GROUP-BLIND DETECTION WITH VERY LARGE ANTENNA ARRAYS
IN THE PRESENCE OF PILOT CONTAMINATION

G. C. Ferrante°*, G. Geraci°, T. Q. S. Quek°, and M. Z. Win°

°SUTD, Singapore, and *MIT, MA

ABSTRACT

Massive MIMO is, in general, severely affected by pilot contamination. As opposed to traditional detectors, we propose a group-blind detector that takes into account the presence of pilot contamination. While sticking to the traditional structure of the training phase, where orthogonal pilot sequences are reused, we use the excess antennas at each base station to partially remove interference during the uplink data transmission phase. We analytically derive the asymptotic SINR achievable with group-blind detection, and confirm our findings by simulations. We show, in particular, that in an interference-limited scenario with one dominant interfering cell, the SINR can be doubled compared to non-group-blind detection.

Index Terms— Group-blind detection, pilot contamination, massive MIMO, interference suppression.

1. INTRODUCTION

Massive MIMO is, along with network densification and increased bandwidth, one of the key technologies promising to tremendously increase the rate per area in future cellular networks [1, 2]. The idea behind massive MIMO is to equip base stations (BSs) with a number of antennas much larger than the number of active users per time-frequency resource block [3, 4]. This allows to increase the uplink SNR through maximal-ratio combining, and make the matched-filter detector optimal with perfect channel state information.

However, the modern cellular network architecture foresees the use of orthogonal pilots to estimate the channel between each user and the BS. The number of orthogonal pilots is upper bounded by the coherence time [5, 6], hence pilots are usually reused in different cells. As a consequence, channel estimation is corrupted by the interference received during the training phase from users in other cells, a phenomenon known as pilot contamination [7, 8, 9]. As the number of antennas at each BS grows, the rate achievable by traditional receivers saturates due to pilot contamination [10, 11, 12, 13]. In order to mitigate pilot contamination through an improved channel estimation, a nonlinear iterative algorithm that jointly estimates channels and transmitted symbols has been proposed in [14]. A step toward the understanding of the fundamental limits of massive MIMO has been recently made by Müller et al. [15, 16, 17], showing that pilot contamination can be removed if the power received from users within the cell is larger than that received from out-of-cell users. However, this assumption requires both power control and a regular cell geometry, and may not hold in a dense cellular network [18].

In this paper, we propose a receiver design for the uplink of massive MIMO multiuser cellular networks. The proposed receiver takes into account pilot contamination, by adapting the group-blind detection scheme originally proposed in [19] to the case of imperfect channel knowledge due to contamination. Unlike [15, 16, 17], we do not assume neither power control nor regular cell tessellation. Moreover, this paper differs from [14] as it focuses on improving data detection rather than modifying the channel estimation phase. We provide an asymptotic analysis of the SINR achievable by our scheme, showing a gain with respect to non-group-blind receivers. In particular, we show an SINR gain equal to two in an interference-limited scenario with one dominant interfering cell. Simulations validate our analysis and confirm the rate improvement attained by our scheme.

2. SYSTEM MODEL

2.1. Received signal

Consider the uplink of a noncooperative multicellular network with $L$ cells. Each cell is equipped with one BS having $n$ antennas, each BS serving $K$ single-antenna users on the same time-frequency signaling resource. Throughout the paper, the reference cell is referred to as cell 1, and interfering cells are labelled with indices $l \in \{2, \ldots, L\}$. Users in the reference cell and in other cells will be referred to as in-cell users and out-of-cell users, respectively. The signal received by the reference BS during symbol period $m$ is:

$$
y(m) = \sum_{l=1}^{L} \sum_{k=1}^{K} h_{lk} \beta_{lk} x_{lk}(m) + n(m), \tag{1}
$$

where: $h_{lk} = [h_{l1k}, \ldots, h_{lkn}]^\top \in \mathbb{C}^n$ is the channel vector between user $k$ in cell $l$, and the reference BS, being $h_{lk}$, the channel coefficient with respect to antenna $r$; $\beta_{lk} > 0$ is the channel gain between the reference BS
and user $k$ in cell $l$, that models pathloss and shadowing effects; $x_{ik}(m)$ is the symbol transmitted by user $k$ in cell $l$; $n(m) \in \mathbb{C}^n$ is additive white Gaussian noise (AWGN) vector. We assume $\{\beta_{kl}\}$ fixed during the coherence time, $h_{lk} \sim \mathcal{CN}(0, I), \mathbb{E}[n(m)n(m)' ] = \delta_{mm'}, \mathbb{E}[h_{lk}h_{lk}'] = I \delta_{kk}, \mathbb{E}[x_{ik}x_{ik}'] = P \delta_{kk}$, where $P$ is the transmitted power, assumed equal for all users. Denote $H_l \triangleq [h_{l1}, \ldots, h_{lK}] \in \mathbb{C}^{n \times K}$, $R_l = \text{diag}(\beta_{1l}, \ldots, \beta_{Kl})$, and $G_l = H_l R_l^{1/2} = [g_{l1}, \ldots, g_{lK}] \in \mathbb{C}^{n \times K}$. Compactely, eq. (1) can be written as follows for the generic symbol period $m$:

$$y = \sum_{l=1}^{L} H_l R_l^{1/2} x_l + n = \sum_{l=1}^{L} G_l x_l + n,$$

(2)

where dependence on $m$ is implicit.

2.2. Channel estimation

By using orthogonal pilots during the training phase, the MMSE estimation $\hat{g}_{1k}$ of $g_{1k}$ is [20, 10]

$$\hat{g}_{1k} = \left( \sum_{l \geq 1} g_{lk} + \sqrt{\nu}_{lk} \right) \varphi_{1k} \beta_{1k}^{-1},$$

(3)

where $1/\epsilon$ is equal to the effective training SNR, $\nu_{lk} \sim \mathcal{CN}(0, I)$, and

$$\varphi_{1k} = \frac{\beta_{1k}^2}{\epsilon + \sum_{l \geq 1} \beta_{lk}}.$$  \hspace{1cm} (4)

We collect in matrix form estimations $\hat{G}_l = [\hat{g}_{l1}, \ldots, \hat{g}_{lK}]$ and estimation errors $\hat{G}_l = G_l - \hat{G}_l$. 

2.3. Achievable rate

Following [5, 10], an achievable rate $R_{1k}$ for in-cell user $k$ can be derived by considering the signal

$$y' = \hat{G}_l x_l + \hat{G}_l \tilde{x}_l + \sum_{l \geq 1} G_l x_l + n,$$

(5)

where $\tilde{x}_l$ is independent on $x_l$ and has same covariance. Let $\hat{w}_{1k}$ denote the linear receiver for user $k$, the rate $R_{1k}$ is given by

$$R_{1k} = \mathbb{E}[\log(1 + \gamma_{1k})],$$

(6)

where the expectation is with respect to estimated channels, and the SINR $\gamma_{1k}$ is given in (7) at the top of the next page.

3. PROPOSED GROUP-BLIND DETECTOR

Blind receivers were developed for multiuser detection and equalization [21], and then generalized to group-blind detection in the context of CDMA [19, 22, 23, 24]. While in blind techniques the receiver knows the signature sequence of the user to decode only, in group-blind techniques it knows the signature sequences of a subset of users. In the uplink of a cellular network, this corresponds to a BS knowing in-cell channels and being unaware of out-of-cell channels. While group-blind detection was originally proposed assuming perfect knowledge of a subset of channels, we hereby extend group-blind detection to the case of contaminated knowledge.

The proposed receiver $\hat{w}_{1k}$ consists of two components. A first component, $\hat{w}_{1k}$, belongs to range $\hat{G}_l$ and is derived on the basis of the signal $y'' = \hat{G}_l x_l + n$; the MMSE criterion yields

$$\hat{w}_{1k} = (\hat{G}_l G_l^{1/2} + \frac{1}{n} I)^{-1} \hat{g}_{1k}.$$  \hspace{1cm} (8)

A second component, $\tilde{w}_{1k}$, belongs to a subspace orthogonal to range $\hat{G}_l$ and that lies within the signal space. Let $\hat{U}_{G_l}$ be a matrix whose columns span such subspace. The component $\tilde{w}_{1k}$ is derived by taking into account the whole received signal. Following the MMSE criterion, a derivation similar to [19] that also accounts for imperfect channel estimation due to pilot contamination yields

$$\tilde{w}_{1k} = - \hat{U}_{G_l} \left( \hat{U}_{G_l}^* C_{y'} \hat{U}_{G_l} \right)^{-1} \hat{U}_{G_l}^* C_{y'} \hat{w}_{1k},$$

(9)

where $C_{y'}$ is the covariance matrix of (5). The group-blind detector $w_{1k} = \hat{w}_{1k} + \tilde{w}_{1k}$ is, therefore, explicitly given by

$$w_{1k} = \left( I - \hat{U}_{G_l} \left( \hat{U}_{G_l}^* C_{y'} \hat{U}_{G_l} \right)^{-1} \hat{U}_{G_l}^* C_{y'} \right) \hat{w}_{1k}.$$  \hspace{1cm} (10)

We can show that simple blanking techniques [2] allow to accurately estimate $C_{y'}$ for the purpose of implementing (10). Details are omitted due to space constraints.

4. ASYMPTOTIC PERFORMANCE ANALYSIS

We derive the asymptotic achievable rate, as $n \rightarrow \infty$. In this limit, the SINR is bounded by non-vanishing interference terms. In order to obtain these asymptotically non-vanishing terms, we use the two following properties [7, 10]:

(i) channels are asymptotically orthogonal in the almost sure sense, i.e., $n^{-1} g_{ik}^* g_{kj} \xrightarrow{a.s.} \beta_{kl} \delta_{kk} \delta_{ij}$, where $\delta_{ij}$ denotes the Kronecker delta;

(ii) in the high-SNR regime, $\hat{g}_{1k} \in \mathcal{S}_k = \text{range} \{g_{lk} : l \geq 1\}$.

A conceptual representation of the structure of the signal space in the high-SNR regime as $n \rightarrow \infty$ is shown on Fig. 1. Properties (i) and (ii) imply that, asymptotically, the signal space $\mathcal{S}_k = \text{range} \{G_l : l \geq 1\}$ is the direct sum $\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_K$. However, the projection of $\hat{g}_{1k}$ onto vectors in $\mathcal{S}_k$ does not asymptotically vanish, in general. In fact, the following limit holds:

$$\frac{1}{n} g_{1k}^* g_{ij} \xrightarrow{a.s.} \varphi_{1k} \beta_{1k}^{-1} \sum_{m \geq 1} g_{mk}^* \beta_{1k}^{-1} \beta_{lj} \delta_{kj},$$

(11)
Theorem 1. Let
\[ \gamma_{1k} = \frac{\mathbb{E} \left[ \left| w_{1k}^\dagger \hat{g}_{1k} \right|^2 \right]}{\mathbb{E} \left[ \left| w_{1k}^\dagger \left( I + \sum_{j \neq k} g_{1j} g_{1j}^\dagger + \sum_{l > 1} \sum_{j > 1} g_{lj} g_{lj}^\dagger \right) w_{1k} \right| \right]}. \] (7)

Let \( \bar{\gamma}_{1k} \) denote the SINR achieved with traditional (non-group-blind) detection, i.e., when \( \hat{w}_{1k} = 0 \), given by [7, 10]
\[ \bar{\gamma}_{1k} = \frac{\beta_{2k}^2}{\sum_{l > 1} \beta_{lk}^2}. \] (14)

We define the asymptotic SINR gain \( \bar{\eta}_{1k} \) provided by the proposed group-blind detector as
\[ \bar{\eta}_{1k} = \frac{\bar{\gamma}_{1k}}{\bar{\gamma}_{1k}}. \] (15)

For \( L = 2 \), (14) reduces to \( \bar{\gamma}_{1k} = \beta_{2k}^2 \beta_{2k}^{-2} \), which combined with (13) and (15) yields
\[ \bar{\eta}_{1k} = 1 + \frac{1}{(1 + \epsilon/\beta_{2k})^2}. \] (16)

Both \( \bar{\gamma}_{1k} \) and \( \bar{\eta}_{1k} \) simplify in the limit \( \epsilon \to 0 \), as specified in the following Corollary.

Corollary 1. Let \( L = 2 \). The SINR gain obtained by using the group-blind receiver for user \( k \) satisfies
\[ \bar{\eta}_{1k} \to 2 \quad \text{as} \quad \epsilon \to 0, \] (17)

hence the asymptotic SINR \( \gamma_{1k} \) reduces to
\[ \bar{\gamma}_{1k} \to 2 \bar{\gamma}_{1k} \quad \text{as} \quad \epsilon \to 0. \] (18)

The above Corollary shows that, in the high-SNR regime, the asymptotic SINR achieved with group-blind detection is doubled compared to traditional detection.

Let \( \Delta R_{1k} \) be the difference between the asymptotic rates achieved by user \( k \) with and without group-blind detection, respectively, given by
\[ \Delta R_{1k} = \log(1 + \bar{\gamma}_{1k}) - \log(1 + \bar{\gamma}_{1k}). \] (19)

In the high-SNR regime, \( \Delta R_{1k} \approx \rho_{1k}^2 \) when \( \beta_{2k} \gg \beta_{1k} \) (strong out-of-cell interference), while \( \Delta R_{1k} \approx 1 \) b/s/Hz when \( \beta_{2k} \ll \beta_{1k} \) (weak out-of-cell interference). Note that the case \( \beta_{2k} \gg \beta_{1k} \) can occasionally occur when BSSs are randomly deployed, resulting in irregular Voronoi cells.

5. NUMERICAL RESULTS

In this section, we give numerical results to validate our analysis and show the performance gain achieved through group-blind detection. We assume \( \beta_{11} = \cdots = \beta_{1K} = 1 \) and denote \( \text{SNR} = P \beta_{1k} = P \) the average received SNR from in-cell users. Solid lines on figures correspond to analytical results,

As a consequence, in-cell user \( k \) is asymptotically interfered only by out-of-cell users who used the same training sequence. On the basis of the above observations, we can derive the following results, that hold for the case of \( L = 2 \) cells. We note that such a scenario approximates an interference-limited network with one dominant interfering cell. Proofs as well as analytical results for \( L \geq 2 \) cells are omitted and will be provided in a journal version of this manuscript. Numerical results that validate our analysis are given in § 5.

Lemma 1. Let \( L = 2 \). Asymptotically, the variable after detection satisfies
\[ \frac{1}{n} w_{1k}^\dagger y' \xrightarrow{a.s.} \varphi_{1k} x_{1k} + \left( (\varphi_{1k} \beta_{1k}^{-1} \beta_{2k}) - \frac{(\varphi_{1k} \beta_{1k}^{-1} \beta_{2k})^3}{\lambda} \right) x_{2k} + \lambda^{-1} (\varphi_{1k} \beta_{1k}^{-1} \beta_{2k})^2 (\beta_{1k} - \varphi_{1k}) \tilde{x}_{1k} \] (12)

where \( \lambda = (\beta_{1k} - \varphi_{1k})^2 + (\varphi_{1k} \beta_{1k}^{-1} \beta_{2k})^2 \).

The SINR achieved with group-blind detection readily follows from Lemma 1.

Theorem 1. Let \( L = 2 \). The SINR \( \gamma_{1k} \) achieved with the proposed group-blind detector satisfies:
\[ \gamma_{1k} \xrightarrow{a.s.} \bar{\gamma}_{1k} = \left[ 1 + \frac{1}{(1 + \epsilon/\beta_{2k})^2} \right] \beta_{1k}^{-1} \] (13)

where \( \rho_{1k} = \beta_{1k} \beta_{2k}^{-1} \).

Fig. 1: Conceptual representation of the asymptotic structure of the signal space in the high-SNR regime. All channels are almost surely orthogonal. For any fixed in-cell user \( k \), both estimated and error vectors belong to \( \mathcal{S}_k = \text{range}\{g_{ik} : l \geq 1\} \).
results, whereas dashed lines connect simulation outputs. All figures confirm that simulations converge to our closed form expressions as the number of antennas grows.

We consider a simple scenario in Fig. 2, with \( L = 2, K = 1, \) \( \text{SNR} = 20 \text{ dB} \) and \( \beta_{11}/\beta_{21} = 0 \text{ dB} \) (strong interference). We compare the achievable rate of non-group-blind (NGB) detection vs. group-blind (GB) detection. GB detection outperforms NGB detection by \( \Delta R_{1k} \approx 0.585 \) (b/s/Hz). This value matches the asymptotic rate gap following from (19), (18), and (14), that is \( \Delta R_{1k} = \log_2(3) - \log_2(2) \) (b/s/Hz).

Figure 3 shows the achievable rate vs. the number of antennas \( n \) for GB detection. We consider several values of \( \epsilon/\beta_{2k} \), that model whether the estimation error is dominated by pilot contamination (\( \epsilon < \beta_{2k} \)) or thermal noise (\( \epsilon > \beta_{2k} \)). Figure 3 is based on the following scenario: \( L = 2, K = 1, \) \( \text{SNR} = 10 \text{ dB} \), and fixed \( \beta_{21} \) with \( \beta_{11}/\beta_{21} = 10 \text{ dB} \) (weak interference). The rate achieved with GB detector in the presence of negligible noise during the training phase is plotted for comparison. The rate achieved with GB detectors decreases as \( \epsilon \) grows, consistently with (13). However, even when the training phase is severely affected by noise, GB detection still outperforms NGB detection with noise-free training phase.

Finally, we consider in Fig. 4 a scenario with \( L = 4 \) cells, \( \text{SNR} = 10 \text{ dB} \), and a number of users per cell equal to either \( K = 1 \) or \( K = 10 \), and plot the achievable rate per user, as a function of the number of antennas \( n \). GB detection is showed for \( n \geq KL \), that is the required minimum number of antennas to implement the detector. In this case also, GB detection outperforms NGB detection. Moreover, the figure shows that GB detection is much more robust to variations of the network load, i.e., the number of users per BS antenna.

6. DISCUSSION

We introduced a receiver for the uplink of multiuser massive MIMO that accounts for the presence of pilot contamination. The proposed scheme performs group-blind detection by exploiting the excess degrees of freedom provided by the large number of antennas per BS. We derived analytical results for the asymptotic achievable rate in an interference-limited scenario with one dominant interfering cell, and confirmed our findings through simulations. We found that group-blind detection outperforms traditional detection irrespective of the noise impairment during the training phase, and it is much more robust to variations of the network load.
7. REFERENCES

[1] F. Boccardi, R.W. Heath, A Lozano, T.L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.

[2] J.G. Andrews, S. Buzzi, W. Choi, S.V. Hanly, A. Lozano, A.C.K. Soong, and J.C. Zhang, “What will 5G be?,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[3] L. Lu, G.Y. Li, A.L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 40–60, 2013.

[4] 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 Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2013.

[5] B. Hassibi and B.M. Hochwald, “How much training is needed in multiple-antenna wireless links?,” *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[6] N. Jindal and A. Lozano, “A unified treatment of optimum pilot overhead in multipath fading channels,” *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2939–2948, 2010.

[7] T.L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.

[8] 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 Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, 2013.

[9] J. Jose, A. Ashikhmin, T.L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multicell TDD systems,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.

[10] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013.

[11] B. Gopalakrishnan and N. Jindal, “An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas,” in *Proc. IEEE Int. Workshop Signal Process. Adv. in Wireless Commun.* (SPAWC), 2011, pp. 381–385.

[12] F. Fernandes, A. Ashikhmin, and T.L. Marzetta, “Inter-cell interference in noncooperative TDD large scale antenna systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192–201, 2013.

[13] N. Krishnan, R.D. Yates, and N.B. Mandayam, “Cellular systems with many antennas: Large system analysis under pilot contamination,” in *Proc. 50th Annu. Allerton Conf. Commun. Control Comput.*, 2012, pp. 1220–1224.

[14] H.Q. Ngo and E.G. Larsson, “EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2012, pp. 3249–3252.

[15] R.R. Muller, L. Cottatellucci, and M. Vehkapera, “Blind pilot decontamination,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, 2014.

[16] R.R. Muller, M. Vehkapera, and L. Cottatellucci, “Analysis of blind pilot decontamination,” in *Proc. Asilomar Conf. on Signals, Systems and Computers (ACSSC)*, 2013, pp. 1016–1020.

[17] L. Cottatellucci, R.R. Muller, and M. Vehkapera, “Analysis of pilot decontamination based on power control,” in *Proc. IEEE Veh. Technol. Conf. (VTC)*, 2013, pp. 1–5.

[18] J.G. Andrews, H. Claussen, M. Doehler, S. Rangan, and M.C. Reed, “Femtocells: Past, present, and future,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, 2012.

[19] X. Wang and A. Host-Madsen, “Group-blind multiuser detection for uplink CDMA,” *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1971–1984, 1999.

[20] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[21] U. Madhow, “Blind adaptive interference suppression for direct-sequence CDMA,” *Proc. IEEE*, vol. 86, no. 10, pp. 2049–2069, 1998.

[22] J. Zhang and X. Wang, “Large-system performance analysis of blind and group-blind multiuser receivers,” *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2507–2523, 2002.

[23] A. Host-Madsen and X. Wang, “Performance of blind and group-blind multiuser detectors,” *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1849–1872, Jul 2002.

[24] Z. Xu and X. Wang, “Large-sample performance of blind and group-blind multiuser detectors: A perturbation perspective,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2389–2401, 2004.