that

$$p_n(0) = \tilde{p}_n(0) + \frac{1}{\sqrt{\alpha}} \tilde{p}_{n-1}(0) \tag{22}$$

Therefore, the asymptotic regime (14) is reached for $p_n(0)$ if it is reached for $\tilde{p}_n(0)$ itself, i.e. if ϕ^{n-1} is negligible. Hence, (13) holds if ϕ^{2n} is small enough. This confirms that, although $N \to \infty$, close to optimal performance can be achieved by finite dimensional reduced rank Wiener filters (see [5]).

The unequal powers and the frequency selective channel cases

Let us now assume that the users in model (3) have different powers, resulting in the model $\mathbf{y} = \mathbf{W}\sqrt{\mathbf{P}}\mathbf{s} + \mathbf{v}$ where $\mathbf{P}$ is the power diagonal matrix. The interference signal covariance matrix is then $\mathbf{R}_I = \mathbf{U}\mathbf{Q}\mathbf{U}^H + \omega^2 \mathbf{I}_N$ where $\mathbf{Q}$ is the matrix which remains after extracting the first row and column of $\mathbf{P}$. The power empirical distribution is assumed to converge almost surely in distribution toward a distribution carried by the interval $[p_{\min}; p_{\max}]$ of $(0; +\infty)$ when $K \to \infty$. In this situation, the eigenvalue empirical distribution of $\mathbf{R}_I$ still converges to a certain probability distribution σ. However, there exist explicit expressions for σ only for a few particular power distributions. Notice that σ must have an almost everywhere strictly positive derivative on $[\delta_1, \delta_2]$ for our results to be true. This puts some slight constraints on the power limit distribution. Furthermore, it is still assumed that $\alpha \geq 1$.

The derivation will be restricted here to the asymptotic regime. More precisely, an upper bound for ϕ will be given. The support of the limit distribution of $\mathbf{U}\mathbf{U}^H$, which is the Marchenko-Pastur distribution, is $[(\sqrt{\alpha} - 1)^2; (\sqrt{\alpha} + 1)^2]$. It is easy to notice then that the support $[\delta_1; \delta_2]$ of σ satisfies $\delta_1 \geq p_{\min}(\sqrt{\alpha} - 1)^2 + \omega^2$ and $\delta_2 \leq p_{\max}(\sqrt{\alpha} + 1)^2 + \omega^2$, so $\frac{\delta_2}{\delta_1} \leq \xi$ with

$$\xi = \frac{p_{\max}(\sqrt{\alpha} + 1)^2 + \omega^2}{p_{\min}(\sqrt{\alpha} - 1)^2 + \omega^2} \ .$$

Now, ϕ can be written $\phi = ((\delta_2/\delta_1) - 1) / \left(\sqrt{\delta_2/\delta_1} + 1\right)^2$, and this increases with δ_2/δ_1. In the asymptotic regime, we thus have

$$\phi^{2n} \leq \left(\frac{\xi - 1}{\left(\sqrt{\xi} + 1\right)^2} \right)^{2n} \ . \tag{23}$$

The narrower is the power spread, the closer ξ is to 1, and the faster is the convergence rate. In particular, the best rate is obtained when powers are equal.

Finally, we consider the case where signals of all the users pass through a frequency selective channel in a downlink setting. This channel $h(z)$ is assumed to be a polynomial channel with a small degree relative to the spreading factor N. In this situation, the Inter Symbol Interference can be neglected (see [3] for the details) and the signal model can be approximated in the large system conditions by

$$\mathbf{y} = \mathbf{H}\mathbf{W}\sqrt{\mathbf{P}}\mathbf{s} + \mathbf{v},$$

where $\mathbf{H}$ is a Toeplitz matrix associated to $h(z)$. The interference and noise covariance matrix is now $\mathbf{R}_I = \mathbf{H}\mathbf{U}\mathbf{Q}\mathbf{U}^H\mathbf{H}^H + \omega^2 \mathbf{I}_N$. The matrix $\mathbf{H}\mathbf{H}^H$, being a Toeplitz matrix associated to the spectral density $|h(e^{2j\pi f})|^2$, has an eigenvalue limit distribution supported by $[\min_f \left(|h(e^{2j\pi f})|^2\right)$;
$\max_f \left(|h(e^{2j\pi f})|^2\right)]$. (23) is still true if we put

$$\xi = \frac{\max_f \left(|h(e^{2j\pi f})|^2\right) p_{\max}(\sqrt{\alpha} + 1)^2 + \omega^2}{\min_f \left(|h(e^{2j\pi f})|^2\right) p_{\min}(\sqrt{\alpha} - 1)^2 + \omega^2} \ .$$

Frequency selectivity slowers the convergence of β_n.

5. REFERENCES

[1] N.I. Akhiezer, *The Classical Moment Problem*, Oliver & Boyd.

[2] A. Bultheel, C. Díaz-Mendoza, P. González-Vera, and R. Orive, "Quadrature on the Half Line and Two-Point Padé Approximants to Stieltjes Functions. Part II : Convergence", *J. Comput. Appl. Math.*, 65, 1996.

[3] J.M. Chaufray, Ph. Loubaton, and W. Hachem, "Asymptotical Analysis of Optimum and Sub-Optimum CDMA Downlink MMSE Receivers", *ISIT 2002*, Lausanne, Switzerland.

[4] J.S. Goldstein, I.S. Reed, and L.L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections", *IEEE Trans. on Information Theory*, Vol. 44, No. 7, November 1998.

[5] M.L. Honig and W. Xiao, "Performance of Reduced-Rank Linear Interference Suppression", *IEEE Trans. on Information Theory*, Vol. 47, No. 5, pp. 1928–1946, July 2001.

[6] H. Stahl and V. Totik, *General Orthogonal Polynomials*, Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992.

[7] G. Szegö, *Orthogonal Polynomials*, Vol. 33 of *Amer. Math. Soc. Colloq. Publ.*, AMS.

[8] L.G. F. Trichard, J.S. Evans and I.B. Collings,"Optimal Linear Multistage Receivers for Synchronous CDMA", Proceedings of IEEE International Conference on Communications, New York City, New York, USA, Apr. 2002, pp. 1461-1465.

[9] L.G.F. Trichard, J.S. Evans, and I.B. Collings, "Optimal Linear Multistage Receivers and the Recursive Large System SIR", *ISIT 2002*, Lausanne, Switzerland.