Analysis of Blind Pilot Decontamination

Ralf R. Müller
Universität Erlangen-Nürnberg
Institute for Digital Communications
91058 Erlangen, Germany
Email: mueller@lnt.de

Mikko Vehkaperä
Aalto University
Dept Signal Processing & Acoustics
02150 Espoo, Finland
Email: mikko.vehkapera@aalto.fi

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

***Abstract*—A nonlinear channel estimation scheme proposed by the authors in earlier work is shown to overcome the pilot contamination problem in massive multiple-input multiple-output (MIMO) systems. The method is based on a subspace projection using a singular value decomposition and is studied both by analytical and simulative means. The analysis presented in this paper is a refined version of an earlier work by the same authors. In addition, simulations based on a cellular network model show that the benefits of the proposed scheme are retained also in more realistic settings.**

***Index Terms*—Multiple antennas, multiple-input multiple-output (MIMO) systems, massive MIMO, channel estimation, principal component analysis, random matrix theory.**

I. Introduction

Strongly asymmetric MIMO systems [1] realize a large array gain by a massive use of antenna elements. This design principle commonly referred to as *massive MIMO* has attracted considerable attention recently [2]. Given perfect channel state information, the signals received at all antenna elements can be combined coherently and the array gain grows without bound with the number of antennas at the access point. Therefore, massive use of antennas elements can overcome both multiuser interference and thermal noise for any given number of users and any given powers of the interfering users.

In [3], however, a pessimistic conclusion about the performance of massive MIMO in cellular systems was drawn. Based on the explicit assumption of no coordination among cells and on the implicit assumption of linear channel estimation [3, Eq. (5)], it was concluded that the array gain can be realized only for data detection, but not for channel estimation. The author argued that channel state information, though not required to be perfect, must have at least a certain quality in order to utilize unlimited array gains. As a result, pilot interference from neighboring cells would limit the ability to obtain sufficiently accurate channel estimates and be the new bottleneck of the system. This effect, commonly referred to as *pilot contamination* [4], was believed by many researchers, e.g. [3]–[8] to be a fundamental effect, despite the lack of a solid proof that it cannot be overcome.

Recent works have indicated that pilot contamination is not as fundamental as it was thought to be: Using Bayesian channel estimation, [9] found that pilot contamination can vanish under certain conditions on the channel covariance matrix if some cooperation among cells is allowed. On the other hand, simulation results in [10] showed that the channel can be estimated blindly using non-linear methods with much greater accuracy than with linear methods for a wide range of system parameters. Finally, using random matrix theory (RMT) the authors showed in [11], [12] that pilot contamination is not a fundamental effect, but a shortcoming of linear channel estimation. Using a subspace approach, the array gain can be attained blindly even before channel estimation.

The schemes proposed in [10]–[12] all start with a singular value decomposition of the received signal matrix. Unlike [10], however, the method considered in [11], [12] does not aim to subsequently estimate the channel matrix before performing data detection. Instead, the received signal is proejected onto an (almost) interference-free subspace where communication is governed by a non-linear compound channel that can be estimated easily.

Using RMT, the authors provided in [11], [12] an approximate analysis to determine for which system parameters the subspace of the signal of interest can be identified blindly when a power-controlled hand-off protocol is applied that ensures a power margin between intra-cell and inter-cell users. A compact condition for the existence of blind pilot decontamination was found in [12]. However, this preliminary result lacks accuracy since it relies on several approximations. In particular, it neglects the repulsion of singular values that belong to signals-of-interest from those ones that belong to interference. The eigenvalue repulsion was accounted for later in [11], but the results were less intuitive and more complicated. In this work, we provide a refined version of the original analysis in [12] and take into account for the singular value repulsion. At the same time, we are able to provide intuitive and compact conditions for blind pilot decontamination. Due to space constraints, however, the derivations have been omitted in this paper and they can be found in [13]. We also provide simulation results that use a more realistic cellular model than the one used in [11]–[13].

In Section II, we introduce the system model. In Section III, we propose the 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, Section VI concludes the paper.

II. System Model

Consider a wireless communication channel. In order to ease notation and for sake of conciseness, let the channel bandwidth be smaller than the coherence bandwidth. Channels whose physical bandwidth is wider than the coherence bandwidth can be decomposed into equivalent parallel narrowband channels

[0]This work was supported in part by the European Commission's 7th framework project HARP under grant agreement no. 318489. R. Müller is also with the Norwegian University of Science and Technolgy, Trondheim, Norway.

by means of orthogonal frequency division multiplexing or related techniques.

Let the frequency-flat, block-fading, narrowband channel from T transmit antennas to $R > T$ receive antennas be described by the matrix equation

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{Z}, \tag{1}$$

where $\boldsymbol{X} \in \mathbb{C}^{T\times C}$ is the transmitted data (eventually multiplexed with pilot symbols), $C \geq R$ is the coherence time in multiples of the symbol interval, $\boldsymbol{H} \in \mathbb{C}^{R\times T}$ is the channel matrix of unknown propagation coefficients, $\boldsymbol{Y} \in \mathbb{C}^{R\times C}$ is the received signal, and $\boldsymbol{Z} \in \mathbb{C}^{R\times C}$ is the total impairment. Furthermore, we assume that channel, data, and impairment have zero mean, i.e. $\mathsf{E}\,\boldsymbol{X} = \mathsf{E}\,\boldsymbol{H} = \mathsf{E}\,\boldsymbol{Z} = \boldsymbol{0}$. The impairment includes both thermal noise and interference from other cells and is, in general, neither white nor Gaussian.

III. Proposed Algorithm

Consider the singular value decomposition

$$\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\dagger \tag{2}$$

with unitary matrices $\boldsymbol{U} \in \mathbb{C}^{R\times R}$ and $\boldsymbol{V} \in \mathbb{C}^{C\times C}$ and the $R \times C$ diagonal matrix $\boldsymbol{\Sigma}$ with diagonal entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_R$ sorted in non-increasing order. As shown in [10], the columns of $\boldsymbol{U}$ are highly correlated with the columns of $\boldsymbol{H}$. Based on this observation [10], proposed two algorithms for improved nonlinear estimation of the channel matrix $\boldsymbol{H}$.

In the sequel, we propose a strategy different from the one in [10]. We decompose the matrix of left singular vectors

$$\boldsymbol{U} = [\boldsymbol{S}|\boldsymbol{N}] \tag{3}$$

into the signal space basis $\boldsymbol{S} \in \mathbb{C}^{R\times T}$ and the null space basis $\boldsymbol{N} \in \mathbb{C}^{R\times(R-T)}$. Now, we project the received signal onto the signal subspace and get

$$\tilde{\boldsymbol{Y}} = \boldsymbol{S}^\dagger\boldsymbol{Y}. \tag{4}$$

The null space basis $\boldsymbol{N}$ is not required in the sequel. In fact, there is no need to compute the full singular value decomposition (2). Only the basis of the signal subspace $\boldsymbol{S}$ is needed and there are efficient algorithms available to exclusively calculate $\boldsymbol{S}$.

Consider now the massive MIMO case, i.e. $R \gg T$: 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 $R \to \infty$.

Using the algorithm above, we can achieve an array gain even without the need for estimating the channel coefficients. In fact, channel estimation can be delayed until the received signal has been projected onto the signal subspace and the dominant part of the white noise has already been suppressed.

In order to save complexity it is sensible not to estimate the channel matrix $\boldsymbol{H}$, at all. Instead, we directly consider the subspace channel

$$\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{H}}\boldsymbol{X} + \tilde{\boldsymbol{Z}} \tag{5}$$

and estimate the much smaller subspace channel matrix $\tilde{\boldsymbol{H}} \in \mathbb{C}^{T\times T}$. Although the data dependent projection (4) implies that the noise $\tilde{\boldsymbol{Z}} = \boldsymbol{S}^\dagger\boldsymbol{Z} \in \mathbb{C}^{T\times C}$ is not independent from the data $\boldsymbol{X}$, neglecting this dependence is an admissible approximation that becomes exact as the number of receive antennas R grows large.

In addition to white noise, there is co-channel interference from L neighboring cells. For sake of notational convenience, we assume that the number of transmit antennas is identical in all cells and equal to T. The interference from neighboring cells is anything but white. It is the more colored, the smaller the ratio

$$\alpha = \frac{T}{R} \tag{6}$$

which will be called *load* in the following. Any R-dimensional channel vector is orthogonal to any other channel vector in the limit $R \to \infty$. This holds regardless whether the two channel vectors correspond to transmitters in the same cell or in different cells. In the limit of zero load, i.e. $\alpha \to 0$, we have an even stronger result: the subspace spanned by the co-channel interference is orthogonal to the signal subspace.[1] That means that in the limit $R/T \to \infty$, the $(L+1)T$ largest singular values of the received signal matrix $\boldsymbol{Y}$ become identical to the Euclidean norms of the $(L+1)T$ channel vectors. If we can identify which singular values correspond to channel vectors from inside the cell as opposed to channel vectors from transmitters in neighboring cells, we can remove the interference from neighboring cells by subspace projection.

Note that for $R \to \infty$, the system has infinite diversity and 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 perfect received power control and a 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 in the limit $(R, \alpha) \to (\infty, 0)$, i.e. the number of receive antennas grows large while the number of transmit antennas does not.

For practical systems with small, but nonzero load, i.e. $0 < \alpha \ll 1$, a certain power margin is required between signals of interest and interfering signals. For most interfering users, such a power margin is created for free by shadowing and path loss. However, there might be few users close to cell boundaries who lack such a power margin. As a kind of countermeasure, a power margin has to be engineered for them. There are various ways to do so. In the sequel, we will exemplarily list two such potential methods.

One way to create an additional power margin is a smart choice of frequency or time re-use patterns. However, this requires coordination among cells. Another way to create an additional power margin is to equip each user with at least two transmit antennas. Then, the few users who suffer from insufficient power margin can form beams that favor one of the base stations or access points over others[2]. This will noticeable

[1] Note that the pairwise orthogonality of channel vectors holds for $R \to \infty$, in general, and does not require $\alpha \to 0$. However, the orthogonality of subspaces requires $\alpha \to 0$ in addition to $R \to \infty$, as the accumulation of $T = \alpha R$ vanishing pairwise correlations is not vanishing, in general.

[2] Note that such beam forming does not require channel state information. One can keep on forming random beams until a sufficient power margin is reached.

increase their power margins. The majority of users will not need to employ such methods and can use the two antennas for spatial multiplexing.

IV. Performance Analysis

We have demonstrated above, that the proposed algorithm works in principle in massive MIMO systems as the number of receive antennas grows much larger than the product of transmit antennas and neighboring cells. 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 as in the classical massive MIMO setting. Then, in real systems the asymptotic properties are only approximated. A useful and insightful approach to understand the behavior of a real network consists in assuming that both T and R grow large with a fixed ratio α. This setting can be studied effectively by RMT. In this section, we will adopt results from RMT to answer the question, how large is large enough in practice.

We decompose the impairment process

$$\boldsymbol{Z} = \boldsymbol{W} + \boldsymbol{H}_{\mathrm{I}}\boldsymbol{X}_{\mathrm{I}} \tag{7}$$

into white noise $\boldsymbol{W}$ and interference from L neighboring cells where interfering data $\boldsymbol{X}_{\mathrm{I}} \in \mathbb{C}^{LT\times R}$ is transmitted in neighboring cells and received in the cell of interest through the channel $\boldsymbol{H}_{\mathrm{I}} \in \mathbb{C}^{R\times LT}$. Combining (1) and (7), we get

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{H}_{\mathrm{I}}\boldsymbol{X}_{\mathrm{I}} + \boldsymbol{W}. \tag{8}$$

Let the entries of the data signal $\boldsymbol{X}$ be iid with zero mean and variance P. Let the entries of the channel matrix $\boldsymbol{H}$ be also iid with zero mean, but have unit variance. Let the entries of the matrix of interfering signals $\boldsymbol{X}_{\mathrm{I}}$ be iid with zero mean and variance P and let the entries of the k^{th} column of the matrix of interfering channels $\boldsymbol{H}_{\mathrm{I}}$ be iid with zero mean and variance I_k/P such that the ratio I_k/P accounts for the relative attenuation between out-of-cell user k and the intracell users. Let the empirical distribution of I_k converge to a limit distribution as $LT \to \infty$ which is denoted by $\mathrm{P}_I(\cdot)$. Furthermore, we assume that the elements of the noise $\boldsymbol{W}$ are independent and identically distributed (iid) with zero-mean and variance W. Finally, we define the normalized coherence time

$$\kappa = \frac{C}{R}. \tag{9}$$

Let us denote the asymptotic eigenvalue distribution of $\boldsymbol{Y}\boldsymbol{Y}^\dagger$ as $\mathrm{P}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(x)$. Using a generalization of the results in [11], it can be shown that this asymptotic eigenvalue distribution obeys

$$\begin{aligned} s\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) + 1 = \\ &- \frac{PTC\alpha\left(s\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) + 1 - \kappa\right)\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s)}{\alpha\kappa - PTC\left(s\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) + 1 - \kappa\right)\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s)} \\ &- \int \frac{xLTC\alpha\left(s\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) + 1 - \kappa\right)\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s)\mathrm{dP}_I(x)}{\alpha\kappa - xTC\left(s\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) + 1 - \kappa\right)\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s)} \\ &- \frac{WC\left(s\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) + 1 - \kappa\right)\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s)}{\kappa} \end{aligned} \tag{10}$$

with

$$\mathrm{G}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(s) = \int \frac{\mathrm{dP}_{\boldsymbol{Y}\boldsymbol{Y}^\dagger}(x)}{x - s} \tag{11}$$

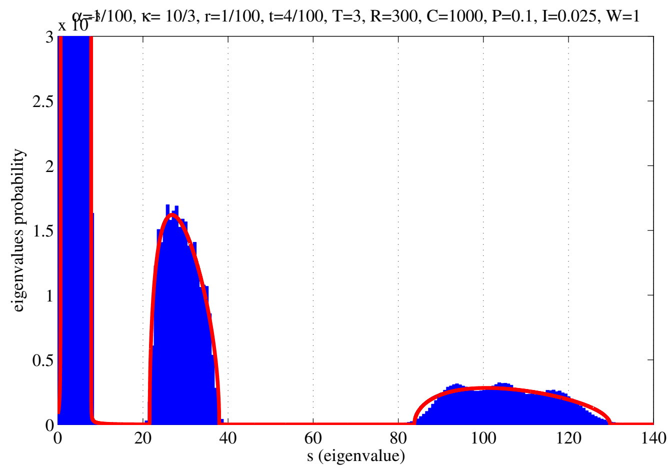

Fig. 1. Asymptotic eigenvalue density of the matrix $\boldsymbol{Y}\boldsymbol{Y}^\dagger/R$ in solid red line for $\alpha = \frac{1}{100}$, $\kappa = \frac{10}{3}$, $P = -10$ dB, $I_k = \frac{P}{4}\,\forall k$, $W = 0$ dB. The asymptotic eigenvalue distribution is compared to the empirical eigenvalue distribution for $T = 3$, $R = 300$, and $C = 1000$ given by the histogram in blue.

denoting its Stieltjes transform. By means of the Stieltjes inversion formula

$$\mathrm{p}(x) = \frac{1}{\pi}\lim_{y\to 0^+} \Im \mathrm{G}(x + \mathrm{j}y) \tag{12}$$

the asymptotic eigenvalue density is obtained.

In Figure 1, the solid line in red shows the asymptotic eigenvalue distribution of $\boldsymbol{Y}\boldsymbol{Y}^\dagger/R$ obtained by (10)-(12). The histogram in blue shows the empirical eigenvalue distribution of $\boldsymbol{Y}\boldsymbol{Y}^\dagger/R$ for $T = 3, R = 300$, and $C = 1000$. We observe that the distribution is decomposed into three disjunct bulks: A noise bulk to the far left, a bulk of the signal of interest to the right, and an interference bulk in between. The fact that the bulks do not overlap enables us to blindly separate the signals of interest from interference and noise as discussed in Section III.

The three bulks are not disjunct in general, but only for certain values of the involved system parameters. It is therefore of utmost importance for practical design of blind pilot decontamination to know which system parameters do lead to bulk separation. The extremely good match between the asymptotic distribution and the empirical distribution for finite matrices corroborate the usefulness to study the support of the asymptotic eigenvalue distribution of $\boldsymbol{Y}\boldsymbol{Y}^\dagger$ and the asymptotic conditions of bulk separability.

The general result for the asymptotic eigenvalue distribution (10) is implicit and not very intuitive. In the following, we develop an approximate analysis for small, but not vanishing loads α. It is based upon the separate calculation of each bulk and subsequent rescaling of the bulks due to pairwise bulk-to-bulk repulsion. We will see that it leads to explicit and intuitive design guidelines.

The co-channel interference is not white but, like the signal of interest, highly concentrated in certain subspaces. The empirical distribution of the squared singular values of the normalized signal of interest, i.e. $\boldsymbol{H}\boldsymbol{X}/\sqrt{TR}$, is shown in [12] to converge, as $R \to \infty$, to a limit distribution which for

$\alpha \ll 1$ is supported in the interval

$$\mathcal{P} = \left[\frac{\kappa P}{\alpha} - 2P\sqrt{\frac{\kappa^2+\kappa}{\alpha}}; \frac{\kappa P}{\alpha} + 2P\sqrt{\frac{\kappa^2+\kappa}{\alpha}} \right]. \tag{13}$$

The empirical distribution of the squared singular values of the normalized co-channel interference, i.e. $\boldsymbol{H}_\text{I}\boldsymbol{X}_\text{I}/\sqrt{TR}$, also converges to a limit distribution. For $\alpha \ll 1$, it is supported in the interval [12]

$$\mathcal{I} = \left[\frac{\kappa I}{\alpha} - 2I\sqrt{L\,\frac{\kappa^2+\kappa}{\alpha}}; \frac{\kappa I}{\alpha} + 2I\sqrt{L\,\frac{\kappa^2+\kappa}{\alpha}} \right] \tag{14}$$

for $I_k = I\,\forall k$. We remark that the condition $I_k = I\,\forall k$ is unrealistic, in practice. However, the general case is not tractable by analytic means. We note, however, that setting all interference powers to the maximum interference power among the users is a worst case scenario and covered by (14).

When separately calculating the eigenvalue spectra of the signal-of-interest, the interference and the noise, the accuracy of the results suffers from the eigenvalues in different bulks repelling each other. In the following, we will correct for this effect up to first order. We decompose one bulk of eigenvalues into single eigenvalues. Then, we introduce correction factors that account for the scaling of one of the single eigenvalues due to the presence of one other bulk of eigenvalues. We will then approximate the influence of several other bulks, e.g. noise bulk and interference bulk, by multiplying the correction factors. This procedure is an approximation, since we neglect the fact that also the scaled bulk of eigenvalues repels the scaling bulk and that the two scaling bulks repel each other.

The presence of additive noise scales the eigenvalues of both the signal of interest and the interference. As shown in [13, Appendix C-A], the scale factors are given for $R \gg T$ by

$$n_\text{P} = \left(1 + \frac{W}{PR}\right)\left(1 + \frac{W}{PC}\right) \tag{15}$$

and

$$n_\text{I} = \left(1 + \frac{W}{IR}\right)\left(1 + \frac{W}{IC}\right), \tag{16}$$

respectively. Note that the two scale factors converge to 1 in the large system limit irrespective of the load α, if the noise power W does not scale with the system size.

The presence of interference scales the eigenvalues of the signal of interest and vice versa. As shown in [13, Appendix C-B], the scale factors for non-overlapping bulks are given for $R \gg T$ by

$$i_\text{P} = \left(1 + \frac{L\alpha/\kappa}{\frac{P}{I} - 1}\right)\left(1 + \frac{L\alpha}{\frac{P}{I} - 1}\right) \tag{17}$$

and

$$i_\text{I} = \left(1 + \frac{\alpha/\kappa}{\frac{I}{P} - 1}\right)\left(1 + \frac{\alpha}{\frac{I}{P} - 1}\right), \tag{18}$$

respectively. Note, however, that these scale factors are only accurate if $P \gg I$. This limits their usefulness in practice.

If the two supporting intervals do not overlap, i.e.

$$n_\text{P} i_\text{P} \mathcal{P} \cap n_\text{I} i_\text{I} \mathcal{I} = \emptyset \tag{19}$$

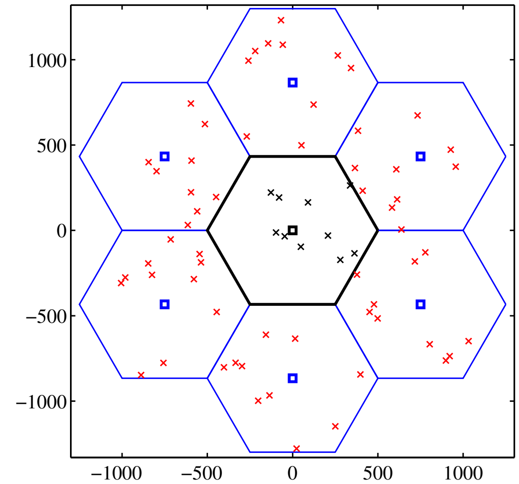

Fig. 2. Hexagonal cellular network considered in the simulations. The bit error rate of the users in the middle cell (depicted in black) is considered when the interfering users (in red) are active. The figure illustrates one specific configuration of user locations in a network with $L = 6$ and $T = 10$.

or equivalently

$$\frac{P}{I} > \frac{n_\text{I} i_\text{I}}{n_\text{P} i_\text{P}} \cdot \frac{1 + 2\sqrt{\alpha L\left(1 + \frac{1}{\kappa}\right)}}{1 - 2\sqrt{\alpha\left(1 + \frac{1}{\kappa}\right)}}, \tag{20}$$

the singular value distribution of the sum of the signal of interest and the interference converges, as $R \to \infty$, to a limit distribution that is composed of two separate non-overlapping bulks [14]. Note that in the limit $\alpha \to 0$, the signal bulk always separates from the interference bulk as long as $P/I > 1$. Therefore, the signal subspace and the interference subspace can be identified blindly. The interference can be nulled out and pilot contamination does *not* happen.

V. Numerical Results

In this section, we provide simulation results for the uncoded bit error rate (BER) and compare the proposed subspace algorithm, with the conventional linear channel and data estimation scheme considered in [3]. Cellular network with hexagonal cells is assumed, as illustrated in Fig. 2 for one specific realization of the user locations. All links are assumed to suffer from path loss, shadowing and fast fading. Long term attenuation is a combination of an exponential path loss model with exponent 4 and log-normal shadow fading with standard deviation 6 dB. Shadowing is assumed to be the same to all R receiving antennas from one specific user but independent among different users. Fast fading follows the Rayleigh distribution and is independent between the receiving antennas. Power control and handover is employed so that the average received power of all users associated with a specific access point is one, while the interference to any other access point is at most unity. We set $P/W = 0.1$ (SNR is -10 dB), that is, assume that the system operates in the low SNR region. Identical set of orthogonal pilot sequences of length T is adopted by all the access points to facilitate channel estimation. The BER is calculated for the users in the middle cell when all six interfering cells have T active users.

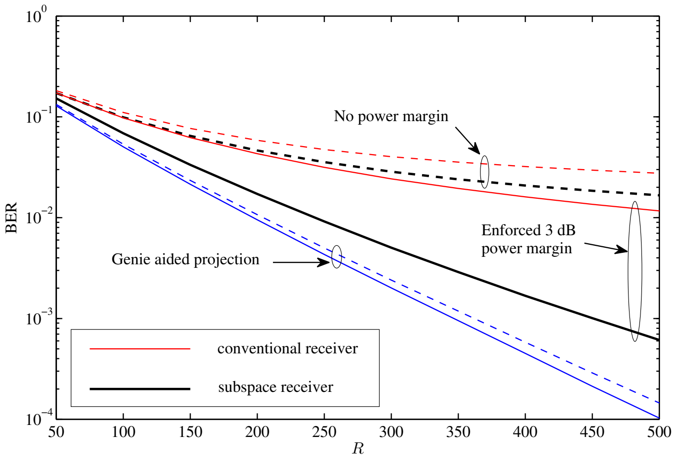

Fig. 3. BER vs. number of receive antennas in a hexagonal cellular network, with $T = 10$, $C = 1000$, $L = 6$, and $P/W = 0.1$ (SNR is -10 dB). Dashed lines for case 1) and solid lines for case 2) setups for interference management (see the main text for details).

To test the performance of the proposed algorithm, we consider the effect of increasing the number of receive antennas while the rest of the parameters are fixed to $T = 10$, $L = 6$, and $C = 1000$. Two interference management scenarios are considered:

1) *No power margin*: Only power-controlled hand-off is employed so that $I_k \leq P$ for all interfering users.
2) *Enforced* 3 *dB power margin*: Additionally, beamforming or scheduling is used so that all interference that would be received with power $I_k > P/2$ is attenuated as $I_k/2$.

Note that the latter case guarantees that all received interference is at least 3 dB below the desired signal. For the current setup, employing the 3 dB power margin required that on average approximately 3% of all users needed to be handled specially. For beamforming this would imply almost a negligible loss in overall spectral efficiency.

As may be observed from Fig. 3, the proposed subspace algorithm outperforms the receiver based on linear channel estimation in [3] (conventional) in both cases. However, naive interference management using only power controlled hand-off is not sufficient to provide a useful performance for either algorithm. If additional interference avoidance is employed so that 3 dB power margin is guaranteed, the subspace method provides a satisfactory performance, while the conventional approach benefits from the reduced interference only marginally. We have also plotted a genie aided projection in Fig. 3 for comparison. In this case, the signal subspace is computed from the true channel $\boldsymbol{H}\boldsymbol{H}^\dagger$ while channel estimation and detection is carried out as before after the projection. While there is a gap between the practical method and genie aided projection, the difference is surprisingly small when 3 dB power margin is guaranteed since the projection is fully blind and based on a mismatched matrix $\boldsymbol{Y}\boldsymbol{Y}^\dagger$.

VI. Summary and Conclusions

We proposed a practical algorithm with polynomial complexity to avoid pilot contamination in cellular systems with power controlled handoff. The dominant complexity of this algorithm is a singular value decomposition of the received signal block. The algorithm was analyzed by means of random matrix theory. The analysis shows that pilot contamination is not a fundamental effect, but is overcome by means of the proposed algorithm.

This paper has focussed solely on the reverse link channel. For the forward link channel, one can exploit channel reciprocity in time-division duplex systems. Similar to the reverse link channel, knowledge of the full channel matrix is not required. Basic considerations of linear algebra show that it is sufficient to know the subspace which the channel vectors of interest span in order to solely require accurate channel estimates for the projected channel (5).

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] 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,” *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[3] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel,” *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[6] 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.

[7] 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.

[8] N. Krishnan, R. D. Yates, and N. B. Mandayam, “Cellular systems with many antennas: Large system analysis under pilot contamination,” in *Proc. of 50th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Oct. 2012.

[9] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[10] 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.

[11] L. Cottatellucci, R. R. Müller, and M. Vehkaperä, “Analysis of pilot decontamination based on power control,” in *Proc. of IEEE Vehicular Technology Conference (VTC)*, Dresden, Germany, Jun. 2013.

[12] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, “Blind pilot decontamination,” in *Proc. of ITG Workshop on Smart Antennas*, Stuttgart, Germany, Mar. 2013.

[13] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, “Blind pilot decontamination,” arXiv:1305.4755 [cs.IT].

[14] R. R. Müller and B. Çakmak, “Channel modelling of MU-MIMO systems by quaternionic free probability,” in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012.

[15] A. Zee, “Law of addition in random matrix theory,” *Nuclear Physics*, pp. 726–744, 1996.