

Performance Analysis of Block and Comb Type Channel Estimation for Massive MIMO Systems

Gábor Fodor^{*†}, Piergiuseppe Di Marco[†], Miklós Telek[‡]

^{*} Ericsson Research, Stockholm, Sweden. E-mail: gabor.fodor@ericsson.com

[†] Royal Institute of Technology, Stockholm, Sweden. E-mail: pidm@kth.se

[‡] Budapest University of Technology and Economics, Budapest, Hungary. E-mail: telek@hit.bme.hu

Abstract—For pilot sequence based multiple input multiple output (MIMO) channel estimation, the arrangements of pilot symbols, such as the *block* or *comb* type arrangement, is known to play an important role. In this paper we compare the performance of block and comb pilot symbol patterns in terms of uplink mean square error (MSE) and spectral efficiency when the receiver at the base station employs least square (LS) or minimum mean square error (MMSE) channel estimation and MMSE equalizer for uplink data reception. For this system, we derive a closed form solution for the MSE and spectral efficiency that allows us to obtain exact results for an arbitrary number of antennas. Our key observation is that the comb pilot arrangement allows for unequal pilot-data power allocation in the frequency domain, which leads to a significant spectral efficiency increase. This spectral efficiency increase is particularly important with LS estimation and as the number of base station antennas grows large. It also gives noticeable gains with MMSE estimation. Our main conclusion is that with a large number of antennas, unequal power allocation facilitated by comb arrangement can give large gains over alternative pilot arrangements.¹

I. INTRODUCTION

Since the seminal work by Hassibi and Hochwald [1], a number of papers investigated the trade-off between the resources used for channel state information (CSI) acquisition and data transmission. For example, assuming a block fading reciprocal channel, a finite number of symbols in the time and frequency domains are available for CSI acquisition, and uplink as well as downlink precoding and data transmission [2], [3]. Also, under a fixed power budget, pilot symbols reduce the transmitted energy for data symbols, as it has been pointed out in [4] and [5] where the optimal pilot-to-data power ratio (PDPR) is investigated for various pilot patterns and receiver structures. The results of [5], for example, indicate that the optimal PDPR provides about 2-3 dB gain compared with equal power for pilot and data symbols. Subsequently, [6] derived a closed form of the optimal PDPR for minimum mean square error (MMSE) channel estimation and showed that a tight bound lying in the quasi-optimal region provides a good approximation for the optimal PDPR. More recently, [7] derived a closed form PDPR that maximizes the capacity bound

¹The work of G. Fodor has been partially performed in the framework of the FP7 project ITC 317669 METIS. G. Fodor has also been supported by the Swedish Foundation for Strategic Research Strategic Mobility SM13-0008 Matthew Project. The work of M. Telek was partially performed when he was visiting the Royal Institute of Technology, Stockholm. M. Telek was partially supported by the OTKA K101150 project.

of MIMO-OFDM systems and studied the impact of carrier frequency offset (CFO) on maximizing power allocation.

In [8] and [9], equivalence conditions for the achievable spectral efficiency between block-fading channels and continuous-fading channels are discussed. An approximate closed-form analytical expression of the spectral efficiency is derived in the hypothesis of MMSE estimation.

In our previous work [10], we investigated the effects of the PDPR on the MSE, assuming a single pilot and a single data symbol under a fixed sum power budget with least square (LS) channel estimation at the base station (BS). While [10] provides insight into the trade-off related to PDPR, it does not consider the trade-off related to the number of symbols used for CSI acquisition and data transmission, which is critical for the spectral efficiency. Therefore, the purpose of the present paper is to devise a methodology to find the optimum number of pilot and data symbols *and* the optimum PDPR. It turns out that the constraints for these trade-offs depend on the pilot pattern that is used in the time and frequency domains. Specifically, the so called *block* type arrangement dedicates all frequency channels within a given time slot to either channel estimation or data transmission whereas the *comb* pilot pattern employs pilot and data symbols mixed in the frequency domain within a single time slot.

The design of the uplink demodulation reference signals (DMRS) specifically in 3GPP Long Term Evolution Advanced (LTE-A) systems is described in [11]. In the LTE uplink, DMRS:s are used to facilitate channel estimation for the coherent demodulation of the physical uplink shared and control channels. The LTE DMRS:s occupy specific OFDM symbols within the uplink subframe according to the block type arrangement and support a large number of user equipment utilizing cyclic extensions of the well known Zadoff-Chu sequences [12].

Our key contribution in the present paper is the derivation of a closed form solution for both the uplink data MSE and spectral efficiency specifically taking into account the constraints of the comb and block type pilot arrangements. As a major difference with respect to previous works (i.e., [1]–[9]), this closed form result allows us to find the close-to-optimum number of pilot symbols and pilot power for a generic channel estimation method. In particular, we compare LS and MMSE channel estimation in block-type and comb-type pilot arrangement, for a BS employing a large number of antennas.

This approach enables us to arrive at some insights that we believe are novel in the massive MIMO literature.

The system model, including the description of the block and comb type pilot patterns, is defined in Sect. II. Next, Sect. III, discusses the LS and MMSE channel estimation algorithms specifically in block or comb systems and, in Sect. IV, we introduce the the MSE and spectral efficiency for the equalized uplink data symbols assuming linear MMSE reception. Our main analytical results are derived in Sect. V. Sect. VI presents numerical results and Sect. VII concludes the paper.

II. SYSTEM MODEL

We consider the uplink transmission of a multi-antenna single cell wireless system, in which users are scheduled on orthogonal frequency channels. It is assumed that each mobile station (MS) employs an orthogonal pilot sequence, so that no interference between pilots is present in the system. This is a common assumption in massive multi-user MIMO systems in which a single MS may have a single antenna. The base station (BS) estimates the channel \mathbf{h} (column vector of dimension N_r , where N_r is the number of receive antennas at the BS) by either LS or MMSE channel estimation to initialize an MMSE equalizer for uplink data reception. Since we assume orthogonal pilot sequences, the channel estimation process can be assumed independent for each MS. We consider a time-frequency resource of T time slots in the channel coherence time, and F subcarriers in the coherence bandwidth, with a total number of symbols $\tau = F \cdot T$. We denote by τ_p the number of symbols allocated to pilots, and by τ_d the number of symbols allocated to data ($\tau_p + \tau_d = \tau$). Moreover, we consider a transmission power level P_p and P for each pilot and data symbol, respectively. With this setup, we consider two pilot symbol allocation methods, namely block type and comb type, which we discuss in the following subsections.

A. Block Type Pilot Allocation

The block type pilot arrangement consists of allocating one or more time slots for pilot transmission, by using all subcarriers in those time slots. This approach is a suitable strategy for slow time-varying channels. Given T slots, a fraction of T_p slots are allocated to the pilot and $T_d = T - T_p$ slots are allocated to the data symbols. Note that a maximum transmission power P_{tot} is allowed in each time slot, among all F subcarriers. This power constraint is then identical for both the pilot (P_p) and data power (P), i.e.,

$$FP_p \leq P_{tot} \quad FP \leq P_{tot}. \quad (1)$$

The power cannot be traded between pilot and data, but the energy budget can be distributed by tuning the number of time slots T_p and T_d , i.e., $\tau_p = FT_p$ and $\tau_d = FT_d$.

B. Comb Type Pilot Allocation

In the comb type pilot arrangement a certain number of subcarriers are allocated to pilot symbols, continuously in time. This approach is a suitable strategy for non-frequency selective channels. Given F subcarriers in the coherence bandwidth, a

fraction of F_p subcarriers are allocated to the pilot and $F_d = F - F_p$ subcarriers are allocated to the data symbols.

Each MS transmits at a constant power P_{tot} , however, the transmission power can be distributed unequally in each subcarrier. In particular, if we consider a transmitted power P_p for each pilot symbol and P for each data symbol transmission, the following constraint is enforced:

$$F_p P_p + (F - F_p) P = P_{tot}. \quad (2)$$

The total number of symbols for pilots is $\tau_p = TF_p$ and for data is $\tau_d = TF_d$. However, with comb type pilot arrangement, the trade-off between pilot and data signals includes the trade-offs between the number of frequency channels *and* between the transmit power levels, which is an additional degree of freedom compared with the block type arrangement.

III. CHANNEL ESTIMATION

Let us consider a MS that transmits an orthogonal pilot sequence $\mathbf{s} = [s_1, \dots, s_{\tau_p}]^T$, where each symbol is scaled as $|s_i|^2 = 1$, for $i = 1, \dots, \tau_p$. Thus, the $N_r \times \tau_p$ matrix of the received pilot signal at the BS from the MS is

$$\mathbf{Y}^p = \alpha \sqrt{P_p} \mathbf{h} \mathbf{s}^T + \mathbf{N}, \quad (3)$$

where we assume that \mathbf{h} is a circular symmetric complex normal distributed vector of r.v. with mean vector $\mathbf{0}$ and covariance matrix \mathbf{C} (of size N_r), denoted as $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$, α accounts for the propagation loss, $\mathbf{N} \in \mathbb{C}^{N_r \times \tau_p}$ is the spatially and temporally additive white Gaussian noise (AWGN) with element-wise variance σ^2 .

In this paper, we consider two techniques, i.e., the least square (LS) and the minimum mean-square error (MMSE) channel estimation that are detailed in the following subsections.

A. LS Estimation

Conventional LS estimation relies on correlating the received signal with the known pilot sequence. The BS estimates the channel based on (3) assuming

$$\hat{\mathbf{h}}_{LS} = \mathbf{h} + \tilde{\mathbf{h}}_{LS} = \frac{1}{\alpha \sqrt{P_p}} \mathbf{Y}^p \mathbf{s}^* (\mathbf{s}^T \mathbf{s}^*)^{-1} = \mathbf{h} + \frac{1}{\alpha \sqrt{P_p} \tau_p} \mathbf{N} \mathbf{s}^*. \quad (4)$$

Note that $\mathbf{N} \mathbf{s}^* = [\sum_{i=1}^{\tau_p} s_i^* n_{i,1}, \dots, \sum_{i=1}^{\tau_p} s_i^* n_{i,N_r}]^T$, then $\mathbf{N} \mathbf{s}^* \sim \mathcal{CN}(\mathbf{0}, \tau_p \sigma^2 \mathbf{I}_{N_r})$.

By considering $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$, it follows that the estimated channel $\hat{\mathbf{h}}_{LS}$ is a circular symmetric complex normal distributed vector $\hat{\mathbf{h}}_{LS} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{LS})$, with

$$\mathbf{R}_{LS} = \mathcal{E}\{\hat{\mathbf{h}}_{LS} \hat{\mathbf{h}}_{LS}^H\} = \mathbf{C} + \frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r}. \quad (5)$$

The channel estimation error is defined as $\tilde{\mathbf{h}}_{LS} = \mathbf{h} - \hat{\mathbf{h}}_{LS}$, so that $\tilde{\mathbf{h}}_{LS} \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_{LS})$ with

$$\mathbf{W}_{LS} = \frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r}$$

and the estimation mean square error (MSE) is derived as

$$\varepsilon_{LS} = \mathcal{E}\{\|\tilde{\mathbf{h}}_{LS}\|_F^2\} = \text{tr}\{\mathbf{W}_{LS}\} = \frac{N_r \sigma^2}{\alpha^2 P_p \tau_p}, \quad (6)$$

where $\|\cdot\|_F$ is the Frobenius norm.

B. MMSE Estimation

We define a training matrix $\mathbf{S} = \mathbf{s} \otimes \mathbf{I}_{N_r}$ (of size $\tau_p N_r \times N_r$), so that $\mathbf{S}^H \mathbf{S} = \tau_p \mathbf{I}_{N_r}$. The $\tau_p N_r \times 1$ vector of received signal (3) can be conveniently rewritten as

$$\tilde{\mathbf{Y}}^p = \alpha \sqrt{P_p} \mathbf{S} \mathbf{h} + \tilde{\mathbf{N}}.$$

where $\tilde{\mathbf{Y}}^p, \tilde{\mathbf{N}} \in \mathbb{C}^{\tau_p N_r \times 1}$.

The MMSE equalizer aims at minimizing the MSE between the estimate $\hat{\mathbf{h}}_{MMSE} = \mathbf{H} \tilde{\mathbf{Y}}^p$ and the channel \mathbf{h} . More precisely,

$$\mathbf{H} = \arg \min_{\mathbf{H}} \mathcal{E}\{\|\mathbf{H} \tilde{\mathbf{Y}}^p - \mathbf{h}\|_F^2\}$$

$$= \alpha \sqrt{P_p} (\sigma^2 \mathbf{I}_{N_r} + \alpha^2 P_p \mathbf{C} \mathbf{S}^H \mathbf{S})^{-1} \mathbf{C} \mathbf{S}^H; \quad \mathbf{H} \in \mathbb{C}^{N_r \times \tau_p N_r}.$$

The MMSE estimate is then expressed as

$$\begin{aligned} \hat{\mathbf{h}}_{MMSE} &= \alpha \sqrt{P_p} (\sigma^2 \mathbf{I}_{N_r} + \alpha^2 P_p \tau_p \mathbf{C})^{-1} \mathbf{C} \mathbf{S}^H (\alpha \sqrt{P_p} \mathbf{S} \mathbf{h} + \tilde{\mathbf{N}}) \\ &= \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C} \right)^{-1} \mathbf{C} \left(\mathbf{h} + \frac{1}{\alpha \sqrt{P_p \tau_p}} \mathbf{S}^H \tilde{\mathbf{N}} \right). \end{aligned} \quad (7)$$

Notice that $\mathbf{S}^H \tilde{\mathbf{N}} \sim \mathcal{CN}(\mathbf{0}, \tau_p \sigma^2 \mathbf{I}_{N_r})$, therefore the estimated channel $\hat{\mathbf{h}}_{MMSE}$ is also a circular symmetric complex normal distributed vector $\hat{\mathbf{h}}_{MMSE} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{MMSE})$, that is

$$\hat{\mathbf{h}}_{MMSE} = \mathbf{h} + \tilde{\mathbf{h}}_{MMSE},$$

and

$$\mathbf{R}_{MMSE} = \mathbf{C}^2 \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C} \right)^{-1}, \quad (8)$$

where we considered $\mathbf{C} = \mathbf{C}^H$ and applied the commutativity of \mathbf{C} and \mathbf{I}_{N_r} to substitute

$$\left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C} \right)^{-1} \mathbf{C} = \mathbf{C} \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C} \right)^{-1}.$$

The channel estimation error is $\tilde{\mathbf{h}}_{MMSE} = \mathbf{h} - \hat{\mathbf{h}}_{MMSE}$ so that $\tilde{\mathbf{h}}_{MMSE} \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_{MMSE})$ with

$$\mathbf{W}_{MMSE} = \mathbf{C} \left(\mathbf{I}_{N_r} + \frac{\alpha^2 P_p \tau_p}{\sigma^2} \mathbf{C} \right)^{-1}$$

and the estimation MSE simply follows as

$$\varepsilon_{MMSE} = \text{tr} \left\{ \mathbf{C} \left(\mathbf{I}_{N_r} + \frac{\alpha^2 P_p \tau_p}{\sigma^2} \mathbf{C} \right)^{-1} \right\}. \quad (9)$$

Notice that for both LS and MMSE channel estimation, the estimation MSE is a monotonically decreasing function of the pilot energy per antenna $P_p \tau_p$.

In the next section, we characterize the receiver and the uplink signal MSE and spectral efficiency based on $\hat{\mathbf{h}}$ which is computed for LS ($\hat{\mathbf{h}}_{LS}$) and MMSE estimation ($\hat{\mathbf{h}}_{MMSE}$)

in this section.

IV. LINEAR MMSE RECEIVER

We consider a MMSE receiver at the BS. For each transmitted data symbol x , the data signal received by the BS can be written as

$$\mathbf{y} = \alpha \sqrt{P} \mathbf{h} x + \mathbf{n},$$

where $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_r})$.

Using a linear detection matrix \mathbf{G} of size $1 \times N_r$, the mean-square error between the estimate $\mathbf{G} \mathbf{y}$ and the transmitted symbol x is $(\mathbf{G} \mathbf{y} - x)^2$. The resulting $\text{MSE}(\mathbf{h}) = \mathcal{E}_{x, \mathbf{n}}\{(\mathbf{G} \mathbf{y} - x)^2\}$ is

$$\begin{aligned} \text{MSE}(\mathbf{h}) &= \mathbf{G} (\alpha^2 P \mathbf{h} \mathbf{h}^H + \sigma^2 \mathbf{I}_{N_r}) \mathbf{G}^H \\ &\quad - \alpha \sqrt{P} (\mathbf{G} \mathbf{h} + \mathbf{h}^H \mathbf{G}^H) + 1. \end{aligned} \quad (10)$$

According to [13], the instantaneous signal-to-noise-ratio (SNR) of the MMSE receiver is given by:

$$\gamma(\mathbf{h}) = \frac{1}{\text{MSE}(\mathbf{h})} - 1.$$

Using this relationship, the achievable spectral efficiency can be expressed as:

$$S(\mathbf{h}) = \frac{(\tau - \tau_p)}{\tau} \left[\log \left(\frac{1}{\text{MSE}(\mathbf{h})} \right) \right]. \quad (11)$$

We recall that the performance of the MMSE receiver depends on the availability of channel state information (CSI) at the receiver. In the considered scenario, the channel \mathbf{h} is not available at the BS. The BS implements the MMSE receiver by using the estimated channel $\hat{\mathbf{h}}$ as if it was the actual channel, [10], i.e., the detection matrix, which is a function of $\hat{\mathbf{h}}$ (more precisely, $\hat{\mathbf{h}}_{LS}$ or $\hat{\mathbf{h}}_{MMSE}$) is calculated as

$$\mathbf{G} = \mathbf{G}(\hat{\mathbf{h}}) = \frac{\alpha \sqrt{P} \hat{\mathbf{h}}^H}{\alpha^2 P \|\hat{\mathbf{h}}\|^2 + \sigma^2}. \quad (12)$$

In the next section, we derive a closed form analytical expression for the uplink MSE and the spectral efficiency for both LS and MMSE channel estimation.

V. ANALYTICAL DERIVATION OF THE SPECTRAL EFFICIENCY

In this section, we propose an analytical model to study the performance of the single-cell multiple antenna system illustrated in the previous section, in order to derive the optimal resource (slot/frequency/power) allocation for pilot and data. We first introduce a useful lemma and then we derive (10) and (11) analytically, by considering the MMSE receiver in (12).

A. Conditional Distribution of the Channel

When implementing the MMSE receiver in (12), the expression of the uplink MSE in (10) contains both the actual channel \mathbf{h} and the estimated channel $\hat{\mathbf{h}}$. However, since the square error is averaged over \mathbf{n} , the MSE depends on the conditional distribution ($\hat{\mathbf{h}}|\mathbf{h}$) or, equivalently, ($\mathbf{h}|\hat{\mathbf{h}}$). Therefore, in order

to compute the unconditional MSE analytically, we introduce the following lemma.

Lemma 1: The conditional distribution of the channel given its estimation $\hat{\mathbf{h}}$ is

$$(\mathbf{h}|\hat{\mathbf{h}}) \sim \mathcal{CN}(\mathbf{D}\hat{\mathbf{h}}, \mathbf{Q}), \quad (13)$$

where

$$\mathbf{D} = \begin{cases} \mathbf{C}\mathbf{R}_{LS}^{-1} & \text{for LS estimation} \\ \mathbf{I}_{N_r} & \text{for MMSE estimation} \end{cases}$$

$$\mathbf{Q} = \begin{cases} \mathbf{C} - \mathbf{C}\mathbf{R}_{LS}^{-1}\mathbf{C} & \text{for LS estimation} \\ \mathbf{C} - \mathbf{R}_{MMSE} & \text{for MMSE estimation} \end{cases}$$

with \mathbf{R}_{LS} and \mathbf{R}_{MMSE} given in (5) and (7), respectively.

Proof: The proof is reported in the appendix. ■

B. Calculating the Uplink MSE

By using Lemma 1, the MSE in (10) can be conveniently expressed as a function of only the estimated channel $\hat{\mathbf{h}}$ as

$$\begin{aligned} \text{MSE}(\hat{\mathbf{h}}) = & \mathbf{G}(\hat{\mathbf{h}}) \left(\alpha^2 P (\mathbf{D}\hat{\mathbf{h}}\hat{\mathbf{h}}^H \mathbf{D}^H + \mathbf{Q}) + \sigma^2 \mathbf{I}_{N_r} \right) \mathbf{G}(\hat{\mathbf{h}})^H \\ & - \alpha \sqrt{P} \left(\mathbf{G}(\hat{\mathbf{h}}) \mathbf{D}\hat{\mathbf{h}} + \hat{\mathbf{h}}^H \mathbf{D}^H \mathbf{G}(\hat{\mathbf{h}}) \right) + 1. \end{aligned} \quad (14)$$

In the special case of independent channel distributions with identical variances (i.e., $\mathbf{C} = c\mathbf{I}_{N_r}$), we exploit the following lemma.

Lemma 2: Assume $\mathbf{C} = c\mathbf{I}_{N_r}$, where $c \in \mathbb{R}^+$, then the matrices \mathbf{D} and \mathbf{Q} are diagonal with $\mathbf{D} = d\mathbf{I}_{N_r}$ and $\mathbf{Q} = q\mathbf{I}_{N_r}$ and the MSE is given by

$$\text{MSE}(\hat{\mathbf{h}}) = \frac{p^2 \|\hat{\mathbf{h}}\|^4 (d-1)^2 + p \|\hat{\mathbf{h}}\|^2 (2\sigma^2 - 2d\sigma^2 + b) + \sigma^4}{(p \|\hat{\mathbf{h}}\|^2 + \sigma^2)^2}, \quad (15)$$

where $p = \alpha^2 P$ and $b = qp + \sigma^2$.

Proof: The proof is reported in the appendix. ■

It is important to notice that (15) depends on the channel only through the norms $\|\hat{\mathbf{h}}\|^2$ and $\|\hat{\mathbf{h}}\|^4$.

C. Calculating the spectral efficiency

The spectral efficiency expression in (11) can be also conveniently rewritten as a function of $\hat{\mathbf{h}}$, by using the results from Lemma 1 and 2. The average spectral efficiency is then

$$\bar{S} = \frac{(\tau - \tau_p)}{\tau} \mathcal{E}_{\hat{\mathbf{h}}} \left\{ \left[\log \left(\frac{1}{\text{MSE}(\hat{\mathbf{h}})} \right) \right] \right\}. \quad (16)$$

which lead to the following results.

Result 1 (Spectral efficiency with LS estimation): Assume $\mathbf{C} = c\mathbf{I}_{N_r}$, where $c \in \mathbb{R}^+$, then the average spectral efficiency with LS channel estimation and MMSE receiver is

$$\bar{S}_{LS} = \frac{(\tau - \tau_p)}{\tau} \left(\frac{2\mathcal{G}(x_0) - \mathcal{G}(x_1) - \mathcal{G}(x_2)}{(N_r - 1)!} - \log(d-1)^2 \right) \quad (17)$$

with $x_{1,2} = \frac{1}{2} \left(-\frac{2\sigma^2 - 2d\sigma^2 + b}{p(d-1)^2} \pm \sqrt{\left(\frac{2\sigma^2 - 2d\sigma^2 + b}{p(d-1)^2} \right)^2 - \frac{4\sigma^4}{p^2(d-1)^2}} \right)$,
 $x_0 = \frac{\sigma^2}{p}$, $p = \alpha^2 P$, $b = qp + \sigma^2$, $q = c(1-c/r)$, $r = c + \frac{\sigma^2}{\alpha^2 P \tau_p}$,
and where

$$\mathcal{G}(x) = \text{MeijerG}_{2,3}^{1,3} \left(\begin{matrix} 0, 1 \\ 0, 0, N_r \end{matrix} \middle| \frac{x}{r} \right), \quad (18)$$

is the Meijer G-function.

Proof: The proof is reported in the appendix. ■

Result 2 (Spectral efficiency with MMSE estimation): Assume $\mathbf{C} = c\mathbf{I}_{N_r}$, where $c \in \mathbb{R}^+$, then the average spectral efficiency with MMSE channel estimation and MMSE receiver is

$$\bar{S}_{MMSE} = \frac{(\tau - \tau_p)}{\tau} \left(\log(pb) + \frac{2\mathcal{G}(x_3) - \mathcal{G}(x_4)}{(N_r - 1)!} \right) \quad (19)$$

with $x_3 = \frac{\sigma^2}{p}$, $x_4 = \frac{\sigma^2}{pb}$, $b = qp + \sigma^2$, $q = \frac{\sigma^2 c}{\sigma^2 + \alpha^2 c P \tau_p}$, and $\mathcal{G}(x)$ defined in (18).

Proof: The proof is reported in the appendix. ■

We notice that for a fixed pilot energy, the average spectral efficiency decreases with τ_p . Intuitively, this means that for a given pilot energy ($\tau_p P_p$), this energy should be concentrated to as few time slots as possible with the maximum allowed pilot power P_p . This is because the estimation MSE in Eqs. (6) and (9) only depend on the τ_p and P_p through the product (that is the pilot energy). The hypothesis of independent and identical channel distributions (i.e., $\mathbf{C} = c\mathbf{I}_{N_r}$) is necessary to obtain a tractable analytical expression and it is widely adopted in the related literature (i.e., [8]). Although correlation among antennas may have an effect on the value of achievable spectral efficiency, the optimal pilot resource allocation is not expected to change significantly with the antenna correlation.

D. Summary

In this section, we derived the spectral efficiency for a single cell multi-antenna system with pilot-based channel estimation that implements a linear MMSE receiver. Moreover, we provided a closed form analytical expression of the spectral efficiency that makes use of the Meijer G-function, for both LS and MMSE channel estimation methods.

We started by deriving the conditional distribution $(\mathbf{h}|\hat{\mathbf{h}})$, which allows for computing the uplink MSE as a function of the norm of the estimated channel $\|\hat{\mathbf{h}}\|^2$ through Lemma 2. By averaging the spectral efficiency over the distribution of $\|\hat{\mathbf{h}}\|^2$, we derived the closed form expression of the spectral efficiency for LS estimation in Result 1 and for MMSE estimation in Result 2.

Under the constraints (1) and (2) for symbols and power in the time-frequency domain, this model reproduces the behavior of the block type and comb type pilot arrangement, respectively, as we show in the following section.

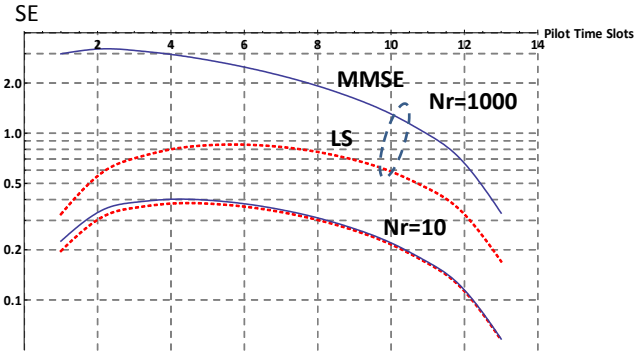


Fig. 1. Spectral efficiency (SE) in log scale of block type channel estimation as a function of the number of pilot time slots with $N_r = 10$ and $N_r = 1000$ antennas at the BS. With block type arrangement and EPA, all $F = 12$ subcarriers in each of the $T = 14$ time slots are dedicated to either pilot or data transmission with $P_{tot} = 250$ mW total transmit power shared equally in the frequency domain. The T_p that maximizes spectral efficiency is clearly different with LS and MMSE estimations.

VI. NUMERICAL RESULTS

A. Equal Power (Power Density for Each Symbol) Allocation

Recall that equal power allocation (EPA) implies that all symbols within a slot are transmitted with P_{tot}/F transmit power. In this case, the pilot-data resource allocation trade-off consists of the number of time slots (T_p out of T) in the case of block type arrangement or the number of frequency channels (F_p out of F) in the case of comb type arrangement used for pilot transmission.

Figure 1 shows the spectral efficiency in bps/Hz as the function of time slots T_p when using LS and MMSE channel estimation with $N_r = 10$ and $N_r = 1000$ BS antennas for a specific value of the large scale fading (50 dB). As expected, the maximum spectral efficiency that can be reached with MMSE estimation (3 bps/Hz) is much higher than that of the optimum spectral efficiency value with LS estimation (1 bps/Hz). Also, the optimum spectral efficiency is reached with different T_p settings: while with MMSE $T_p = 2$ maximizes the spectral efficiency, with LS we need to spend more slots on CSI acquisition ($T_p = 5$).

Figure 2 shows a similar tendency in terms of the necessary frequency channels (subcarriers) that optimize the spectral efficiency with comb type pilot arrangement. The similar (almost identical) behavior that can be seen in Figure 1 and 2 can be explained by noticing that under the assumption of channel coherence in the time and frequency domains of a resource block of $F = 12$ frequency channel and $T = 14$ time slots, block and comb type arrangements with equal power allocation among the $F \times T = 168$ resource elements result in the same total pilot energy. That is under the coherence assumption, the roles of the time and frequency domains from the perspective of pilot energy are the same.

B. Optimum Power Allocation

When using the comb pilot symbol pattern, it is possible to use F_p subcarriers for transmitting pilot symbols and the

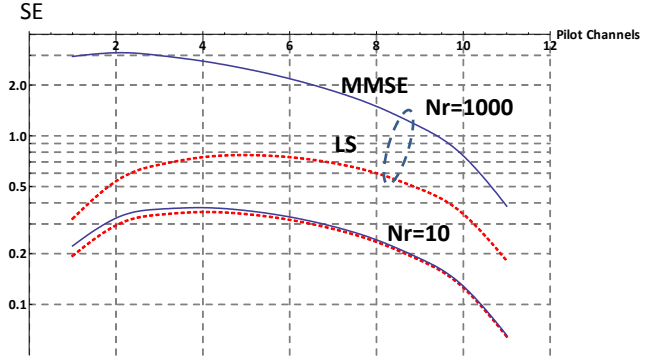


Fig. 2. Spectral efficiency (SE) in log scale of comb type channel estimation as a function of the number of pilot channels in the frequency domain with $N_r = 10$ and $N_r = 1000$ antennas at the BS. With comb type arrangement and EPA, F_p subcarriers (each with P_{tot}/F transmit power) are used to transmit pilot symbols in each of the $T = 14$ time slots. The F_p that maximizes spectral efficiency is clearly different with LS and MMSE estimations.

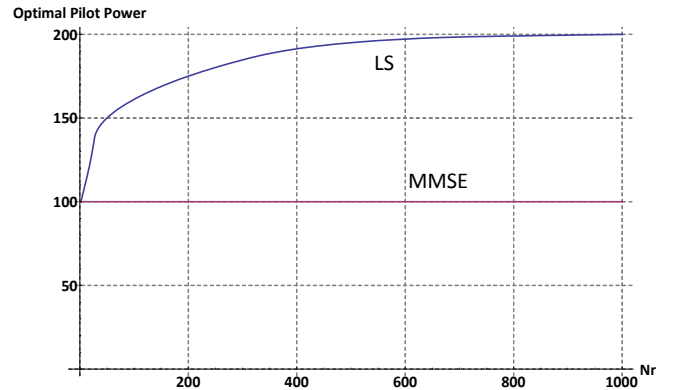


Fig. 3. Optimum pilot power in mW as the function of the number of receive antennas at the BS when using LS (upper curve) or MMSE (lower curve) channel estimation. With LS estimation, the optimum pilot power increases with the number of antennas, whereas with MMSE estimation, the optimum pilot power is constant (staying at 40% of the total power budget P_{tot} in each time slot).

remaining $F - F_p$ symbols for data transmission in each time slot. In this case, it is also possible to use unequal total power for pilot (P_p) and data symbol transmission as long as the sum over the F symbols in each time slot does not exceed the P_{tot} power budget. In this case, each of the F_p pilot subcarriers are transmitted with F_p/P_p transmit power and both F_p and P_p are design parameters.

Figure 3 shows the value of the pilot power P_p in mW that maximizes the spectral efficiency when using the comb arrangement and employing LS (upper) and MMSE (lower) channel estimation as the number of antennas grows from $N_r = 2$ to $N_r = 1000$. With LS, the optimal pilot power grows from about 40% to around 80% of the P_{tot} total power budget, whereas with MMSE estimation, the optimal pilot power remains the same.

Figure 4 shows the achieved spectral efficiency as the

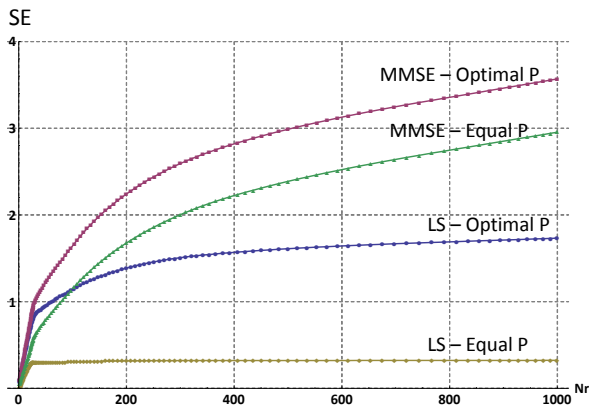


Fig. 4. The spectral efficiency (SE) as the function of the number of receive antennas at the BS, when employing LS (lower 2 curves) and MMSE (upper 2 curves) channel estimation. In both cases, optimum pilot power allocation is compared with equal power allocation between pilot and data transmission. With LS estimation, optimum pilot power allocation gives large gains, whereas with MMSE estimation, this spectral efficiency gain obtained by optimum pilot power allocation is less, although still significant.

function of the antennas with LS and MMSE estimation with equal (EPA) and optimal pilot power allocation. Optimizing the pilot power is clearly beneficial with both LS and MMSE estimations. With LS estimation, optimizing the pilot power is particularly important as the number of antennas grows large, but even with MMSE estimation, the spectral efficiency increases by 20%. Recall that unequal power allocation over the subcarriers in each time slot requires comb type arrangement, implying that with block type pilot pattern, MMSE estimation gives large gains over LS estimation when the number of antennas is large.

Figures 5 and 6 compare the achievable spectral efficiency as the function of the pilot power and the number of pilot frequency channels with $N_r = 10$ and $N_r = 1000$ receive antennas. With $N_r = 10$, the achievable spectral efficiency is similar with LS and MMSE channel estimation, whereas with $N_r = 1000$ antennas, the optimal spectral efficiency is roughly twice as high with MMSE as with LS.

VII. CONCLUSIONS

In this paper, we considered the trade-off between the time, frequency and power resources allocated to the transmission of pilot and data symbols and its impact on the MSE and spectral efficiency of the uplink of a single cell system, in which the number of receive antennas grows large. We made the point that the joint allocation of frequency, time and power resources is subject to constraints that depend on the specific pilot pattern, such as the pattern used by the block and comb type pilot arrangements. In this rather general setting, we provided an analytical method to calculate the MSE and the uplink spectral efficiency that enabled us to derive exact numerical results when the receiver at the base station employs LS or MMSE channel estimation and MMSE equalizer for uplink data reception. We found that with a large number of antennas, exploiting the engineering freedom of tuning *both* the

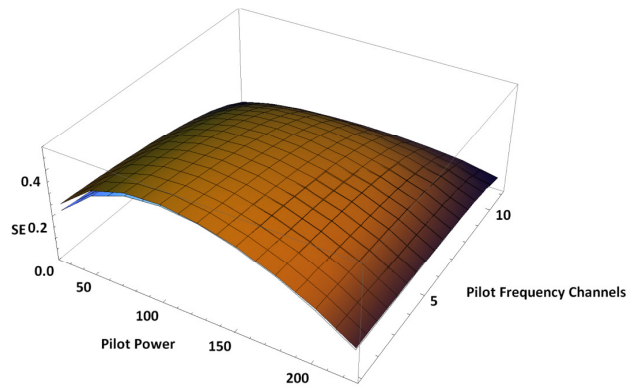


Fig. 5. Spectral efficiency with comb pilot arrangement and LS (lower) and MMSE (upper) channel estimation as a function of the number of frequency channels and the total pilot power (out of the P_{tot}) with $N_r = 10$ receive antennas. The pilot power that maximizes spectral efficiency is around $P_p^{opt} = 100$ mW with both LS and MMSE.

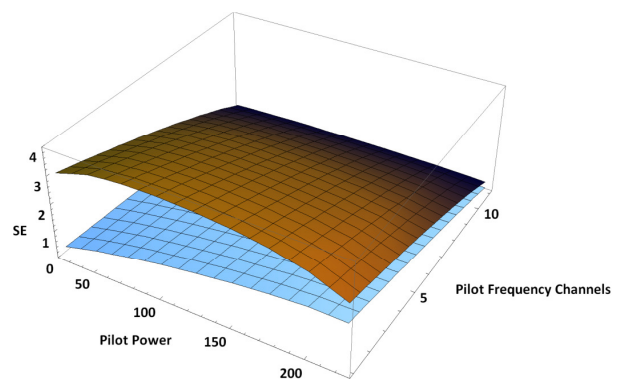


Fig. 6. Spectral efficiency with comb pilot arrangement and LS (lower) and MMSE (upper) channel estimation as a function of the number of frequency channels and the total pilot power with $N_r = 1000$ receive antennas. The pilot power that maximizes spectral efficiency is around $P_p^{opt} = 200$ mW with LS and 100 mW with MMSE estimation.

number of pilot symbols and the pilot transmit power levels become increasingly important, especially if the relatively simple LS estimator is used at the base station. Also, the gain of using MMSE estimation (preferably with optimized pilot power allocation) increases over LS estimation. Interestingly, the optimal PPDR is different when using MMSE and LS estimators and the gain in terms of spectral efficiency when optimizing *both* the number of pilot symbols and the transmit power levels increases as the number of antennas increases. We believe that our methodology as well as the obtained insights are new and provide useful guidelines for designing practical large antenna systems.

ACKNOWLEDGMENTS

M. Telek was partially supported by the OTKA K101150 project.

PROOF OF LEMMA 1

Given $\mathbf{h} \sim \mathcal{CN}(0, \mathbf{C})$, and $\hat{\mathbf{h}} \sim \mathcal{CN}(0, \mathbf{R})$, from (10.24)–(10.27) of [14], it follows that $\mathbf{h}|\hat{\mathbf{h}}$ is a complex normal distributed random vector with the following mean and covariance

$$\begin{aligned}\mathcal{E}(\mathbf{h}|\hat{\mathbf{h}}) &= \mathcal{E}(\mathbf{h}\hat{\mathbf{h}}^H)\mathbf{R}^{-1}\hat{\mathbf{h}}, \\ \mathbf{C}_{\mathbf{h}|\hat{\mathbf{h}}} &= \mathbf{C} - \mathcal{E}(\mathbf{h}\hat{\mathbf{h}}^H)\mathbf{R}^{-1}\mathcal{E}(\hat{\mathbf{h}}\mathbf{h}^H).\end{aligned}$$

For the LS channel estimation model in (4), we derive

$$\mathcal{E}(\mathbf{h}\hat{\mathbf{h}}^H) = \mathbf{C} + \frac{1}{\alpha\sqrt{P_p\tau_p}}\mathcal{E}(\mathbf{h}\mathbf{s}^T\mathbf{N}^H) = \mathbf{C}.$$

For the MMSE channel estimation model in (7), we derive

$$\begin{aligned}\mathcal{E}(\mathbf{h}\hat{\mathbf{h}}^H) &= \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C}\right)^{-1} \mathbf{C} \left(\mathbf{C} + \frac{1}{\alpha\sqrt{P_p\tau_p}}\mathcal{E}(\mathbf{h}\tilde{\mathbf{N}}^H\mathbf{S})\right) \\ &= \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C}\right)^{-1} \mathbf{C}^2 = \mathbf{C}^2 \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p} \mathbf{I}_{N_r} + \mathbf{C}\right)^{-1}.\end{aligned}$$

The expressions of \mathbf{D} and \mathbf{Q} in (13) are obtained by substitution.

PROOF OF LEMMA 2

Let us assume $\mathbf{C} = c\mathbf{I}_{N_r}$. In the LS estimation case, from (5) we have

$$\mathbf{R}_{LS} = \left(c + \frac{\sigma^2}{\alpha^2 P_p \tau_p}\right) \mathbf{I}_{N_r}$$

and therefore by using Lemma 1, $\mathbf{D} = d\mathbf{I}_{N_r}$ and $\mathbf{Q} = q\mathbf{I}_{N_r}$, where

$$d = c \left(c + \frac{\sigma^2}{\alpha^2 P_p \tau_p}\right)^{-1} \quad \text{and} \quad q = c - c^2 \left(c + \frac{\sigma^2}{\alpha^2 P_p \tau_p}\right)^{-1}.$$

In the MMSE estimation case, from (8) we have

$$\mathbf{R}_{MMSE} = c^2 \left(c + \frac{\sigma^2}{\alpha^2 P_p \tau_p}\right)^{-1} \mathbf{I}_{N_r},$$

and therefore by using Lemma 1, $\mathbf{D} = d\mathbf{I}_{N_r}$ and $\mathbf{Q} = q\mathbf{I}_{N_r}$, where

$$d = 1 \quad \text{and} \quad q = c - c^2 \left(c + \frac{\sigma^2}{\alpha^2 P_p \tau_p}\right)^{-1}.$$

By replacing $\mathbf{D} = d\mathbf{I}_{N_r}$, $\mathbf{Q} = q\mathbf{I}_{N_r}$ and (12) in (14), we obtain

$$\begin{aligned}\text{MSE}(\hat{\mathbf{h}}) &= 1 - \frac{2\|\hat{\mathbf{h}}\|^2 d \alpha^2 P}{\|\hat{\mathbf{h}}\|^2 \alpha^2 P + \sigma^2} + \frac{\|\hat{\mathbf{h}}\|^2 \alpha^2 P}{\left(\|\hat{\mathbf{h}}\|^2 \alpha^2 P + \sigma^2\right)^2} \\ &\quad \times \left[\|\hat{\mathbf{h}}\|^2 d^2 \alpha^2 P + q \alpha^2 P + \sigma^2\right] = \\ &= \frac{\left(\|\hat{\mathbf{h}}\|^2 p + \sigma^2\right)^2}{\left(\|\hat{\mathbf{h}}\|^2 p + \sigma^2\right)^2} - \frac{2\|\hat{\mathbf{h}}\|^2 dp \left(\|\hat{\mathbf{h}}\|^2 p + \sigma^2\right)}{\left(\|\hat{\mathbf{h}}\|^2 p + \sigma^2\right)^2} +\end{aligned}$$

$$+ \frac{\|\hat{\mathbf{h}}\|^2 p \left(\|\hat{\mathbf{h}}\|^2 d^2 p + \overbrace{qp + \sigma^2}^b\right)}{\left(\|\hat{\mathbf{h}}\|^2 p + \sigma^2\right)^2}.$$

$$\text{MSE}(\hat{\mathbf{h}}) = \frac{p^2 \|\hat{\mathbf{h}}\|^4 (d-1)^2 + p \|\hat{\mathbf{h}}\|^2 (2\sigma^2 - 2d\sigma^2 + b) + \sigma^4}{\left(\|\hat{\mathbf{h}}\|^2 p + \sigma^2\right)^2},$$

where $p = \alpha^2 P$, $b = qp + \sigma^2$.

PROOF OF RESULT 1

The key step to prove Result 1 is to derive the expectation of the log term in Eq. (16).

By considering the MSE in (15), the log term in Eq. (16) can be written as:

$$\begin{aligned}\log\left(\frac{1}{\text{MSE}(\hat{\mathbf{h}})}\right) &= -\log(\text{MSE}(\hat{\mathbf{h}})) \\ &= -\log(p^2 Y^2 (d-1)^2 + pY(2\sigma^2 - 2d\sigma^2 + b) + \sigma^4) \\ &\quad + 2\log(Yp + \sigma^2)\end{aligned}$$

where $Y = \|\hat{\mathbf{h}}\|^2$.

We therefore need to calculate the following expectation:

$$\begin{aligned}\mathcal{E}_{\hat{\mathbf{h}}}\left[-\log(\text{MSE}(\hat{\mathbf{h}}))\right] &= 2 \int_{x=0}^{\infty} \log(xp + \sigma^2) f_Y(x) dx \quad (20) \\ &\quad - \int_{x=0}^{\infty} \log\left(p^2 x^2 (d-1)^2 + px(2\sigma^2 - 2d\sigma^2 + b) + \sigma^4\right) f_Y(x) dx,\end{aligned}$$

where the density function of Y – being the sum of exponentially distributed random variables of parameter r , where r is the diagonal element of \mathbf{R} –, is given by [10]:

$$f_Y(x) = \frac{r^{-N_r} x^{N_r-1} e^{-x/r}}{(N_r - 1)!} \quad x > 0.$$

Notice that for LS estimation, $d = c \left(c + \frac{\sigma^2}{\alpha^2 P_p \tau_p}\right)^{-1} \neq 1$, therefore the terms of the integral can be rearranged as follows

$$\begin{aligned}\mathcal{E}_{\hat{\mathbf{h}}}\left[-\log(\text{MSE}(\hat{\mathbf{h}}))\right] &= 2 \int_{x=0}^{\infty} \log(p(x + x_0)) f_Y(x) dx \\ &\quad - \int_{x=0}^{\infty} \log\left(p^2 (d-1)^2 (x^2 + a_1 x + a_0)\right) f_Y(x) dx \\ &= 2 \int_{x=0}^{\infty} \log(x + x_0) f_Y(x) dx - \log(d-1)^2 \\ &\quad - \int_{x=0}^{\infty} \log(x^2 + a_1 x + a_0) f_Y(x) dx \quad (21)\end{aligned}$$

where $x_0 = \frac{\sigma^2}{p}$, $a_0 = \frac{\sigma^4}{p^2(d-1)^2}$, and $a_1 = \frac{2\sigma^2 - 2d\sigma^2 + b}{p(d-1)^2}$.

The last integral can be further simplified by considering

$$\log(x^2 + a_1 x + a_0) = \log(x - x_1) - \log(x - x_2),$$

where $x_{1,2} = 0.5 \left(-a_1 \pm \sqrt{a_1^2 - 4a_0}\right)$.

In conclusion, we have to compute integrals of the form $\int_{x=0}^{\infty} \log(x + A) f_Y(x) dx$, which can be solved in Mathematica [15] via the Meijer G-function.

Specifically:

$$\int_{x=0}^{\infty} \log(x+A) f_Y(x) dx = \frac{A^{N_r}}{(N_r-1)! r^{N_r}} \text{MeijerG}_{2,3}^{1,3} \left(\begin{matrix} -N_r, -N_r+1 \\ -N_r, -N_r, 0 \end{matrix} \middle| \frac{A}{r} \right) \quad (22)$$

where $A, N_r > 0$, $r > 1$ and $\text{MeijerG}_{p,q}^{m,n} \left(\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| z \right)$ is the Meijer G-function with parameters p, q, m, n .

Recognizing that

$$\begin{aligned} z^{N_r} \text{MeijerG}_{2,3}^{1,3} \left(\begin{matrix} -N_r, -N_r+1 \\ -N_r, -N_r, 0 \end{matrix} \middle| z \right) \\ = \text{MeijerG}_{2,3}^{1,3} \left(\begin{matrix} 0, 1 \\ 0, 0, N_r \end{matrix} \middle| z \right), \end{aligned}$$

(22) is equivalent with:

$$\int_{x=0}^{\infty} \log(x+A) f_Y(x) dx = \frac{\text{MeijerG}_{2,3}^{1,3} \left(\begin{matrix} 0, 1 \\ 0, 0, N_r \end{matrix} \middle| z \right)}{(N_r-1)!} \quad (23)$$

where $z = \frac{A}{r}$.

By substituting the result of (23) in (21), (17) follows.

PROOF OF RESULT 2

The proof follows with similar steps as for Result 1. However, in the MMSE case, we have $d = 1$, and the expectation in (20) can be conveniently rewritten as

$$\begin{aligned} \mathcal{E}_{\hat{\mathbf{h}}} \left[-\log(\text{MSE}(\hat{\mathbf{h}})) \right] &= \\ &= 2 \int_{x=0}^{\infty} \log(xp + \sigma^2) f_Y(x) dx - \int_{x=0}^{\infty} \log(xb + \sigma^4) f_Y(x) dx \\ &= \log(pb) + 2 \int_{x=0}^{\infty} \log(x+x_3) f_Y(x) dx - \int_{x=0}^{\infty} \log(x+x_4) f_Y(x) dx, \end{aligned}$$

with $x_3 = \frac{\sigma^2}{p}$, $x_4 = \frac{\sigma^2}{pb}$, $b = qp + \sigma^2$, $q = \frac{\sigma^2 c}{\sigma^2 + \alpha^2 c P_p \tau_p}$,

which is solved in Mathematica by using the Meijer G-function defined in (18).

REFERENCES

- [1] B. Hassibi and B. M. Hochwald, "How Much Training is Needed in Multiple Antenna Wireless Links," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.
- [2] T. Marzetta, "How Much Training is Needed for Multiuser MIMO ?" *40th Asilomar Conference*, October 2006.
- [3] L. Li, A. Ashikhmin, and T. Marzetta, "Pilot contamination precoding for interference reduction in large scale antenna systems," *51th Annual Allerton Conference*, vol. Allerton House, Illinois, USA, pp. 226–232, October 2013.
- [4] T. Kim and J. G. Andrews, "Optimal Pilot-to-Data Power Ratio for MIMO-OFDM," *IEEE Globecom*, 2005.
- [5] —, "Balancing Pilot and Data Power for Adaptive MIMO-OFDM Systems," *IEEE Globecom*, 2005.
- [6] E. Golovins and N. Ventura, "Optimal Training for the SM-MIMO-OFDM Systems with MMSE Channel Estimation," *6th Annual Communication Networks and Services Research Conference*, pp. 470–477, 2008.
- [7] V. K. V. Gottomukkala and H. Minn, "Capacity Analysis and Pilot-Data Power Allocation for MIMO-OFDM With Transmitter and Receiver IQ Imbalances and Residual Carrier Frequency Offset," *IEEE Trans. Veh. Techn.*, pp. 553–565, 2012.

- [8] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels," *IEEE Transactions on Communications*, vol. 58, no. 10, pp. 2939–2948, October 2010.
- [9] K. T. Truong, A. Lozano, and R. W. H. Jr., "Optimal training in continuous block-fading massive mimo systems," *20th European Wireless, Barcelona, Spain*, May 2014.
- [10] G. Fodor and M. Telek, "On the Pilot-Data Power Trade Off in Single Input Multiple Output Systems," *European Wireless '14, Barcelona, Spain*, May 2014.
- [11] B. Furht and S. A. Ahson, *Long Term Evolution: 3GPP LTE Radio and Cellular Technology*. Auerbach Publications, ISBN-10: 1420072102, April 2009.
- [12] X. Hou, Z. Zhang, and H. Kayama, "DMRS Design and Channel Estimation for LTE-Advanced MIMO Uplink," *70th IEEE Vehicular Technology Conference (Fall)*, 2009.
- [13] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*. Prentice Hall, 1993, no. ISBN: 0133457117.
- [15] S. Wolfram, *The Mathematica Book*. Wolfram Media, ISBN 1-57955-022-3, 2003.