Asymptotic Moments for Interference Mitigation in Correlated Fading Channels
Jakob Hoydis, Merouane Debbah, Mari Kobayashi

To cite this version:

HAL Id: hal-00648020
https://hal-supelec.archives-ouvertes.fr/hal-00648020
Submitted on 4 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Asymptotic Moments for Interference Mitigation in Correlated Fading Channels

Jakob Hoydis†‡, Mériouane Debbah†, and Mari Kobayashi‡

†Department of Telecommunications, Supélec, 91192 Gif-sur-Yvette, France
‡Alcatel-Lucent Chair on Flexible Radio, Supélec, 91192 Gif-sur-Yvette, France
{jakob.hoydis, meriouane.debbah, mari.kobayashi}@supelec.fr

Abstract—We consider a certain class of large random matrices, composed of independent column vectors with zero mean and different covariance matrices, and derive asymptotically tight deterministic approximations of their moments. This random matrix model arises in several wireless communication systems of recent interest, such as distributed antenna systems or large antenna arrays. Computing the linear minimum mean square error (LMMSE) detector in such systems requires the inversion of a large covariance matrix which becomes prohibitively complex as the number of antennas and users grows. We apply the derived moment results to the design of a low-complexity polynomial expansion detector which approximates the matrix inverse by a matrix polynomial and study its asymptotic performance. Simulation results corroborate the analysis and evaluate the performance for finite system dimensions.

I. INTRODUCTION

Distributed antenna systems and large antenna arrays have recently attained significant research interest [1], [2]. Both are considered as promising solutions to counter intercell interference and to increase the spectral efficiency of current cellular networks. Since these techniques rely in essence on a significant increase of the number of coordinated antennas, the computational complexity of the joint precoding/detection of the transmitted/received signals grows. This calls for low-complexity solutions. In this paper, we address this need by assessing the performance of a polynomial expansion detector [3] adapted to the following general channel model.

Consider a discrete-time $N \times K$ multiple-input multiple-output (MIMO) channel with output vector $y \in \mathbb{C}^N$:

$$y = Hx + n$$  \hspace{1cm} (1)

where $x = [x_1, \ldots, x_K]^\top$ is the complex channel input vector satisfying $\mathbb{E}[xx^\dagger] = I_K$, $H = [h_1 \cdots h_K] \in \mathbb{C}^{N \times K}$ is the random channel matrix and $n \sim \mathcal{CN}(0, \sigma^2 I_N)$ is a vector of additive noise. The $j$th column $h_j \in \mathbb{C}^N$ of $H$ is modeled as

$$h_j = \frac{1}{\sqrt{K}} R_j w_j, \hspace{0.5cm} j = 1, \ldots, K$$  \hspace{1cm} (2)

where $R_j \in \mathbb{C}^{N \times N}$ is a deterministic matrix and the elements of $w_j \in \mathbb{C}^N$ are independent and identically distributed (i.i.d.) random variables with zero mean, unit variance and finite eighth moment. This channel model captures different types of wireless communication systems and generalizes several well-known channel models as discussed below:

### Distributed Antenna Systems

Let $R_j = \text{diag}(r_{1j}, \ldots, r_{Nj})$ with elements $r_{ij} = \sqrt{p_j/d_{ij}^{\beta}}$, where $d_{ij}$ is the (normalized) distance between transmitter $j$ and receiver antenna $i$, $\beta$ is the path loss exponent and $p_j$ is the transmit power of transmitter $j$. This model is suitable for distributed antenna systems [1] where each transmitter sees a different path loss to each of the receive antennas since $d_{1j}, \ldots, d_{Nj}$ are different.

### Large-scale MIMO

Assume a receiver equipped with a very large antenna array ($N \gg 1$) as in [2]. Unless the antenna spacing is sufficiently large, it is likely that the received signals at different receive antennas are correlated. Our model allows to assign a different correlation matrix $R_j$ to each transmitter.

### MIMO Multiple Access Channel (MAC)

Consider a MIMO MAC from $M$ transmitters equipped with $K_m, m = 1, \ldots, M$, antennas to a receiver with $N$ antennas. Each point-to-point link has a different transmit and receive correlation matrix [4]:

$$y = \sum_{m=1}^{M} \Phi_{R,m}^\frac{1}{2} W_m \Phi_{T,m}^\frac{1}{2} x_m + n$$

where $\Phi_{R,1}, \ldots, \Phi_{R,M} \in \mathbb{C}^{N \times N}$ are deterministic correlation matrices, $\Phi_{T,1} \in \mathbb{C}^{K_1 \times K_1}, \ldots, \Phi_{T,M} \in \mathbb{C}^{K_M \times K_M}$ are nonnegative diagonal matrices, $W_1 \in \mathbb{C}^{N \times K_1}, \ldots, W_M \in \mathbb{C}^{N \times K_M}$ are random channel matrices with i.i.d. entries with zero mean and variance $1/K$, and $x_m \in \mathbb{C}^{K_m}, x_m \in \mathbb{C}^{K_m}$ are the transmit vectors. Let $\sum_{m=1}^{M} K_m = K$. Setting $R_j = \Phi_{R,m}^{1/2} [\Phi_{T,m}]_i$ for $j \in \{1 + \sum_{l=1}^{m-1} K_l, \ldots, \sum_{l=1}^{m} K_l\}$ and $i = j - \sum_{l=1}^{m-1} K_l$, we fall back to the model in (2).

In the sequel, we will study the asymptotic behavior of the moments $\mu_n$ of the matrix $B \overset{\triangle}{=} HH^\dagger$, defined as

$$\mu_n \overset{\triangle}{=} \frac{1}{N} \text{tr} B^n, \hspace{0.5cm} n = 0, 1, 2, \ldots$$  \hspace{1cm} (3)

under the assumption that $N$ and $K$ grow infinitely large at the same speed. In particular, we will derive deterministic approximations $\mu_n$ of $\mu_n$, such that $\mu_n - \mu_n \to 0$ almost surely, for $N, K \to \infty$. This result can be used, for example, to compute low-complexity approximations of the matrix inverse $(B + \sigma^2 I_N)^{-1}$. The computation of this matrix arises in many practical applications, such as for linear multiuser detectors and beamforming strategies. We will focus exemplary on the linear minimum mean square error (LMMSE) detector.
The LMMSE estimate \( \hat{x} \) of \( x \), assuming perfect knowledge of \( H \) at the receiver, is given as [5]

\[
\hat{x} = H^H(B + \sigma^2 I_N)^{-1} y. \tag{4}
\]

The computational complexity of this estimate is of order \( O(r^2) \) [6], where \( r = \min(N, K) \). A reduced complexity estimate can be obtained by approximating the matrix inverse in (4) by the following matrix polynomial [3]

\[
(B + \sigma^2 I_N)^{-1} \approx \sum_{l=0}^{L-1} w_l B^l \tag{5}
\]

for some coefficients \( w_l \), where the filter rank \( L \leq r \) is chosen according to the allowable complexity. For a given transmitter \( k \), the above polynomial expansion detector can be seen as a projection of \( y \) on the \( L \)th Krylov subspace associated to the pair \((B, h_k)\), i.e., the subspace of \( \mathbb{C}^N \) spanned by the vectors \( \{h_k, Bh_k, \ldots, B^{L-1}h_k\} \), and a weighting of the joint projections by the coefficients \( w_l \). Depending on \( L \), the polynomial expansion detector achieves a performance between the matched filter \((L = 1)\) and the LMMSE detector \((L = r)\) [3] and allows, consequently, to trade-off performance for complexity. Moreover, (5) allows for an efficient multistage implementation [3], [7], [6], where each stage \( l \) consists of a matched filter \( H^l \) and subsequent "re-spreading" by the matrix \( H \). In [8], it was shown that the signal-to-interference-plus-noise ratio (SINR) at the filter output converges in certain cases exponentially in the filter rank \( N \) to the SINR of the matched filter \((L = 1)\) and the LMMSE detector \((L = r)\) [3] and allows, consequently, to trade-off performance for complexity. Moreover, (5) allows for an efficient multistage implementation [3], [7], [6], where each stage \( l \) consists of a matched filter \( H^l \) and subsequent "re-spreading" by the matrix \( H \). In [8], it was shown that the signal-to-interference-plus-noise ratio (SINR) at the filter output converges in certain cases exponentially in the filter rank \( N \) to the SINR of the matched filter \((L = 1)\) and the LMMSE detector \((L = r)\) [3] and allows, consequently, to trade-off performance for complexity. Moreover, (5) allows for an efficient multistage implementation [3], [7], [6], where each stage \( l \) consists of a matched filter \( H^l \) and subsequent "re-spreading" by the matrix \( H \). In [8], it was shown that the signal-to-interference-plus-noise ratio (SINR) at the filter output converges in certain cases exponentially in the filter rank \( N \) to the SINR of the matched filter \((L = 1)\) and the LMMSE detector \((L = r)\) [3] and allows, consequently, to trade-off performance for complexity. Moreover, (5) allows for an efficient multistage implementation [3], [7], [6], where each stage \( l \) consists of a matched filter \( H^l \) and subsequent "re-spreading" by the matrix \( H \). In [8], it was shown that the signal-to-interference-plus-noise ratio (SINR) at the filter output converges in certain cases exponentially in the filter rank \( N \) to the SINR of the matched filter \((L = 1)\) and the LMMSE detector \((L = r)\) [3] and allows, consequently, to trade-off performance for complexity. Moreover, (5) allows for an efficient multistage implementation [3], [7], [6], where each stage \( l \) consists of a matched filter \( H^l \) and subsequent "re-spreading" by the matrix \( H \).
Denote by $S$ the class of functions $f$ analytic over $\mathbb{C}\setminus \mathbb{R}_+$, such that, for $z \in \mathbb{C}_+, f \in \mathbb{C}_+, zf \in \mathbb{C}_+$ and $\lim_{y \to \infty} -iy f'(iy) < \infty$. Such functions are known to be Stieltjes transforms of finite measures supported by $\mathbb{R}_+$ [13, Theorem 2.2].

**Theorem 1 (14, Theorem 1):** Let $D \in \mathbb{C}^{N \times N}$ be a Hermitian non-negative definite matrix and assume that $D$ and the matrices $R_j$, $j = 1, \ldots, K$, have uniformly bounded spectral norms (with respect to $N$). Let $N, K \to \infty$, such that $0 < \lim \inf \frac{K}{N} \leq \lim \sup \frac{K}{N} < \infty$. Then, for any $z \in \mathbb{C} \setminus \mathbb{R}_+$,

$$\frac{1}{N} \text{tr} D (B - zI_N)^{-1} - \frac{1}{N} \text{tr} DT(z) \xrightarrow{a.s.} 0$$

where $T(z) \in \mathbb{C}^{N \times N}$ is defined as

$$T(z) = \left( \frac{1}{K} \sum_{j=1}^{K} R_j R_j^H T(z) \right)^{-1}$$

and the following set of $K$ implicit equations

$$\delta_j(z) = \frac{1}{K} \text{tr} R_j R_j^H T(z), \quad j = 1, \ldots, K$$

admits a unique solution $(\delta_1(z), \ldots, \delta_K(z)) \in \mathbb{R}^K$. Moreover, denote by $F$ the distribution function whose Stieltjes transform is given by $m(z) = \frac{1}{N} \text{tr} T(z)$. Then, almost surely,

$$F^B - F \to 0.$$

**III. Asymptotic Moments**

In this section, we state our main results. Outlines of the proofs of Theorems 2 and 3 are provided in the appendix.

**Theorem 2:** Let $F$ be the distribution function as defined in Theorem 1 and denote by $\overline{\pi}_0, \overline{\pi}_1, \ldots$ the successive moments of $F$, defined as $\overline{\pi}_n \doteq \int_0^\infty \lambda^n dF(\lambda)$. These moments can be calculated as

$$\overline{\pi}_n = \frac{(-1)^n}{n!} \frac{1}{N} \text{tr} T_n$$

where $T_n$ is defined recursively by the following set of equations for $n \geq 0$:

$$T_{n+1} = \sum_{i=0}^{n} \sum_{j=0}^{i} \binom{n}{i} \binom{i}{j} T_{n-i} Q_{i,j} T_j$$

$$Q_{n+1} = \frac{n + 1}{K} \sum_{k=1}^{K} f_{k,n} R_k R_k^H$$

$$f_{k,n+1} = \sum_{i=0}^{n} \sum_{j=0}^{i} \binom{n}{i} \binom{i}{j} (n - i + 1) f_{k,j} f_{k,i-j} \delta_{k,n-i}$$

$$\delta_{k,n+1} = \frac{1}{K} \text{tr} R_k R_k^H T_{n+1}$$

where $T_0 = I_N$, $f_{k,0} = -1$ and $\delta_{k,0} = \frac{1}{K} \text{tr} R_k R_k^H \forall k$.

**Remark 3.1:** While Theorem 2 allows to compute the moments $\overline{\pi}_n$ of $F$, it does not imply the a.s. convergence of $\mu_n$ and $\overline{\mu}_n$ in general. Theorem 3 provides some sufficient conditions for which this convergence holds.

**Remark 3.2:** Although difficult to show analytically, one can verify numerically that Theorem 2 coincides with [10, Theorem 1] for $R_j = \text{diag}(T_{ij}, \ldots, T_{nj})$, $j = 1, \ldots, K$.

If the matrices $R_j$ are drawn from a finite set of matrices, we get the following stronger result:

**Theorem 3:** For fixed $M > 0$, let $\mathcal{R} = \{ \tilde{R}_1, \ldots, \tilde{R}_M \}$ be a set of complex $N \times N$ matrices and let $D \in \mathbb{C}^{N \times N}$ be a non-negative definite Hermitian matrix. Assume that $D$ and $R_m$, $m = 1, \ldots, M$, have uniformly bounded spectral norms (with respect to $N$). Let $R_j \in \mathcal{R}$ for $j = 1, \ldots, K$. Assume $N, K \to \infty$, such that $0 < \lim \inf \frac{K}{N} \leq \lim \sup \frac{K}{N} < \infty$. Then, for $n = 0, 1, 2, \ldots$,

$$\frac{1}{N} \text{tr} DB^n - \frac{(-1)^n}{n!} \frac{1}{N} \text{tr} DT_n \xrightarrow{a.s.} 0$$

where $T_n$ is given by Theorem 2. This implies in particular,

$$\mu_n - \overline{\mu}_n \xrightarrow{a.s.} 0.$$

Loosely speaking, Theorem 1 states that, for large matrix dimensions, the e.s.d. $F^B$ of the matrix $B$ can be closely approximated by a deterministic distribution function $F$. Thus, the optimal weighting vector $w$ can be approximated by replacing the moments $\mu_n$ of $F^B$ in (8) by the moments $\overline{\mu}_n$ of $F$. Using the result of Theorem 2, we can compute an approximate weight vector $\mathbf{w} = [\overline{\mu}_0, \ldots, \overline{\mu}_{L-1}]$ as

$$\mathbf{w} = \overline{\Phi}^{-1} \overline{\varphi}$$

where $\overline{\Phi} \in \mathbb{R}_+^{L \times L}$ and $\overline{\varphi} \in \mathbb{R}_+^{L}$ are defined by

$$[\overline{\Phi}]_{ij} = \overline{\pi}_{i+j} + \sigma^2 \overline{\pi}_{i+j-1}$$

$$[\overline{\varphi}]_i = \overline{\pi}_i.$$

**IV. Asymptotic Performance Analysis**

We consider now the asymptotic performance of the polynomial expansion receiver in terms of the received SINR $\gamma_k$ for a given transmitter $k$. With weight vector $\mathbf{w}$, the $k$th element $\hat{x}_k$ of the estimated vector $\hat{x}$ reads

$$\hat{x}_k = h_k^H \sum_{l=0}^{L-1} w_l B^l (Hx + n).$$

One can easily show that the associated SINR $\gamma_k$ can be expressed as [6, Eq. (18)]

$$\gamma_k = \frac{\mathbf{w}^T \varphi_k \varphi_k^T \mathbf{w}}{\mathbf{w}^T (\overline{\Phi}_k - \varphi_k \varphi_k^T) \mathbf{w}}$$

where $\overline{\Phi}_k \in \mathbb{R}_+^{L \times L}$ and $\varphi_k \in \mathbb{R}_+^{L}$ are given as

$$[\overline{\Phi}_k]_{ij} = [B^{i+j}]_{kk} + \sigma^2 [B^{i+j-1}]_{kk}$$

$$[\varphi_k]_i = [B^i]_{kk}.$$

The next theorem provides a tight deterministic approximation of the terms $[B^n]_{kk} = h_k^H B^{n-1} h_k$ in the asymptotic limit.
Let $\mathbf{R}_j = \Theta_j^{1/2}$ and $\Theta_j \in \mathbb{C}^{N \times N}$ be defined as
\[
(\Theta_j)_{kl} = \frac{1}{\phi_{\max} - \phi_{\min}} \int_{\phi_{\min}}^{\phi_{\max}} \exp \left( \frac{2\pi i}{\lambda} d_{kl} \cos(x) \right) dx
\]
where $d_{kl} = 2\lambda(k-l)$ and $\phi_{\min}^j, \phi_{\max}^j$ are drawn independently from the intervals $[-\pi, 0]$ and $[0, \pi]$, respectively. The interval $[\phi_{\min}^j, \phi_{\max}^j]$ can be seen as the angular spread of the signal from transmitter $j$, $\lambda$ is the wave length, and $d_{kl}$ is the spacing between the receive antennas $k$ and $l$. We assume Rayleigh fading channels, i.e., $w_j$ in (2) are independent standard complex Gaussian vectors. The covariance matrices $\Theta_j$ are chosen at random at the beginning and then kept fixed while we average over many realizations of the channel matrix $\mathbf{H}$. We denote by $\text{SNR} = 1/\sigma^2$ the transmit signal-to-noise ratio.

Fig. 1 shows the average received SINR $\mathbf{E}[\gamma_k]$ of a randomly chosen transmitter as a function of the SNR for the matched filter, the LMMSE detector and the polynomial expansion detector with approximate weights for $L = \{2, 3, 6\}$. Markers correspond to simulation results and solid lines to the deterministic SINR approximations. The error bars indicate one standard deviation in each direction.

Fig. 2 depicts the theoretical average bit error rate (BER) over SNR for the different detectors. Assuming binary phase-shift keying (BPSK) modulation and Gaussian interference, the BER is given as $\mathbf{E}[Q(\sqrt{\gamma_k})]$ where $Q(x)$ is the Gaussian tail function. We can clearly see a performance increase of the polynomial expansion detector with $L$, although the BER saturates at high SNR. Although not explicitly shown here, one can even observe a performance decrease for large values of $L$. As mentioned before, this is due to the low accuracy of the approximate weights caused by a slow convergence of the higher-order moments to their deterministic approximations.
VI. CONCLUSION

We have derived asymptotically tight deterministic approximations of the moments of a certain class of large random matrices, useful for the study of distributed antenna systems and large antenna arrays. We have applied these moment results to the design of a polynomial expansion detector which significantly reduces the computational complexity of multiuser detection compared to the LMMSE detector. Moreover, we have derived an explicit expression of the asymptotic SINR at the output of this detector and verified its accuracy and performance for finite system dimensions by simulations.

APPENDIX

Outline of the proof of Theorem 2: From Definition 2, it is easy to see that the moments \( \bar{\mu}_n \) of the distribution function \( F \) can be obtained through successive differentiation of the function \( \frac{1}{z}m(-\frac{1}{z}) \), i.e.,

\[
\bar{\mu}_n = \frac{(-1)^n}{n!} \frac{d^n}{dz^n} \left( \frac{1}{z} m \left( \frac{1}{z} \right) \right) \bigg|_{z=0} = \frac{(-1)^n}{n!} \int \frac{d^n}{dz^n} \left( \frac{1}{z\lambda + 1} \right) dF(\lambda) \bigg|_{z=0} = \int \lambda^n dF(\lambda).
\]

Consider now the function \( \eta(z) = \frac{1}{z} m(-\frac{1}{z}) \) for \( z \geq 0 \) and denote by \( \eta_n(z) \) its \( n \)th derivative. From Theorem 1, we have

\[
\eta(z) = \frac{1}{N} \text{tr} T_0(z)
\]

where

\[
T_0(z) = \left( z \frac{1}{K} \sum_{j=1}^{K} \frac{R_j R_j^H}{1 + z \delta_{j,0}(z)} + I_N \right)^{-1}
\]

and \( (\delta_{1,0}(z), \ldots, \delta_{K,0}(z)) \in \mathbb{R}_+^K \) is the unique solution to the \( K \) implicit equations:

\[
\delta_{j,0}(z) = \frac{1}{K} \text{tr} R_j R_j^H T_0(z), \quad j = 1, \ldots, K.
\]

Denoting \( T_n(z) = \frac{d^n T_0(z)}{dz^n} \), we have \( \eta_n(z) = \frac{1}{n} \text{tr} T_n(z) \). Explicit expressions of \( T_n(z) \) can be found by re-writing \( T_0(z) \) with the help of some auxiliary functions and repeated use of the Leibniz-rule for the \( n \)th derivative of the product of two functions, i.e.,

\[
\frac{d^n}{dz^n}(uv)(z) = \sum_{i=0}^{n} \binom{n}{i} \frac{d^{n-i}u(z)}{dz^{n-i}} \frac{d^i v(z)}{dz^i}.
\]

The resulting set of implicit equations simplifies to a system of recursive equations for \( z = 0 \). One can easily see that \( T_0(0) = T_0 = I_N \) and, consequently, \( \delta_{j,0}(0) = \delta_{j,0} = \frac{1}{K} \text{tr} R_j R_j^H \forall j \).

Outline of the proof of Theorem 3: Both \( \frac{1}{N} \text{tr} D(B - 2I_N)^{-1} \) and \( \frac{1}{N} \text{tr} D^2 T(z) \) as defined in Theorem 1 are Stieltjes transforms of finite measures which we denote by \( \pi \) and \( \tau \). Theorem 1 implies that, almost surely, \( \pi - \tau \to 0 \). Similar to the proof of Theorem 2 the moments of \( \pi \) and \( \tau \) can be respectively expressed as

\[
\int \lambda^n \pi(d\lambda) = \frac{1}{N} \text{tr} DB^n
\]

and

\[
\int \lambda^n \tau(d\lambda) = \frac{1}{n!} \frac{1}{N} \text{tr} DT_n.
\]

The a.s. weak convergence of \( \pi \) and \( \tau \) implies by [16, Theorem 25.8 (iii)] that \( \int f(\lambda) \tau(d\lambda) = \int f(\lambda) \pi(d\lambda) \to 0 \), for any bounded continuous function \( f(\lambda) \). Relying on [17], one can prove that the support of \( \pi \) is almost surely compact since \( D \) has bounded spectral norm and the spectral norm of \( B \) can be shown to be almost surely bounded. Following similar steps as in [4, Proof of Theorem 2, Part B], one can also show that \( \tau \) has compact support. Thus, we can relax the assumption of \( f(\lambda) \) to be bounded and choose \( f(\lambda) = \lambda^n \) to establish the a.s. convergence of the moments of \( \pi \) and \( \tau \).

REFERENCES


