

Analysis of Pilot Decontamination Based on Power Control

Laura Cottatellucci
EURECOM

Mobile Communications Department
BP193, F-06560 Sophia Antipolis, France
Email: cottatel@eurecom.fr

Ralf R. Müller

Norwegian University of Science & Tech.
Dept Electronics & Telecommunications
7491 Trondheim, Norway
Email: mueller@iet.ntnu.no

Mikko Vehkaperä
Aalto University

Dept Signal Processing & Acoustics
02150 Espoo, Finland
Email: mikkov@kth.se

Abstract—A subspace method for channel estimation is proposed for asymmetric antenna array systems. The so-called pilot contamination problem reported in [1] is found to be due to the linearity of channel estimation in [2]. We show that it does not occur in cellular systems with power control and power-controlled handoff when the nonlinear channel estimation method proposed in this paper is used. Power-control hand-off is needed to guarantee separability between signal and interference subspaces. We derive the transmission conditions for subspace separability based on free probability and perturbation theory.

Index Terms—Multiple antennas, multiple-input multiple-output (MIMO) systems, massive MIMO, spread-spectrum, channel estimation, principal component analysis

I. INTRODUCTION

Two pioneering works [1], [2] introduced the concept of large antenna array networks, commonly referred to as *massive MIMO*, and showed that the use of very large antenna arrays in multiuser networks is even more appealing than the use of single user MIMO systems. Under the assumptions that the number of antennas at the access nodes is much higher than the number of served user terminals, the effects of uncorrelated noise and fast fading can be made arbitrarily small (vanishing as the number of antennas goes to infinity). More interestingly, with perfect channel state information (CSI), the total radiated power can be made inversely proportional to the number of radiating antennas without degradation of the quality of the received signals, and, as the number of antennas increases and radiated power is reduced, simple single cell processing becomes asymptotically optimum. The total throughput is asymptotically independent of the cell-size, and the spectral efficiency is independent of the bandwidth. This system design has attracted considerable attention, see e.g. [3] and references therein, and other appealing features have been pointed out (e.g. [9], [10]).

According to the results in [2], the only impairment that does not vanish in asymptotic conditions is the inter-cell interference due to the re-use of orthogonal pilot sequences in other cells. This impairment is dubbed *pilot contamination* and is based on the implicit assumption of linear channel estimation in [2, Eq. (5)]. As a result, orthogonal pilot interference from interfering cells would limit the accuracy of the channel estimation regardless of the number of antenna elements at the access point. This effect was believed by many researchers, e.g. [4]–[6] to be a fundamental effect, despite the lack of a solid proof that it cannot be overcome.

Using Bayesian channel estimation, [7] found that pilot contamination can vanish under the condition that the channel angular spreads of active users do not overlap.

In this paper, we show that pilot contamination is not a fundamental feature of large antenna array networks, but a shortcoming of linear channel estimation. We show, that the array gain can easily be utilized to obtain a channel estimation whose accuracy grows unboundedly with the number of antennas. Our approach is based on the utilization of randomly generated pilot sequences in all the base stations. Under this assumption, the transmitted signals (pilots and data) become asymptotically (i.e. when the number of antennas in the arrays grows large) orthogonal. More interestingly, the signal subspace of each cell becomes orthogonal to the interference subspace. Then, a projection of the received signal onto the signal subspace would remove interference from adjacent cells including pilot contamination. The remaining problem is to distinguish the signal subspace from the interference subspace at each access point. We propose a low complexity approach based on power-controlled hand-off. Thanks to power-controlled hand-off, we keep the powers of subspace of the signals of interest in a range non overlapping with the range of the powers of interferers' subspace. This allows to identify the signal subspace very easily and perform projection. The complexity of subspace identification and projection has polynomial complexity in the number of antenna elements.

Additionally, we analyze the network under the assumption of Rayleigh fading and provide expressions for the ranges of the signals of interest and interference based on free probability and perturbation theory. Transmission conditions for nonzero gap between the two ranges are also provided. Simulations and numerical results corroborate the effectiveness of the proposed method and the correctness of the corresponding analysis.

This contribution is structured as follows. In Section II, we introduce the system model. In Section III, we propose an algorithm for nonlinear channel estimation utilizing the array gain. In Sections IV and V, we investigate the performance of this algorithm by analytic and simulative means, respectively. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODEL

We consider a wireless cellular system in uplink with a cell of interest and L interfering cells. Each access point (base station) in a cell is equipped with R receive antennas and each cell is populated by T active transmit antennas with $T < R$.

By observing that a channel with bandwidth larger than the coherence bandwidth can be always decomposed into equivalent parallel narrowband channels with bandwidth narrower than the coherence bandwidth by orthogonal frequency division multiplexing or related techniques, without loss of

generality we focus on a frequency flat block fading channel. A coherence time interval consists of C symbol intervals. The baseband signal received at the access point of interest during a coherence time interval is given by

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{H}_I\mathbf{X}_I + \mathbf{W}, \quad (1)$$

\mathbf{Y} is the $R \times C$ matrix of received signals in $\mathbb{C}^{R \times C}$, \mathbf{X} is the $T \times C$ matrix of transmitted data of interest in $\mathbb{C}^{T \times C}$. It contains information symbols eventually multiplexed with pilot symbols. Similarly, \mathbf{X}_I is the $LT \times C$ matrix of transmitted symbols by the LT active transmit antennas in the L interfering cells. The $R \times T$ matrix \mathbf{H} consists of the unknown propagation coefficients from antennas in the cell of interest to the access point of interest. Similarly, \mathbf{H}_I is the $R \times LT$ matrix of the propagation coefficients from interfering antennas to the access point of the cell of interest. Finally, \mathbf{W} is the $R \times C$ matrix of the white additive Gaussian noise.

In order to keep exposition and derivations simple, we assume that $C \geq R$. However, it is worth noticing that the results presented in the following hold in general also for $C < R$ although their derivation may require some modification.

For the sake of analysis, we assume that the entries of the matrices \mathbf{X} , \mathbf{X}_I , \mathbf{H} , \mathbf{H}_I , and \mathbf{W} are independent and identically distributed (i.i.d.) zeromean complex random variables. The entries of \mathbf{X} and \mathbf{X}_I have unit variance while the elements of \mathbf{H} and \mathbf{H}_I , have variance P and I , respectively. The variance of the additive white noise is W .

We define the normalized coherence time

$$\kappa = \frac{C}{R} \quad (2)$$

and the cell load

$$\alpha = \frac{T}{R}. \quad (3)$$

Additionally, sometimes, it is useful to consider a unique matrix of noise including both the thermal noise and interference

$$\mathbf{Z} = \mathbf{H}_I\mathbf{X}_I + \mathbf{W} \quad (4)$$

such that (1) reduces to

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}. \quad (5)$$

Note that (5), can also be understood as a code-division multiple-access (CDMA) system with the columns of \mathbf{H} denoting the spreading sequences and R denoting the processing gain. It is well-known that CDMA can be demodulated without knowledge of the spreading sequences by means of blind algorithms, see e.g. [8]. Those algorithms can also be applied in massive MIMO systems as proposed in [9]. In the following section, we introduce an algorithm particularly suitable for large scale antenna array systems.

III. PROPOSED ALGORITHM

Subspace methods are very suitable for dimensionality and noise reduction. They rely on the underlying assumption that the useful information is contained in a subspace with smaller dimensions than the space of observations. A significant part of the total signal power is concentrated in such a subspace, whereas additive white noise is typically distributed through the larger space isotropically. Then, by projecting the received signal onto the signal subspace, it is possible to reduce the

dimensionality of the problem while removing a considerable amount of noise.

In the simple case of a single active transmit antenna in additive noise, i.e. $T = 1$ and $L = 0$ in (1), the linear receiver \mathbf{m}^* that maximizes the signal to noise ratio (SNR) is equivalent the one that maximizes the total received power normalized to the power gain of the filter, i.e.,

$$\mathbf{m}^* = \operatorname{argmax}_m \frac{\mathbf{m}^\dagger \mathbf{Y} \mathbf{Y}^\dagger \mathbf{m}}{\mathbf{m}^\dagger \mathbf{m}}. \quad (6)$$

It is a well-known result of linear algebra that the vector \mathbf{m}^* maximizing the right hand side of (6) is the eigenvector of $\mathbf{Y} \mathbf{Y}^\dagger$ corresponding to the largest eigenvalue.

Similarly, for T transmit antennas in white noise, i.e. $L = 0$ in (1), the projection onto the signal subspace requires the Singular Value Decomposition (SVD) of the matrix \mathbf{Y} with singular values sorted in non-increasing order, i.e.

$$\mathbf{Y} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\dagger \quad (7)$$

with unitary matrices $\mathbf{U} \in \mathbb{C}^{R \times R}$ and $\mathbf{V} \in \mathbb{C}^{C \times C}$ and the $R \times C$ diagonal matrix $\mathbf{\Sigma}$ with diagonal entries $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_R$. Then, it is performed the decomposition of the matrix of left singular vectors into the signal space basis $\mathbf{S} \in \mathbb{C}^{R \times T}$ and the null space basis $\mathbf{N} \in \mathbb{C}^{R \times (R-T)}$, i.e.

$$\mathbf{U} = [\mathbf{S} | \mathbf{N}]. \quad (8)$$

Finally, the received signal is projected onto the signal subspace to obtain

$$\begin{aligned} \tilde{\mathbf{Y}} &= \mathbf{S}^\dagger \mathbf{Y} \\ &= \mathbf{S}^\dagger \mathbf{H} \mathbf{X} + \mathbf{S}^\dagger \mathbf{W} \\ &= \tilde{\mathbf{H}} \mathbf{X} + \tilde{\mathbf{W}} \end{aligned} \quad (9)$$

where $\tilde{\mathbf{Y}}$ is a $T \times C$ matrix with dimensions reduced compared to \mathbf{Y} and does not contain the noise lying in the null space \mathbf{N} ; $\tilde{\mathbf{H}} = \mathbf{S}^\dagger \mathbf{H}$ is a $T \times T$ matrix and $\tilde{\mathbf{W}} = \mathbf{S}^\dagger \mathbf{W}$ is the noise component in the projection subspace. Then, the channel estimation problem reduces to the estimation of the $T \times T$ channel matrix $\tilde{\mathbf{H}}$.

In the following, we discuss the implications that the assumption of large antenna array system, i.e., $R \gg T$ ideally $\alpha \rightarrow 0$, has on the application of a projection subspace method:

- The T -dimensional signal subspace is much smaller than the R -dimensional full space, which the noise lives in. White noise is evenly distributed in all dimensions of the full space. Thus, the influence of white noise onto the signal subspace becomes negligible as $R \rightarrow \infty$. Using the algorithm above, we can achieve an array gain even without the need for estimating the channel coefficients. In fact, the channel can be estimated solely after projection onto the signal subspace when the dominant part of the white noise has already been suppressed;
- Although the data dependent projection (9) implies that the noise $\tilde{\mathbf{W}} \in \mathbb{C}^{T \times C}$ is not independent from the data \mathbf{X} , neglecting this dependence is an admissible approximation that becomes exact as the number of receive antennas R grows large;
- Any R -dimensional channel vector from any transmitter, be it in the cell of interest or in a neighboring cell, to the receive array in the cell of interest becomes orthogonal to

any other channel vector as $R \rightarrow \infty$. This implies that the subspace spanned by transmit antennas in interfering cells is asymptotically orthogonal to the subspace spanned by the signal of interest. Then, in the cases where we are able to discriminate between the subspace of the interfering signals and the subspace of the signals of interest, the proposed projection subspace method is applicable to system (1) even with $L > 0$ by replacing \mathbf{W} and $\tilde{\mathbf{W}}$ in (9) with \mathbf{Z} and $\tilde{\mathbf{Z}} = \mathbf{S}^\dagger \mathbf{Z}$, respectively.

We conclude this section discussing conditions and techniques to identify and distinguish between interference subspace and subspace of the signal of interest. To this aim, we observe that, for $R \rightarrow \infty$ and $\alpha \rightarrow 0$ and in absence of noise, the $(L+1)T$ largest singular values of the received signal matrix \mathbf{Y} become identical to the Euclidean norms of the $(L+1)T$ channel vectors multiplied by the average amplitude of the transmitted signals, i.e. \sqrt{P} or \sqrt{I} for the signal of interest or the interferers. Then, if we keep these values in two orthogonal intervals for the two different sets of signals, we are able to identify the two subspaces by the singular values of the matrix \mathbf{Y} .

This is possible by power control and power controlled hand-off. Note that for $R \rightarrow \infty$, the system has infinite diversity, thus the effect of short-term fading (Rayleigh fading) vanishes. Thus, the norm of a channel vector is solely determined by path loss and long-term fading (shadowing). In a cellular system with power control and power-controlled handoff strategy, the norm of channel vectors from neighboring cells can never be greater than the norm of channel vectors from the cell of interest. We conclude that the identification of singular values belonging to transmitters within the cell of interest is possible by means of ordering them by magnitude. Admittedly, such orderings might be inaccurate for a small fraction of transmitters that happen to experience similar channel conditions to more than a single access point. In practice, this ambiguity can be overcome by a smart scheduling that ensure a certain margin of power separation between users within the cell of interest and users from neighboring cells.

IV. PERFORMANCE ANALYSIS

In the previous section we showed that the interference subspace and the user-of-interest subspace are orthogonal for $R \rightarrow \infty$ and $\alpha \rightarrow 0$. Additionally, adopting power control, power-controlled hand-off and a smart scheduling, it is possible to maintain the signal powers of the two subspaces into two disjoint intervals and distinguish and separate the signal-of-interest subspace based only on the singular values of the matrix \mathbf{Y} . However, in practical systems the number of transmit and receive antennas is finite and the load α can be made very small but not arbitrarily small. Then, in real systems the asymptotic properties are only approximated. A useful and insightful approach to understand the behaviour of a real network consists in assuming that both T and R grow large with a small but finite ratio α . This setting can be studied effectively by random matrix theory. Intuitively, while for $\alpha \rightarrow 0$ all the singular values related to the signals-of-interest and interferers are equal to \sqrt{P} and \sqrt{I} , respectively, in the practical case with finite α the singular values get spread around the ideal values. Objective of our analysis is to determine the distribution of the singular values of \mathbf{Y} , or equivalently the eigenvalue distribution of $\mathbf{Y}\mathbf{Y}^\dagger$ for small but finite α and $R, T \rightarrow \infty$.

Let us denote the asymptotic eigenvalue distribution of $\mathbf{Y}\mathbf{Y}^\dagger$ as $P(x)$. In [11], we showed that this asymptotic eigenvalue distribution obeys

$$sG(s) = -1 - \frac{PTC\alpha(sG(s) + 1 - \kappa)G(s)}{\alpha\kappa - PTC(sG(s) + 1 - \kappa)G(s)} - \frac{ILTC\alpha(sG(s) + 1 - \kappa)G(s)}{\alpha\kappa - ITC(sG(s) + 1 - \kappa)G(s)} - \frac{WC(sG(s) + 1 - \kappa)G(s)}{\kappa} \quad (10)$$

with

$$G(s) = \int \frac{dP(x)}{x-s} \quad (11)$$

denoting its Stieltjes transform.

By means of the Stieltjes inversion formula

$$p(x) = \frac{1}{\pi} \lim_{y \rightarrow 0^+} \Im G(x + jy) \quad (12)$$

the asymptotic eigenvalue distribution is obtained.

The fixed point equation (10) is not insightful to determine under which transmission conditions the two bulks of the eigenvalues related to the signal-of-interest and interference subspaces are disjoint. In the following, we derive conditions under which the two bulks are disjoint.

In order to find the support of asymptotic eigenvalue distributions we follow [12, Chapter 7]. There it is shown that the boundaries of the support of the asymptotic eigenvalue distribution are extrema of the inverse of the Stieltjes transform. Note that the fixed point equation for the Stieltjes transform in (10) can be rewritten as

$$sG + 1 + \frac{\zeta(sG + 1 - \kappa)G}{\kappa} + \frac{\alpha(sG + 1 - \kappa)G}{r\kappa - (sG + 1 - \kappa)G} + \frac{\alpha L(sG + 1 - \kappa)G}{t\kappa - (sG + 1 - \kappa)G} = 0 \quad (13)$$

with $r = (PRC)^{-1}$, $t = (IRC)^{-1}$, and $\zeta = WC$. The inverse function $s(G)$ can be obtained as a zero of the numerator of (13), $N(s) = N_0(s) + N_p(s)$ where

$$N_0(s) = (G(\kappa + \zeta G)s + \zeta(1 - \kappa)G + \kappa)(r\kappa - (sG + 1 - \kappa)G)(t\kappa - (sG + 1 - \kappa)G)$$

is a cubic function in s and

$$N_p(s) = \alpha\kappa G(sG + 1 - \kappa)(\kappa(t + Lr) - (L+1)(sG + 1 - \kappa)G),$$

is a quadratic function in s proportional to α . Then, $N(s)$ is a cubic function in s obtained as perturbation of N_0 by $N_p(s)$. Then, for very small α the zeros of the original numerator $N(s)$ can be computed as a perturbed version of the zeros in $N_0(s)$ given by

$$s_{0,1} = -\frac{\zeta(1 - \kappa)G + \kappa}{(\kappa + \zeta G)Gs} \quad (14)$$

$$s_{0,2} = -\frac{(1 - \kappa)G - \kappa r}{G^2} \quad (15)$$

$$s_{0,3} = -\frac{(1 - \kappa)G - \kappa t}{G^2}. \quad (16)$$

Due to lack of space, we do not go into the details of the derivation of the boundaries of the two bulks but just provide

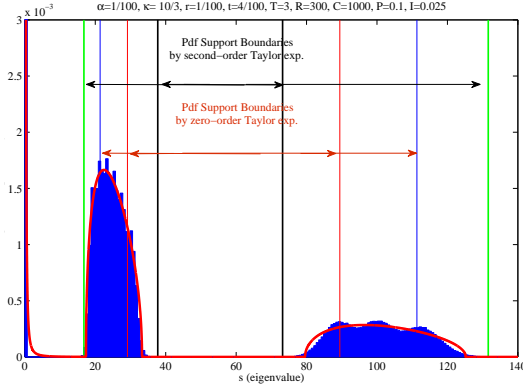


Fig. 1. Eigenvalue pdf of a network with $T = 3$, $L = 2$, $R = 300$, and $W = 0$ with approximation of its support boundaries

that guidelines for their derivation. In order to obtain a good approximation of $N(s)$ in the surroundings of its extremes, we consider the second order Taylor expansions of $N(s)$ around $s_{0,1}$, $s_{0,2}$ and $s_{0,3}$. Each of these expansions has two zeros $s_{\alpha}^{(2,1)}(G)$ and $s_{\alpha}^{(2,2)}(G)$ and provide a set of 6 functions whose extremes should be derived. To this aim, the expansions with respect to $s_{0,2}$ and $s_{0,3}$ are especially useful since they enable analytical approximations of the values of G corresponding to the extremes.

In the following, in order to keep equations shorter we set $\zeta = 0$, i.e. we consider the high SNR regime. However, similar expressions are available for any value of ζ . Then, for $\zeta = 0$ the proposed derivation of the extremes yield the two pairs

$$G_2^{(2)} = -\kappa r(t-r) \left(\frac{\kappa(t-r) + \alpha(t + (L-2)r)}{(\alpha t + \alpha Lr - \kappa t + \kappa r)^2 + 4\alpha\kappa Lr(t-r)} \pm \frac{2\sqrt{\alpha\kappa(t-r)^2 - \alpha^2 r(t + (L-1)r)}}{(\alpha t + \alpha Lr - \kappa t + \kappa r)^2 + 4\alpha\kappa Lr(t-r)} \right) \quad (17)$$

for the expansion around $s_{0,2}$ and

$$G_3^{(2)} = -\kappa t(t-r) \left(\frac{\kappa(t-r) + \alpha((2L-1)t - Lr)}{(\alpha t + \alpha Lr - \kappa t + \kappa r)^2 + 4\alpha\kappa Lr(t-r)} \pm \frac{2\sqrt{\alpha\kappa L(t-r)^2 + \alpha^2 L t((L-1)t - Lr)}}{(\alpha t + \alpha Lr - \kappa t + \kappa r)^2 + 4\alpha\kappa Lr(t-r)} \right) \quad (18)$$

for the expansion around $s_{0,3}$. It can be shown that under the assumptions of physical interest that $L \in \mathbb{N}^+$ and $t \geq r \geq 0$, $G_2^{(2)}$ and $G_2^{(3)}$ are all negative real zeros if

$$\frac{t}{r} \geq 1 + \frac{\alpha}{2\kappa} + \frac{\alpha}{2\kappa} \sqrt{1 + \frac{4L\kappa}{\alpha}} \quad (19)$$

or equivalently

$$0 \leq \frac{\alpha}{\kappa} \leq \frac{(t-r)^2}{r(t + (L-1)r)}. \quad (20)$$

For $x = 2, 3$, let us denote by $G_{x,-}^{(2)}$ and $G_{x,+}^{(2)}$ the instances of $G_x^{(2)}$ obtained by selecting the sign minus or plus, respectively. Then, by simple inspection of (17) and (18), we observe that

$$G_{x,+}^{(2)} \leq G_{x,-}^{(2)}. \quad (21)$$

By inspection, it is easy to verify that $G_{3,-}^{(2)} \leq G_{2,+}^{(2)}$ under the above mentioned conditions of physical interest. However, we are interested in determining under which conditions the two intervals $[G_{2,+}^{(2)}, G_{2,-}^{(2)}]$ and $[G_{3,+}^{(2)} \leq G_{x,-}^{(3)}]$ do not intersect, i.e. when $G_{3,-}^{(2)} \leq G_{2,+}^{(2)}$. It can be verified that this last condition is satisfied for

$$0 \leq \frac{\alpha}{\kappa} \leq \frac{(Lr^2 + 3Lrt + 3rt + t^2 - 2(t+r)\sqrt{3Lrt})}{(t-r)^2} \cdot \frac{1}{L^2 r^2 (3t+r)^2 - 6r^3 t L - 4r^2 L t^2 - 6L r t^3 + t^2 (3r+t)^2}.$$

Finally, let us define

$$s_2^{(2)}(G) = \frac{2\alpha\kappa(L+1) - 2\alpha(L+1) + 2\kappa(1-\kappa)}{2((1+L)\alpha - \kappa)G + 2\kappa(t-2r)} + \frac{\kappa(\kappa(t-5r) + \alpha(t+Lr) + 4r-2t)G + \kappa^2 r(t-3r)}{2G^2(((1+L)\alpha - \kappa)G + \kappa(t-2r))}$$

and

$$s_3^{(2)}(G) = \frac{2\alpha\kappa(L+1) - 2\kappa(\kappa-1) - 2\alpha(L+1)}{2((\alpha(L+1) - \kappa)G - \kappa(2t-r))} + \frac{\kappa((4-5\kappa)t + (\kappa-2)r + \alpha(t+Lr))G + \kappa^2 t(r-3t)}{2G^2((\alpha(L+1) - \kappa)G - \kappa(2t-r))}$$

Then, an approximation of the boundaries based on the second order Taylor expansion is given by $s_{x,-}^{(2)} = s_x^{(2)}(G_{x,-}^{(2)})$ and

$$s_{x,+}^{(2)} = s_x^{(2)}(G_{x,+}^{(2)}).$$

A rougher approximation of the position of the the two bulks is obtained by defining

$$s^{(0)} = -\frac{1}{G} \quad (22)$$

and assuming as boundaries of the eigenvalue pdf support $s_{x,-}^{(0)} = s^{(0)}(G_{x,-}^{(2)})$ and $s_{x,+}^{(0)} = s^{(0)}(G_{x,+}^{(2)})$.

The tightness of the proposed approximation is assessed by numerical simulations. In Figure 1 we show the histogram of the eigenvalues for a finite systems with $T = 3$ transmitted signals per cell, $L = 2$ interfering cells, and $R = 300$ receive antennas at the base station. The signal of interest is received with power $P = -10$ dBW and the interfering signal is received with power $I = \frac{P}{4}$. Additionally, the asymptotic eigenvalue probability density function (pdf) obtained by (10) is drawn by a solid line. The vertical lines indicate the approximation of the boundaries of the asymptotic pdf obtained by perturbation analysis. As expected, the approximation based on $s^{(0)}(G)$ provides a quite rough estimation of the support of the eigenvalue pdf which is strictly contained in the actual support while the approximation based on the second order Taylor expansion provides a much tighter estimation which includes the actual asymptotic support. Thus, the estimation of the gap between the two bulks based on this estimation is conservative and of practical interest.

V. NUMERICAL RESULTS

In this section, results of simulations based on realistic settings with a finite number of receive antennas corroborate the usefulness of the proposed method. In all the simulations we adopt as performance metric the uncoded bit error rate (BER) and compare the proposed algorithm with the linear channel and data estimation proposed in [2].

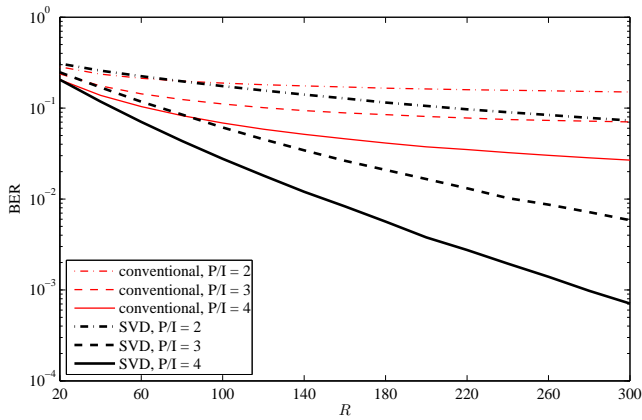


Fig. 2. BER versus number of receive antennas for 1 pilot symbol per transmit antenna and cell. Single-user matched filtering for data detection.

Fig. 2 shows the system improvement in terms of BER when the number of receive antennas increases. In the simulations we set $L = 3$, $T = 5$, $C = 200$, and unit ratio $\frac{P}{W}$, i.e. the SNR is equal to 0dB. The ratio $\frac{P}{T}$ assumes the three different values over the set $\{2, 3, 4\}$. We transmit one pilot symbol per transmit antenna, i.e. 5 pilot symbols in total, and an identical set of orthogonal training sequences is adopted by all the access points as proposed in [2]. For data detection, single user matched filtering is implemented. The proposed algorithm widely outperforms the receiver based on linear channel estimation in [2] (conventional receiver). More interestingly, for any value of the ratio $\frac{P}{T}$ the performance of the conventional receiver flattens due to pilot contamination while the BER of the proposed receiver, for sufficiently large $\frac{P}{T}$ ratios, decreases unboundedly with a slope that, not surprisingly, increases with the ratio¹ $\frac{P}{T}$. Thus, this trend confirms that the pilot contamination problem is overcome, in principle.

Fig. 3 shows the dependence of the BER on the SNR ($\frac{P}{W}$) for $T = \{4, 6, 8\}$, $R = C = 200$, $L = 3$, $\frac{P}{T} = 4$. As in the previous set of simulations, all the cells adopt the same set of orthogonal pilot sequences with one pilot symbol per transmit antenna. As expected, the performance of all systems flattens out at high SNR regime since the thermal noise becomes negligible compared to the interference. The curves for different values of T show the dependence of the performance on α . As pointed out in Section IV the value of α should be below a certain threshold to guarantee separability between the interference and signal-of-interest subspaces.

VI. SUMMARY AND CONCLUSIONS

We proposed a practical algorithm to avoid pilot contamination in cellular systems with random generated pilot sequences, power-controlled handoff and appropriate scheduling. The dominant complexity of this algorithm is a singular value decomposition of received signal block. We conclude, that pilot contamination is not a fundamental effect, but an effect of linear channel estimation. An asymptotic analysis based on free probability and perturbation theory has been proposed.

¹We recall that the ratio $\frac{P}{T}$ determines the separability between the interference and the signal-of-interest subspaces.

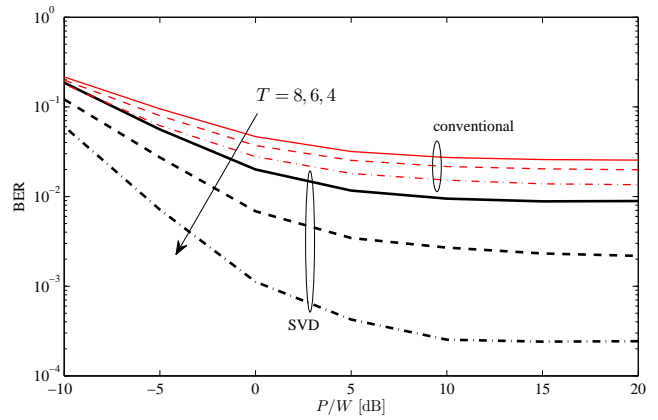


Fig. 3. BER versus SNR for $T = \{4, 6, 8\}$, $R = C = 200$, $L = 3$, $\frac{P}{T} = 4$, one pilot symbol per transmit antenna and cell. Single-user matched filtering for data detection.

ACKNOWLEDGEMENTS

The work of Cottatellucci and Müller was supported by the European research project HARP, which is partly funded by the European Union under its FP7 ICT Programme.

REFERENCES

- [1] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Fortieth Asilomar Conf. on Signals, Systems, & Computers*, Pacific Grove, CA, USA, Oct. 2006.
- [2] —, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *To appear in IEEE Signal Processing Magazine*.
- [4] H. Q. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011.
- [5] B. Gopalakrishnan and N. Jindal, "An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, San Francisco, CA, USA, Jun. 2011.
- [6] F. Fernandez, A. Ashikhmin, and T. Marzetta, "Interference reduction on cellular networks with large antenna arrays," in *Proc. of IEEE International Conference on Communications (ICC)*, Ottawa, Canada, Jun. 2012.
- [7] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *Submitted to IEEE Journal on Selected Areas in Communications*, Mar. 2012, arXiv:1203.5924.
- [8] U. Madhow, "Blind adaptive interference suppression for direct-sequence CDMA," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2049–2069, Oct. 1998.
- [9] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012.
- [10] Christoph Studer and Erik G. Larsson "PAR-Aware Multi-User Precoder for the Large-Scale MIMO-OFDM Downlink", in *Proc. of International Symposium on Wireless Communication Systems (ISWCS)*, Paris, France, Aug. 2012.
- [11] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, "Blind Pilot Decontamination," *Prod. of 17th International ITG Workshop on Smart Antennas (WSA 2013)*, Stuttgart, Germany, Mar. 2013.
- [12] Romain Couillet and Merouane Debbah, "Random matrix methods for wireless communications", *Cambridge University Press*, 2011.