On the Impact of Antenna Correlation on the Pilot-Data Balance in Multiple Antenna 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*—We consider the uplink of a single cell single input multiple output (SIMO) system, in which the mobile stations use intra-cell orthogonal pilots to facilitate uplink channel estimation. In such systems, the problem of transmission power balancing between pilot and data is known to have a large impact on the mean square error (MSE) for the received signal and, consequently, on the achievable uplink data rate. In this paper, we derive a closed form expression of the MSE for the received signal as a function of the pilot and data power levels under a per-user sum pilot-data power constraint. As a major contribution, our model is developed for arbitrary channel covariance matrices and it enables us to study the impact of the number of antennas and antenna correlation structures, including the popular 3GPP spatial channel model. Numerical results suggest that the effect of the antenna spacing is limited, but the angle of arrival and angular spread have a strong and articulated impact on the MSE performance. Moreover, as the number of antennas at the base station grows large, we show that a higher percentage of the power budget should be allocated to pilot signals than with a lower number of antennas.** [1]

I. Introduction

In multiple input multiple output (MIMO) orthogonal frequency division multiplexing (OFDM) systems pilot symbols facilitate channel estimation, but they reduce the transmitted energy for data symbols under a fixed power budget. This fundamental tradeoff has been studied in [1] and [2] where the optimal pilot-to-data power ratio (PDPR) is investigated for various pilot patterns and receiver structures. The results of [2], for example, indicate that the optimal PDPR provides about 2-3 dB gain compared with the equal power for pilot and data symbols. Subsequently, [3] derived a closed form of the optimal PDPR in MIMO-OFDM spatial multiplexing systems with 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, [4] derived a closed form PDPR that maximizes the capacity bound of MIMO-OFDM systems and studied the impact of carrier frequency offset (CFO) on the maximizing power allocation. In [5] and [6], an approximate closed-form analytical expression of the spectral efficiency is derived assuming MMSE estimation.

[1]The 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.

In our previous works [7], [8], [9], we investigated the optimal PDPR in the case of uncorrelated antennas giving rise to a diagonal covariance matrix. An isolated cell was considered in [7], [8]. The multi-cell case was studied in [9], where it was found that the so called distributed iterative channel inversion (DICI) algorithm originally proposed by [10] can be advantageously extended taking into account the pilot-data power trade off. However, none of aforementioned works capture the impact of antenna correlation on the performance of single input multiple output (SIMO) systems.

In this paper, we consider a SIMO system in which the mobile station balances its PDPR, while the base station uses maximum likelihood (ML) channel estimation to initialize a linear MMSE equalizer. Our contribution is to derive a closed form for the mean square error (MSE) of the equalized data symbols for arbitrary correlation structure between the antennas by allowing any covariance matrix of the uplink channel. This form is powerful, because it considers not only the pilot and data transmit power levels and the number of receive antennas at the base station (N_r) as independent variables, but it also explicitly takes into account antenna spacing and the statistics of the angle of arrivals (AoAs), including the angular spread as a parameter. For example, the methodology enables us to study the impact of the PDPR on the uplink performance for the popular 3GPP spatial channel model (SCM) often used to model the wireless channels in cellular systems [11]. The closed form formula takes into account the impact of N_r, AoA and angular spread on the MSE and thereby on the PDPR that minimizes the MSE. To the best of our knowledge the analytical result as well as the insights obtained in the numerical section are novel.

The system model is defined in Section II. In this section, for the sake of completeness and readability, we restate and reuse some results of [8]. Next, Section III describes the channel estimation models for least square and minimum mean square error channel estimators. Section IV is concerned with deriving the conditional mean square error of the uplink equalized data symbols using either of the channel estimation techniques and assuming MMSE equalization. Based on the results of this section, the unconditional MSE with arbitrary channel covariance matrix is determined in Section V. Numerical results are studied in Section VI. Section 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 estimators 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 $\tau_p + \tau_d = F \cdot T,$ where we denote by τ_p the number of symbols allocated to pilot, and by τ_d the number of data 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.,

$$FP_p \leq P_{tot} \qquad FP \leq P_{tot}. \tag{1}$$

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:

$$\frac{\tau_p}{T} P_p + \frac{\tau_d}{T} P = P_{tot} \tag{2}$$

The total number of symbols for pilot 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. With fixed (given or standardized) τ_p and τ_d the engineering freedom includes the tuning of the P_p and P power levels, which is the topic of the present paper.

III. Channel Estimation

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

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

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

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

Note that $\mathbf{N}\,\mathbf{s}^* = \left[\sum_{i=1}^{\tau_p} s_i^* n_{i,1}, ..., \sum_{i=1}^{\tau_p} s_i^* n_{i,N_r}\right]^T$, 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

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

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

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

where $|| \cdot ||_F^2$ is the Frobenius norm.

B. MMSE Estimation

In this case we find it convenient to 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

$$\tilde{\mathbf{Y}}^p = \alpha\sqrt{P_p}\mathbf{S}\mathbf{h} + \tilde{\mathbf{N}}. \tag{7}$$

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,

$$\begin{aligned}\mathbf{H} =& \arg\min_{\mathbf{H}} \mathcal{E}\{||\mathbf{H}\tilde{\mathbf{Y}}^p - \mathbf{h}||_F^2\}\\ =& \alpha\sqrt{P_p}(\sigma^2\mathbf{I}_{N_r} + \alpha^2 P_p \mathbf{C}\mathbf{S}^H\mathbf{S})^{-1}\mathbf{C}\mathbf{S}^H; \quad \mathbf{H} \in \mathbb{C}^{N_r \times \tau_p N_r}.\end{aligned}$$

The MMSE estimate is then expressed as

$$\begin{aligned}\hat{\mathbf{h}}_{MMSE} &= \alpha\sqrt{P_p}(\sigma^2\mathbf{I}_{N_r} + \alpha^2 P_p \tau_p \mathbf{C})^{-1}\mathbf{C}\mathbf{S}^H(\alpha\sqrt{P_p}\mathbf{S}\mathbf{h} + \tilde{\mathbf{N}})\\ &= \left(\frac{\sigma^2}{\alpha^2 P_p \tau_p}\mathbf{I}_{N_r} + \mathbf{C}\right)^{-1}\mathbf{C}\left(\mathbf{h} + \frac{1}{\alpha\sqrt{P_p}\tau_p}\mathbf{S}^H\tilde{\mathbf{N}}\right). \end{aligned}\tag{8}$$

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

$$\hat{\mathbf{h}}_{MMSE} = \mathbf{h} + \tilde{\mathbf{h}}_{MMSE}, \tag{9}$$

and

$$\mathbf{R}_{MMSE} = \mathbf{C}^2\left(\frac{\sigma^2}{\alpha^2 P_p \tau_p}\mathbf{I}_{N_r} + \mathbf{C}\right)^{-1}, \tag{10}$$

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

$$\varepsilon_{MMSE} = \mathrm{tr}\left\{\mathbf{C}\left(\mathbf{I}_{N_r} + \frac{\alpha^2 P_p \tau_p}{\sigma^2}\mathbf{C}\right)^{-1}\right\}. \tag{11}$$

Notice that for both LS and MMSE channel estimation, the estimation MSE is a monotonically decreasing function of the pilot energy per antenna $P_p\tau_p$. The quantity $P_p\tau_p$ can be regarded as the total pilot power (energy) budget, that - assuming a fixed τ_p - can be tuned by tuning P_p.

IV. Determining the Conditional Mean Square Error

A. A Key Observation

Equations (4) and (9) imply that $\mathbf{h}$ and $\hat{\mathbf{h}}$ are jointly circular symmetric complex Gaussian (multivariate normal) distributed random variables [12], [13] . Specifically, we recall from [12] that the covariance matrix of the joint PDF is composed by autocovariance matrices $\mathbf{C}_{\mathbf{h},\mathbf{h}}$, $\mathbf{C}_{\hat{\mathbf{h}},\hat{\mathbf{h}}}$ and cross covariance matrices $\mathbf{C}_{\mathbf{h},\hat{\mathbf{h}}}$, $\mathbf{C}_{\hat{\mathbf{h}},\mathbf{h}}$ as

$$\begin{bmatrix}\mathbf{C}_{\mathbf{h},\mathbf{h}} & \mathbf{C}_{\mathbf{h},\hat{\mathbf{h}}}\\ \mathbf{C}_{\hat{\mathbf{h}},\mathbf{h}} & \mathbf{C}_{\hat{\mathbf{h}},\hat{\mathbf{h}}}\end{bmatrix} = \begin{bmatrix}\mathbf{C} & \mathbf{C}\\ \mathbf{C} & \mathbf{R}\end{bmatrix},$$

and

$$\mathbf{R} = \begin{cases}\mathbf{R}_{LS} & \text{for LS estimation,}\\ \mathbf{R}_{MMSE} & \text{for MMSE estimation.}\end{cases}$$

B. Determining the Conditional Channel Distribution

From the joint PDF of $\mathbf{h}$ and $\hat{\mathbf{h}}$ we can compute the following conditional distributions.

Result 4.1: Given a random channel realization $\mathbf{h}$, the estimated channel $\hat{\mathbf{h}}$ conditioned to $\mathbf{h}$ can be shown to be distributed as

$$(\hat{\mathbf{h}} \mid \mathbf{h}) \sim \mathbf{h} + \mathcal{CN}(\mathbf{0}, \mathbf{R} - \mathbf{C}). \tag{12}$$

Result 4.2: The distribution of the channel realization $\mathbf{h}$ conditioned to the estimate $\hat{\mathbf{h}}$ is normally distributed as follows:

$$(\mathbf{h} \mid \hat{\mathbf{h}}) \sim \mathbf{D}\hat{\mathbf{h}} + \mathcal{CN}\Big(\mathbf{0}, \mathbf{Q}\Big), \tag{13}$$

where $\mathbf{D} = \mathbf{C}\mathbf{R}^{-1}$ and $\mathbf{Q} = \mathbf{C} - \mathbf{C}\mathbf{R}^{-1}\mathbf{C}$.
Both of these results can be easily verified by exploiting the basic characteristics of multivariate Gaussian random variables [12]. To capture the tradeoff between the pilot and data power, we need to calculate the mean square error of the equalized data symbols. To this end, we consider an equalization model in the next subsection.

C. Equalizer Model Based on LS or MMSE Channel Estimation

The data signal received by the BS is

$$\mathbf{y} = \alpha\sqrt{P}\mathbf{h}x + \mathbf{n}, \tag{14}$$

where $|x|^2 = 1$. We assume that the BS employs a naive MMSE equalizer, where the estimated channel (either $\hat{\mathbf{h}}_{LS}$ or $\hat{\mathbf{h}}_{MMSE}$) is taken as if it was the actual channel:

$$\mathbf{G} = \alpha\sqrt{P}\hat{\mathbf{h}}^H(\alpha^2 P\hat{\mathbf{h}}\hat{\mathbf{h}}^H + \sigma^2\mathbf{I})^{-1}. \tag{15}$$

Under this assumption, we recall the following result from [8] as a first step towards determining unconditional MSE.

Result 4.3: Let $\mathrm{MSE}(\hat{\mathbf{h}}) = \mathcal{E}_{\mathbf{h}|\hat{\mathbf{h}}}\Big\{\mathrm{MSE}(\mathbf{h}, \hat{\mathbf{h}})\Big\}$ be the MSE for each equalized data symbol, given the estimated channel realization $\hat{\mathbf{h}}$. It satisfies [14]:

$$\begin{aligned}\mathrm{MSE}(\hat{\mathbf{h}}) = \mathbf{G}\Big(\alpha^2 P(\mathbf{D}\hat{\mathbf{h}}\hat{\mathbf{h}}^H\mathbf{D}^H + \mathbf{Q}) + \sigma^2\mathbf{I}\Big)\mathbf{G}^H -\\ - 2\alpha\sqrt{P}Re\{\mathbf{G}\mathbf{D}\hat{\mathbf{h}}\} + 1.\end{aligned}\tag{16}$$

V. Derivation of the Unconditional Mean Square Error

Based on the conditional MSE expression of the preceding section, we are now interested in deriving the unconditional expectation of the MSE. To this end, the following two lemmas turn out to be useful.

Lemma *1:* Given a channel estimate instance $\hat{\mathbf{h}}$, the MMSE weighting matrix $\mathbf{G}$, as a function of the number of receive antennas at the base station (N_r) can be expressed as follows

$$\mathbf{G} = \frac{\alpha\sqrt{P}}{\|\hat{\mathbf{h}}\|^2\alpha^2 P + \sigma^2}\ \hat{\mathbf{h}}^H, \tag{17}$$

where $\|\hat{\mathbf{h}}\|^2 = \hat{\mathbf{h}}^H\hat{\mathbf{h}} = \sum_{i=1}^{N_r}|\hat{h}_i|^2$.
The proof of the lemma is straightforward [7] and omitted due to space limitations.

Using $z = \frac{\alpha\sqrt{P}}{\|\hat{\mathbf{h}}\|^2\alpha^2 P+\sigma^2}$ and this simple expression of $\mathbf{G}$ we can express the conditional expectation of the MSE of the MMSE equalized data symbols as a function of the channel covariance, $\mathbf{C}$.

Lemma *2:*

$$\begin{aligned}\text{MSE}(\hat{\mathbf{h}}) &= z^2\alpha^2 P\ \hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\hat{\mathbf{h}}^H\mathbf{D}^H\hat{\mathbf{h}} + z^2\alpha^2 P\ \hat{\mathbf{h}}^H\mathbf{Q}\hat{\mathbf{h}}+\\ &+ z^2\sigma^2\hat{\mathbf{h}}^H\hat{\mathbf{h}} - 2z\alpha\sqrt{P}Re\{\hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\} + 1.\end{aligned}$$

where $z = \frac{\alpha\sqrt{P}}{\|\hat{\mathbf{h}}\|^2\alpha^2 P+\sigma^2}$ is a function of $\|\hat{\mathbf{h}}\|^2$.
The proof is in the Appendix.

A. Computing z

Computing z in Lemma 2 is essential in calculating the unconditional MSE. This can be done using the following lemma.

Lemma *3:*

$$z = \frac{1}{\alpha\sqrt{P}}\cdot\int_{x=0}^{\infty} e^{-x(\hat{\mathbf{h}}^H\hat{\mathbf{h}}+\sigma^2/(\alpha^2P))}dx$$

and

$$z^2 = \frac{1}{\alpha^2 P}\cdot\int_{x=0}^{\infty} xe^{-x(\hat{\mathbf{h}}^H\hat{\mathbf{h}}+\sigma^2/(\alpha^2P))}dx.$$

The proof is in the Appendix.

B. Singular Value Decomposition

To take the expectation with respect to $\hat{\mathbf{h}}$ of the terms of MSE($\hat{\mathbf{h}}$) as in Lemma 2, we need the following decomposition. $\hat{\mathbf{h}}$ is $\mathcal{CN}(\mathbf{0},\mathbf{R})$ distributed with $\mathbf{R} = \mathbf{C} + \frac{\sigma^2}{P_p\alpha^2}\ \mathbf{I}$ and we recall from Result 4.2 that $\mathbf{D} = \mathbf{C}\mathbf{R}^{-1}$. Let $\mathbf{C} = \mathbf{\Theta}^H\mathbf{S}_{\mathbf{C}}\mathbf{\Theta}$ be the singular value decomposition of $\mathbf{C}$. Then $\mathbf{R} = \mathbf{\Theta}^H\mathbf{S}_{\mathbf{R}}\mathbf{\Theta}$, $\mathbf{D} = \mathbf{\Theta}^H\mathbf{S}_{\mathbf{D}}\mathbf{\Theta}$ and $\mathbf{Q} = \mathbf{\Theta}^H\mathbf{S}_{\mathbf{Q}}\mathbf{\Theta}$ with $\mathbf{S}_{\mathbf{R}} = \mathbf{S}_{\mathbf{C}} + \frac{\sigma^2}{P_p\alpha^2}\ \mathbf{I}$, $\mathbf{S}_{\mathbf{D}} = \mathbf{S}_{\mathbf{C}}\mathbf{S}_{\mathbf{R}}^{-1}$, and $\mathbf{S}_{\mathbf{Q}} = \mathbf{S}_{\mathbf{C}} - \mathbf{S}_{\mathbf{C}}\mathbf{S}_{\mathbf{R}}^{-1}\mathbf{S}_{\mathbf{C}}$ where matrices $\mathbf{S}_{\bullet}$ are real non-negative diagonal matrices. Specifically, we will refer to the diagonal elements of $\mathbf{S}_{\mathbf{D}}$ and $\mathbf{S}_{\mathbf{R}}$ using the notations $d_k = \mathbf{S}_{\mathbf{D}kk}$ and $r_k = \mathbf{S}_{\mathbf{R}kk}$, respectively. Let $\mathbf{v} = \mathbf{\Theta}\hat{\mathbf{h}}$, then $\mathbf{v}$ is a random vector with distribution $\mathcal{CN}(\mathbf{0},\mathbf{S}_{\mathbf{R}})$, since

$$\begin{aligned}\mathcal{E}(\mathbf{v}\mathbf{v}^H) &= \mathcal{E}(\mathbf{\Theta}\hat{\mathbf{h}}\hat{\mathbf{h}}^H\mathbf{\Theta}^H) = \mathbf{\Theta}\mathcal{E}(\hat{\mathbf{h}}\hat{\mathbf{h}}^H)\mathbf{\Theta}^H = \\ &= \mathbf{\Theta}\mathbf{R}\mathbf{\Theta}^H = \mathbf{\Theta}\mathbf{\Theta}^H\mathbf{S}_{\mathbf{R}}\mathbf{\Theta}\mathbf{\Theta}^H = \mathbf{S}_{\mathbf{R}}\ .\end{aligned}$$

That is, the elements of $\mathbf{v}$ are independent, but they have different variances.

C. Terms of $\mathcal{E}\{MSE\} = \mathcal{E}_{\hat{\mathbf{h}}}\left\{MSE(\hat{\mathbf{h}})\right\}$

To compute $\mathcal{E}_{\hat{\mathbf{h}}}\left\{\text{MSE}(\hat{\mathbf{h}})\right\}$ based on Lemma 2, we need to calculate the following terms (T_1, T_2 and T_3):

$$\begin{aligned}\mathcal{E}_{\hat{\mathbf{h}}}\left\{\text{MSE}(\hat{\mathbf{h}})\right\} = &\underbrace{\alpha^2 P\ \mathcal{E}_{\hat{\mathbf{h}}}\left\{z^2\ \hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\hat{\mathbf{h}}^H\mathbf{D}^H\hat{\mathbf{h}}\right\}}_{T_1} + \\ &+\underbrace{\mathcal{E}_{\hat{\mathbf{h}}}\left\{z^2\ \hat{\mathbf{h}}^H\left(\alpha^2 P\mathbf{Q}+\sigma^2\mathbf{I}\right)\hat{\mathbf{h}}\right\}}_{T_2} + \\ &\underbrace{-2\alpha\sqrt{P}\ \mathcal{E}_{\hat{\mathbf{h}}}\left\{z\ Re\{\hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\}\right\}}_{T_3} + 1.\end{aligned}$$

With Lemmas 1, 2 and 3 in our hands, we can state our main result on calculating the terms T_1, T_2 and T_3.

Theorem *1:* The mean square error (MSE) of the uplink received data with arbitrary correlation matrix $\mathbf{C}$ of the uplink channel can be calculated as in the sum of the terms T_1, T_2 and T_3 plus 1, where

$$\begin{aligned}T_1 = &\sum_k\sum_{\ell,\ell\neq k} d_k d_\ell\ \cdot\\ &\cdot\int_{x=0}^{\infty} xe^{-x\sigma^2/(\alpha^2P)}\frac{1}{x+r_k}\ \frac{1}{x+r_\ell}\ \prod_i\frac{r_i}{x+r_i}\ dx+\\ &+\sum_k d_k^2\int_{x=0}^{\infty} xe^{-x\sigma^2/(\alpha^2P)}\ \frac{2}{(x+r_k)^2}\ \prod_i\frac{r_i}{x+r_i}\ dx;\end{aligned}$$

$$T_2 = \frac{1}{\alpha^2 P}\sum_k m_k\int_{x=0}^{\infty} xe^{-x\frac{\sigma^2}{\alpha^2P}}\frac{1}{x+r_k}\prod_i\frac{r_i}{x+r_i}dx;$$

and

$$T_3 = 2\ \sum_k d_k\ \int_{x=0}^{\infty} e^{-x\frac{\sigma^2}{\alpha^2P}}\frac{1}{x+r_k}\prod_i\frac{r_i}{x+r_i}dx,$$

where $\mathbf{S}_{\mathbf{M}} = \alpha^2 P\mathbf{S}_{\mathbf{Q}}+\sigma^2\mathbf{I}$ is a diagonal matrix with diagonal elements $m_k = \mathbf{S}_{\mathbf{M}kk} = \alpha^2 Pq_k + \sigma^2$, where $q_k = \mathbf{S}_{\mathbf{Q}kk}$.

The proof of Theorem 1 is omitted due to the space limitation but can be found in [14].

VI. Numerical Results

A. Channel Model and Covariance Matrix

In this section we consider a single cell system, in which MSs use orthogonal pilots to facilitate the estimation of the uplink channel by the BS. Recall from Section III that the channel estimation process is independent for each MS and we can therefore focus on a single user. The covariance matrix $\mathbf{C}$ of the channel $\mathbf{h}$ as the function of the antenna spacing,

TABLE I
SYSTEM PARAMETERS

Parameter	Value
Number of antennas	$N_r = 2, 4, 8, 10, 20$
Path Loss	$\alpha = 50, 55, 60$ dB
Power budget	$\tau_p P_p + \tau_d P = P_{tot}$ mW, as in Eq. (2).
Antenna spacing	$D/\lambda = 0.15, ..., 1.5$
Mean Angle of Arrival (AoA)	$\theta = 70°$
Angular spread	$2 \cdot \theta_\Delta = 5, ..., 45°$

mean angle of arrival and angular spread is modeled as by the well known spatial channel model, which is known to be accurate in non-line-of-sight environment with rich scattering and all antenna elements identically polarized, see [15]. For uniformly distributed angle of arrivals, the (m, n) element of the covariance matrix $\mathbf{C}$ are given by

$$\mathbf{C}_{m,n} = \frac{1}{2\theta_\Delta} \int_{-\theta_\Delta}^{\theta_\Delta} e^{j \cdot 2\pi \cdot \frac{D}{\lambda}(n-m)\cos(\bar{\theta}+x)} dx,$$

where the system parameters are given in Table I. The covariance matrix $\mathbf{C}$ becomes practically diagonal as the antenna spacing and the angular spread grows beyond $D\lambda > 1$ and $\theta_\Delta > 30°$. In contrast, with critically spaced antennas $D\lambda = 0.5$ and $\theta_\Delta < 10°$, the antenna correlation in terms of the off-diagonal elements of $\mathbf{C}$ can be considered strong. We included a validation of the proposed channel model with data from a realistic channel simulator in [14].

B. Numerical Results

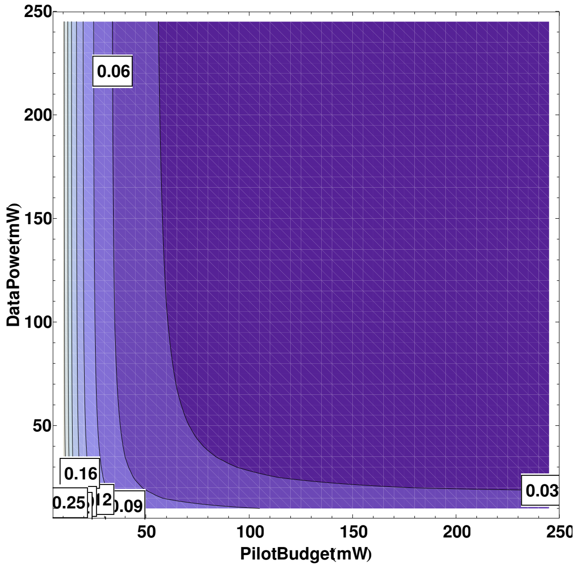

Fig. 1. Contour plot of the MSE achieved by specific pilot and data power settings of a SIMO system with $N_r = 20$ *uncorrelated* receiver antennas at a fix path loss position of 50 dB. For example an MSE value less than 0.03 can be reached by setting the pilot power budget to $\tau_p P_p = 100$ mW and the data power $P = 60$ mW, or by $P_p = 200$ mW and $P = 30$ mW. We can see that with a total power budget of 250 mW, and with proper pilot-data balancing, a minimum MSE that is clearly less than 0.03 can be reached.

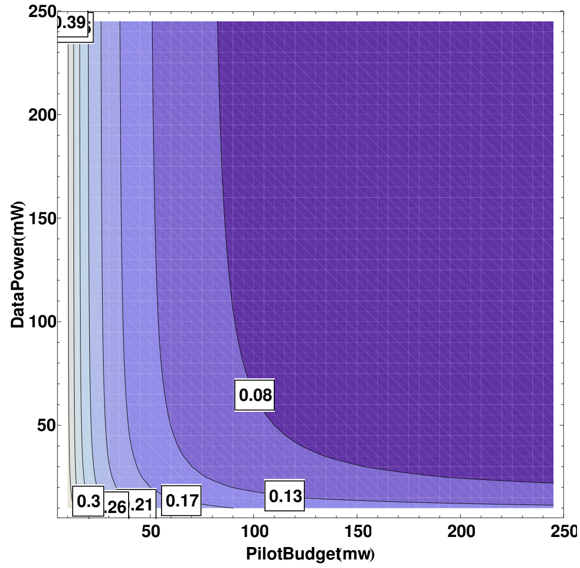

Fig. 2. Contour plot of the MSE achieved by specific pilot and data power settings of a SIMO system with $N_r = 20$ *correlated* receiver antennas at a fix path loss position of 50 dB. Compared with Figure 1, we can see that with similar sum power budget, the MSE value that can be reached is somewhat higher. For example, with a power budget of 250 mW, an MSE value that is less than 0.08 can be realized ($\tau_p P_p = 150$ mW and $P = 100$ mW).

Recall that the MSE of the received data symbols according to Theorem 1 depends on the pilot through the product $t_p P_p$ (which we call the pilot budget) and the data power P for each transmitted data symbol. Figures 1 and 2 compare the MSE that can be achieved by a particular setting of the pilot and data power levels in the case of uncorrelated and highly correlated antennas (for the case with $N_r = 20$ receive antennas). The impact of high antenna correlation is that in order to reach the same MSE, a higher power level for both the pilot and data transmission must be used.

Figures 3 and 4 compare the MSE that can be achieved by a particular setting of the data power levels in the case of uncorrelated and highly correlated antennas (for the case with $N_r = 20$ receive antennas) at different path loss positions. Although the impact of high antenna correlation is clearly high, the proper setting of the pilot-to-data-power ratio has a more pronounced effect.

Figure 5 shows the percentage of the total power budget that minimizes the MSE as a function of the path loss between the MS and the BS when the number of BS antennas is varied as $N_r = 2, 4, 8, 10, 20$. Here we can clearly see the at a given position (path loss), the pilot-to-data power ratio (PDPR) increases with the number of antennas. The intuition behind this behavior is that at a larger number of antennas, the data transmit power can be lower than with a few antennas and a higher pilot power level (a higher PDPR) can be afforded by the MS. Interestingly, the PDPR that minimizes the MSE, in this example (and also in other examples not reported here) is virtually the same for the uncorrelated and correlated antenna cases.

The value of the minimum MSE, assuming the optimal

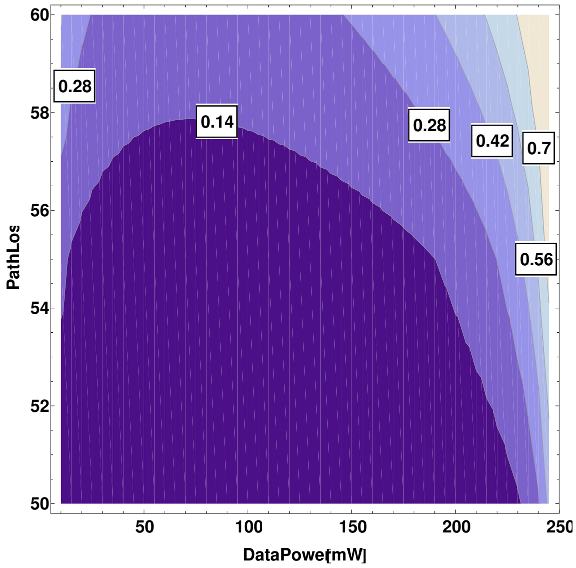

Fig. 3. Contour plot of the MSE as the function of the data power and the path loss under a total power (250 mW) constraint with $N_r = 20$ uncorrelated antennas. For example, with the near optimal data power setting and MSE value of 0.14 can be reached at 58 dB path loss.

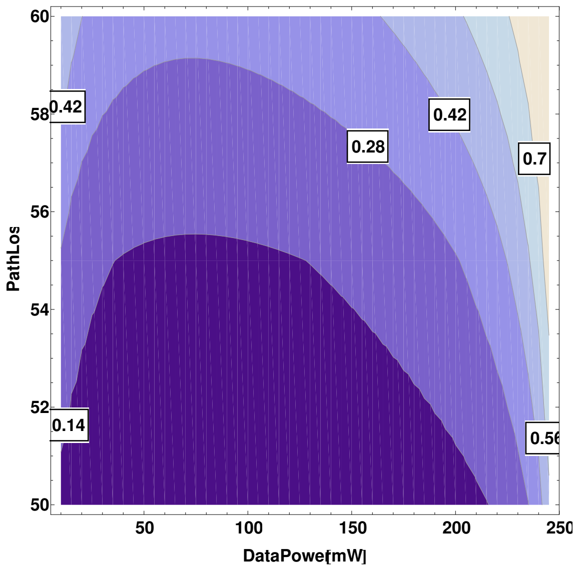

Fig. 4. Contour plot of the MSE as the function of the data power and the path loss under a total power (250 mW) constraint with $N_r = 20$ highly correlated antennas. For example, with the near optimal data power setting and MSE value of 0.14 can be reached at 55 dB path loss.

PDPR setting is shown by Figure 6 and Figure 7 for the case of uncorrelated and correlated antennas respectively. By examining Figures 6-7, we can see that the impact of (high) antenna correlation is significant. For example with $N_r = 20$ receive antennas, at 50 dB path loss, the minimum MSE (that is with the optimal PDPR) is around 0.004, whereas with correlated antennas, the minimum MSE is around 0.010.

Next, we examine the impact of antenna spacing and compare that with the impact of PDPR setting on the MSE in Figure 8. In this figure, the uppermost and the middle curves

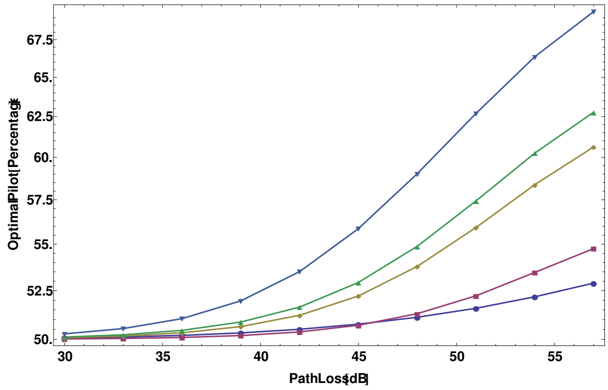

Fig. 5. The percentage of the total power budget that minimizes the MSE as a function of the path loss for $N_r = 2, 4, 8, 10, 20$ antennas. ($N_r = 20$ corresponds to the uppermost curve.) At a given path loss, the percentage of the power budget that minimizes the MSE increases with the number of antennas.

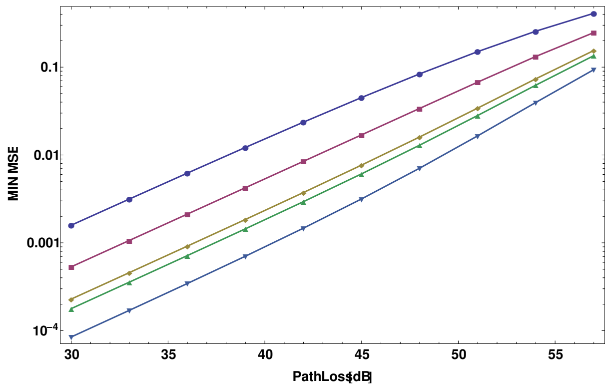

Fig. 6. The minimum MSE that can be reached by setting the percentage of the pilot power according to Figure 5 when the antennas are uncorrelated. ($N_r = 20$ corresponds to the lowermost curve.) For example, at path loss 50 dB an order of magnitude lower MSE can be reached with $N_r = 20$ antennas than with $N_r = 2$.

correspond to suboptimal PDPR settings. The insight provided by this figure is important because it suggests that with correct PDPR setting, the negative impact of antenna corrleation can be compensated. In fact, the correct setting of the PDPR is more important than decreasing the antenna correlation by, for example, increasing the antenna spacing between antenna elements.

VII. Conclusions

In this paper we developed a model of a single cell system, in which MSs use orthogonal pilots to facilitate uplink channel estimation by the BS. We developed a methodology to calculate the MSE of the uplink equalized data symbols and derived a closed form for the MSE as a function of not only the number of antennas, the pilot power and the data transmit power, but also the path loss and other parameters that determine the covariance matrix of the fast fading channel between the MS and the BS. With this methodology, through numerical examples, we found that although the impact of antenna correlation on the MSE performance can be significant, this

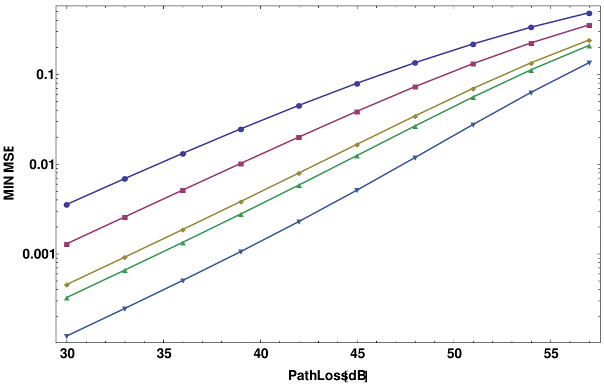

Fig. 7. The minimum MSE that can be reached by setting the percentage of the pilot power according to Figure 5 when the antennas are correlated. ($N_r = 20$ corresponds to the lowermost curve.) Compared with Figure 6, we can see that the MSE values get somewhat higher.

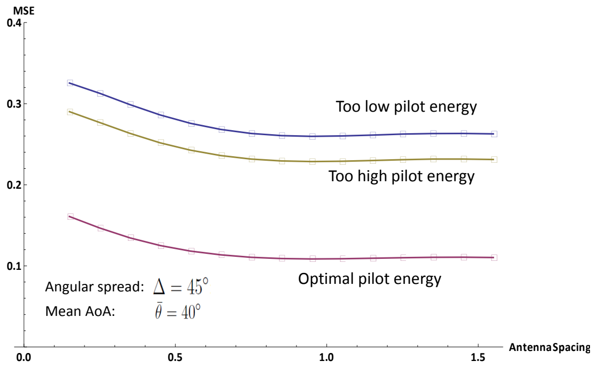

Fig. 8. The impact of antenna spacing ($N_r = 20$) on the MSE as a function of antenna element spacing (in terms of multiple of the wavelength λ) at the constant path loss of 50 dB. The three curves correspond to the optimal pilot power setting (lowest curve), too high pilot power (middle) and too low pilot power levels (uppermost).

impact can be compensated by setting the correct pilot-to-data power ratio (PDPR) in case of a total power budget. Furthermore, we found that as the number of antennas grows, a higher ratio of the power budget should be spent on pilots, virtually independently of the antenna correlation. This can be seen in line with the findings of massive MIMO systems that suggest that the data transmit power at the MS can be significantly lower as the number of antennas at the BS grows large.

Appendix

A. Proof of Lemma 2

$$\begin{aligned}\text{MSE}(\hat{\mathbf{h}}) &= z\ \hat{\mathbf{h}}^H\Big(\alpha^2 P(\mathbf{D}\hat{\mathbf{h}}\hat{\mathbf{h}}^H\mathbf{D}^H + \mathbf{Q}) + \sigma^2\mathbf{I}\Big)z\ \hat{\mathbf{h}} - \\ &- 2\alpha\sqrt{P}Re\{z\ \hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\} + 1 = \\ &= z^2\alpha^2 P\ \hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\hat{\mathbf{h}}^H\mathbf{D}^H\hat{\mathbf{h}} + z^2\alpha^2 P\ \hat{\mathbf{h}}^H\mathbf{Q}\hat{\mathbf{h}} + \\ &+ z^2\sigma^2\hat{\mathbf{h}}^H\hat{\mathbf{h}} - 2z\alpha\sqrt{P}Re\{\hat{\mathbf{h}}^H\mathbf{D}\hat{\mathbf{h}}\} + 1.\end{aligned}$$

B. Proof of Lemma 3

First, we notice that according to equation 3.326 (2) of [16]:

$$\int_{x=0}^{\infty} x^m e^{-yx^n}dx = \frac{\Gamma(\gamma)}{ny^\gamma}; \quad \gamma = \frac{m+1}{n},$$

which specifically for $n = 1$ means:

$$\frac{1}{y^\gamma} = \int_{x=0}^{\infty}\frac{x^{\gamma-1}}{\Gamma(\gamma)}e^{-xy}dx.$$

that is, for $\gamma = 1$ and $y = \|\hat{\mathbf{h}}\|^2 + \sigma^2/(\alpha^2 P)$:

$$\begin{aligned}z &= \frac{\alpha\sqrt{P}}{\|\hat{\mathbf{h}}\|^2\alpha^2 P + \sigma^2} = \frac{1}{\alpha\sqrt{P}}\ \frac{1}{\|\hat{\mathbf{h}}\|^2 + \sigma^2/(\alpha^2 P)} = \\ &= \frac{1}{\alpha\sqrt{P}\,\underbrace{\Gamma(1)}_{1}}\cdot\int_{x=0}^{\infty} e^{-x\left(\hat{\mathbf{h}}^H\hat{\mathbf{h}} + \sigma^2/(\alpha^2 P)\right)}dx\end{aligned}$$

and, for $\gamma = 2$:

$$\begin{aligned}z^2 &= \frac{\alpha^2 P}{(\|\hat{\mathbf{h}}\|^2\alpha^2 P + \sigma^2)^2} = \\ &= \frac{1}{\alpha^2 P\,\underbrace{\Gamma(2)}_{1}}\cdot\int_{x=0}^{\infty} x e^{-x\left(\hat{\mathbf{h}}^H\hat{\mathbf{h}} + \sigma^2/(\alpha^2 P)\right)}dx.\end{aligned}$$

References

[1] T. Kim and J. G. Andrews, "Optimal Pilot-to-Data Power Ratio for MIMO-OFDM," *IEEE Globecom*, 2005.

[2] ——, "Balancing Pilot and Data Power for Adaptive MIMO-OFDM Systems," *IEEE Globecom*, 2005.

[3] E. Golovins and N. Ventura, "Optimal Training for the SM-MIMO-OFDM Systems with MMSE Channel Estimation," *6th Annual Comm. Networks and Services Research Conf.*, pp. 470–477, 2008.

[4] V. K. V. Gottumukkala and H. Minn, "Capacity Analysis and Pilot-Data Power Allocation for MIMO-OFDM With Transmitter and Receiver IQ Inbalances and Residual Carrier Frequency Offset," *IEEE Trans. Veh. Techn.*, pp. 553–565, 2012.

[5] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhear in multipath fading channels," *IEEE Transactions on Communications*, vol. 58, no. 10, pp. 2939–2948, October 2010.

[6] K. T. Truong, A. Lozano, and R. W. H. Jr., "Optimal training in continuous block-fading massive mimo systems," 20^{th} *European Wireless, Barcelona, Spain*, May 2014.

[7] G. Fodor and M. Telek, "On the pilot-data power trade off in simo systems," *http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-139408*, January 2014.

[8] ——, "On the Pilot-Data Power Trade Off in Single Input Multiple Output Systems," *European Wireless, Barcelona, Spain*, May 2014.

[9] P. Soldati, M. Johansson, G. Fodor, and S. Sorrentino, "On Pilot Dimensioning in Multicell Single Input Multiple Output Systems," *IEEE Workshop on Broadband Wireless Access*, Houston, TX, USA,2011.

[10] R. Chen, J. G. Andrews, R. W. Heath, and A. Ghosh, "Uplink Power Control in Multi-cell Spatial Multiplexing Wireless Systems," *IEEE Trans. on Wireless Comm.*, vol. 6, no. 7, pp. 2700–2711, July 2007.

[11] *3GPP Technical Specification Group RAN: Spatial Channel Model for MIMO Simulations*, 2003. [Online]. Available: http://3gpp.org

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*. Prentice Hall, 1993, no. ISBN: 0133457117.

[13] R. Gallager, "Circularly-Symmetric Gaussian Complex Vectors," *http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf*, 2008.

[14] G. Fodor, M. Telek, and P. Di Marco, "On the impact of antenna correlation on the pilot-data balance in multiple antenna systems," *http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-143530*, March 2014.

[15] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” *http://arxiv.org/abs/1203.5924*, no. arXiv:1203.5924v1, March 2012.
[16] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, 2007, no. ISBN-13: 978-0-12-373637-6.