Massive MIMO for Crowd Scenarios: A Solution Based on Random Access

Jesper H. Sørensen, Elisabeth de Carvalho and Petar Popovski
Aalborg University, Department of Electronic Systems, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark
E-mail: {jhs,edc,petarp}@es.aau.dk

***Abstract*—This paper presents a new approach to intra-cell pilot contamination in crowded massive MIMO scenarios. The approach relies on two essential properties of a massive MIMO system, namely near-orthogonality between user channels and near-stability of channel powers. Signal processing techniques that take advantage of these properties allow us to view a set of contaminated pilot signals as a graph code on which iterative belief propagation can be performed. This makes it possible to decontaminate pilot signals and increase the throughput of the system. The proposed solution exhibits high performance with large improvements over the conventional method. The improvements come at the price of an increased error rate, although this effect is shown to decrease significantly for increasing number of antennas at the base station.**

I. Introduction

Multiple-input multiple-output (MIMO) has been identified as a key technology to improve spectral efficiency of wireless communication systems and is finding its way into practical systems, like LTE and its successor LTE-Advanced. The research in MIMO has recently taken a new turn, when the advantage of having a massive number of antennas at a base station (BS) was asserted in [1]. As defined in [1], a massive MIMO system refers to a multi-cell multi-user system with a massive number of antennas at the BS that serves multiple users. The number of users is much smaller than the number of BS antennas, defining an under-determined multi-user system with a massive number of extra spatial degrees of freedom (DoF). Exploiting those extra DoF and assuming an infinite number of antennas at the BS, the multi-user MIMO channel can be turned into an orthogonal channel and the effect of small-scale fading and thermal noise can be eliminated. Based on those excellent properties, massive MIMO is acknowledged as a promising technology for very high system throughput and energy efficiency [2].

When the number of antennas becomes massive, acquiring the channel state information (CSI) becomes a severe bottleneck. Downlink channel training requires a training length that is proportional to the number of antennas at the BS and is thus impractical. One solution promoted in [1] restricts massive MIMO operations to time-division duplex (TDD) for which channel reciprocity is exploited. As the downlink and uplink channels are equal, CSI is acquired at the BS based on uplink training and then used for DL transmission. The benefit is that the training length is proportional to the number of users, which is much smaller than the number of BS antennas.

As described in [1], CSI is acquired using orthogonal training sequences, but, due to the shortage of orthogonal sequences, the same pilot sequences must be reused in neighboring cells, causing pilot contamination. The problem of pilot contamination is considered as one of the major challenges in massive MIMO systems [3]. Mitigation of pilot contamination has been the focus of several works recently. These include [4], where it is utilized that the desired and interfering signals can be distinguished in the channel covariance matrices, as long as the angle-of-arrival spreads of desired and interfering signals do not overlap. A pilot sequence coordination scheme is proposed to help satisfying this condition. The work in [5] utilizes coordination among base stations to share downlink messages. Each BS then performs linear combinations of messages intended for users applying the same pilot sequence. This is shown to eliminate interference when the number of base station antennas goes to infinity. A multi-cell precoding technique is used in [6] with the objective of not only minimizing the mean squared error of the signals of interest within the cell, but also minimizing the interference imposed to other cells.

The survey of the related work indicates that the pilot contamination problem has been seen as an inter-cell problem that arises when the users associated with two neighboring cells use the same pilot sequence. An implicit assumption associated with it is that the pilot sequences of the users associated with the same cell are perfectly scheduled, such that no intra-cell pilot contamination occurs. These assumptions fall apart when one considers very dense, crowded scenarios as envisioned in 5G wireless scenarios [7]. In such a setting, orthogonal scheduling of the users belonging to the same BS becomes infeasible. This is the scenario considered in this paper, where each user needs to select a pilot sequence from a small pool of pilot sequences shared by all the users. Since the users are not coordinated, the pilot contamination problem can be cast as a specific form of a *random access* problem. We identify two features specific to massive MIMO: (1) asymptotic orthogonality between any pair of users; and (2) asymptotic invariance of the power received from a user over a short time interval. We use these two features in order to formulate a *pilot access protocol* using the framework of *coded random access* [8, 9] and boost the performance by applying successive interference cancellation (SIC). The

The research presented in this paper was supported by the Danish Council for Independent Research (Det Frie Forskningsråd) DFF - 1335 – 00273

difference from the existing approaches for coded random access is that the proposed protocol combines decoding of the data in the uplink with estimation of the channel, which can be used for downlink transmission. Overall, the solution proposed in this paper is a radical departure from the usual treatment of the pilot contamination problem and introduces an important link to the area of random access protocols.

II. System Model

In this work we denote scalars in lower case, vectors in bold lower case and matrices in bold upper case. A superscript "T" denotes the transpose and a superscript "H" denotes the conjugate transpose.

We consider a random access system consisting of a single base station with M antennas and K users with a single antenna, where M and K are in the hundreds or thousands. Communication is performed on a time slotted basis, where each time slot consists of four phases; an uplink pilot phase, an uplink data phase, a downlink pilot phase and a downlink data phase, see Fig. 1. The channel between the k'th user and the base station in the n'th time slot is denoted $\boldsymbol{h}_{nk} = [h_{nk}(1)\ h_{nk}(2)\dots h_{nk}(M)]^T$, where $h_{nk}(i) \sim \mathcal{CN}(0,1)\ \forall\ i$. It is assumed that $\boldsymbol{h}_{nk}\ \forall\ k$ are mutually orthogonal, which is justified by the range of M. Moreover, it is assumed that channel coefficients in different time slots are i.i.d, while the channel power, $\boldsymbol{h}_{nk}^H\boldsymbol{h}_{nk} = ||\boldsymbol{h}_{nk}||^2$ remains constant within a period of β time slots. Note that the channel power varies due to path loss and shadowing effects, which causes it to vary much slower than the channel coefficients.

In each time slot, each user is active with probability p_a. If a user is active, a random pilot sequence, $\boldsymbol{s}_k = [s_k(1)\ s_k(2)\dots s_k(\tau)]^T$, is chosen among a set of size τ with mutually orthogonal pilot sequences. Note that multiple users may choose the same pilot sequence. See Fig. 2 for an example of a random pilot schedule with $\tau = 2$ and $K = 3$. By $\mathcal{A}_n$, we denote all active users in time slot n and by $\mathcal{A}_n^j$, we denote the set of users applying $\boldsymbol{s}_j$ in the n'th time slot. If $\boldsymbol{Y}_{nj}^{pu}$ denotes the uplink pilot signal received in time slot n from users applying $\boldsymbol{s}_j$, we have

$$\boldsymbol{Y}_{nj}^{pu} = \sum_{k\in\mathcal{A}_n^j} \boldsymbol{h}_{nk}\boldsymbol{s}_j^T + \boldsymbol{z}_{nj}^{pu}, \tag{1}$$

where $\boldsymbol{z}_{nj}^{pu}$ is a vector of i.i.d. Gaussian noise components, hence $\boldsymbol{z}_n(i) \sim \mathcal{CN}(0,\sigma_n^2)\ \forall\ i$. Any future instances of a vector $\boldsymbol{z}$ with different sub- or superscripts follow the same definition. All active users transmit a message of length L in the uplink data phase. The message from the k'th user is denoted $\boldsymbol{x}_k^u = [x_k^u(1)\ x_k^u(2)\dots x_k^u(L)]$. Denoting the received uplink signal in time slot n as $\boldsymbol{Y}_n^u$, we then have

$$\boldsymbol{Y}_n^u = \sum_{k\in\mathcal{A}_n} \boldsymbol{h}_{nk}\boldsymbol{x}_k^u + \boldsymbol{z}_n^u. \tag{2}$$

In the downlink phase we rely on channel reciprocity, such that the uplink channel estimate is assumed to be a valid estimate of the downlink channel. The base station transmits a precoded downlink pilot sequence, such that the k'th user receives a downlink pilot signal, $\boldsymbol{y}_{nk}^{pd}$, given by

$$\boldsymbol{y}_{nk}^{pd} = \boldsymbol{h}_{nk}^T\boldsymbol{w}_{nk}\boldsymbol{s}_j^T + \boldsymbol{z}_{nk}^{pd}, \tag{3}$$

where $\boldsymbol{w}_{nk} = [w_{nk}(1)\ w_{nk}(2)\dots w_{nk}(M)]^T$ is the precoding vector for user k in the n'th time slot. The base station is able to schedule the downlink messages, $\boldsymbol{x}_k^d = [x_k^d(1)\ x_k^d(2)\dots x_k^d(L)]$, such that the received signal in the downlink data phase is

$$\boldsymbol{y}_{nk}^d = \boldsymbol{h}_{nk}^T\boldsymbol{w}_{nk}\boldsymbol{x}_k^d + \boldsymbol{z}_{nk}^d. \tag{4}$$

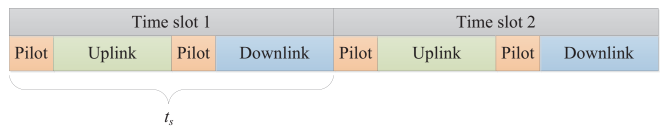

Fig. 1. An example of a transmission schedule.

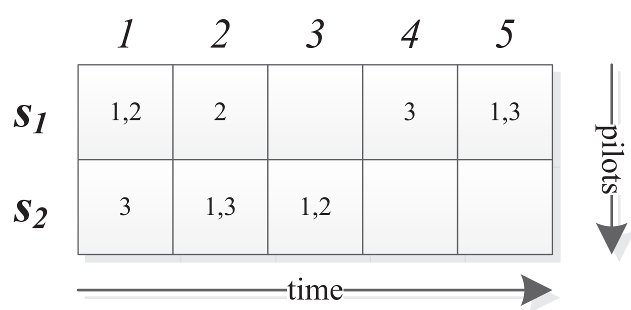

Fig. 2. An example of a pilot schedule.

III. Coded Pilot and Data Transmissions

This section describes the proposed method of communication in the system described in section II. The main focus of this work is the uplink phase, however, a subsection is dedicated to describing the operation in the downlink phase.

A. Uplink

Based on the uplink pilot signals from (1), it is possible to estimate the channels between the users and the base station. However, since multiple user may apply the same pilot sequence, it is only possible to estimate a sum of the involved channels. The least squares estimate, $\boldsymbol{\phi}_{nj}$, based on the pilot signal received in time slot n from users applying $\boldsymbol{s}_j$ is found as

$$\begin{aligned}\boldsymbol{\phi}_{nj} &= ((\boldsymbol{s}_j^H\boldsymbol{s}_j)^{-1}\boldsymbol{s}_j^H\boldsymbol{Y}_{nj}^{pu\,T})^T \\ &= \sum_{k\in\mathcal{A}_n^j} \boldsymbol{h}_{nk} + \boldsymbol{z}_{nj}^{pu'}.\end{aligned} \tag{5}$$

where $z_{nj}^{pu'}$ is the impairment of the estimate caused by the noise, z_{nj}^{pd}. Any future instances of a vector z with a prime follow the same definition.

The problem of interfering users applying the same, or a non-orthogonal, pilot sequence is often called *pilot contamination*. If we proceed to detect the data in the uplink phase using a contaminated channel estimate, the result will be a summation of data messages. If ψ_{nj} is the data estimate based on the channel estimate ϕ_{nj}, we have

$$\begin{aligned} \psi_{nj} &= ((\phi_{nj}^H \phi_{nj})^{-1} \phi_{nj}^H Y_n^u)^T \\ &= \sum_{k \in \mathcal{A}_n^j} \frac{||h_{nk}||^2}{||h_{nj}||^2} x_k^u + z_n^{u'}. \end{aligned} \tag{6}$$

Hence, a pilot collision leads to a data collision. In our system, one way to deal with this problem is to carefully select p_a, such that the probability of having one and only one user applying a particular pilot sequence in a particular time slot is maximized. Hence, we have

$$\begin{aligned} &\underset{p_a}{\text{maximize}} && \Pr(|\mathcal{A}_n^j| = 1) \\ &\text{subject to} && 0 \leq p_a \leq 1 \end{aligned} \tag{7}$$

This will maximize the number of non-contaminated channel estimates, and in turn maximize the number of successful uplink data transmissions. This approach is reminiscent of the framed slotted ALOHA protocol for conventional random access. We consider this a reference scheme in this work and refer to it simply as ALOHA.

ALOHA has been state-of-the-art for many years within random access protocols, but recently a paradigm shift has started with the advent of coded random access [8, 9]. In this work, we view the problem of pilot contamination as a random access problem and apply newly developed tools in this area to solve the problem. Two features from the massive MIMO scenario are essential to our solution; near-orthogonality between user channels and near-stability of channel powers. Through signal processing techniques they allow us resolve pilot collisions and thereby utilize otherwise wasted resources. The solution can be viewed as a two-stage processing approach:

1) **Matched filter:** The received uplink pilot and data signals, as expressed in (1) and (2), are processed using matched filters, which are constructed from the contaminated estimates in (5). More specifically, we multiply the received signals with ϕ_{nj}^H creating filtered signals, which we denote f_{nj} and g_{nj} for data and pilots respectively. These signals contain linear combinations of the data and pilots transmitted by the users contributing to the contaminated estimate, ϕ_{nj}, see (8) and (9). The relationship between the variables we wish to estimate and the filtered signals can be viewed as a factor graph, see Fig. 3 and Fig. 4.
2) **Successive interference cancellation:** The coefficients of the linear combinations arising from the matched filtering are the two-norms, $||h_{nk}||^2$, of the involved channels. In a massive MIMO system, with M in the hundreds, this can be assumed slowly fading, contrary to the fast fading channel coefficients. Hence, successive interference cancellation can be applied on the filtered signals in order to reduce the linear combinations to data signals from individual users.

$$\begin{aligned} f_{nj} &= (\phi_{nj}^H Y_n^u)^T \\ &= \sum_{k \in \mathcal{A}_n^j} ||h_{nk}||^2 x_k^u + z_n^{u'} \end{aligned} \tag{8}$$

$$\begin{aligned} g_{nj} &= (\phi_{nj}^H Y_{nj}^{pu})^T \\ &= \sum_{k \in \mathcal{A}_n^j} ||h_{nk}||^2 s_j + z_{nj}^{pu'} \end{aligned} \tag{9}$$

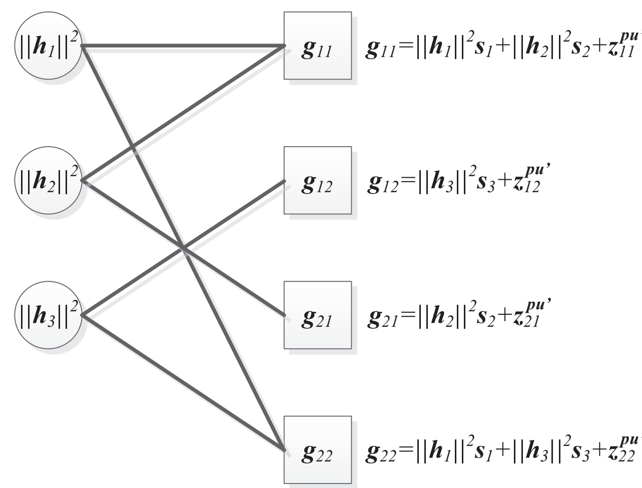

Fig. 3. A graph representation of pilot collisions.

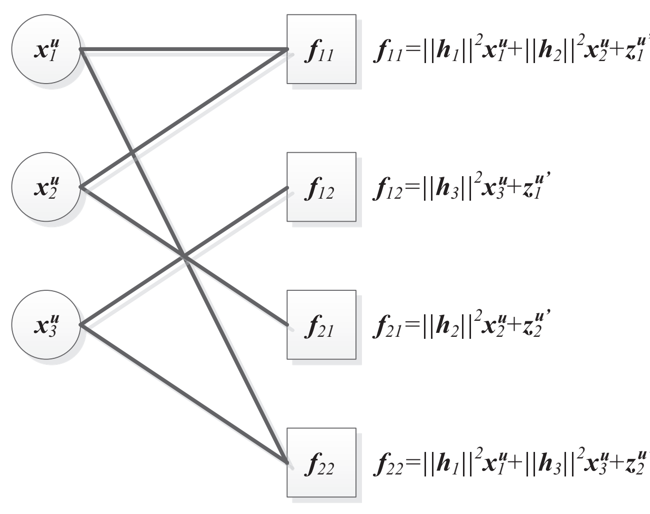

Fig. 4. A graph representation of data collisions.

The purpose of the matched filters is to transform the received signals from linear combinations with fast fading coefficients (the channel coefficients) into linear combinations with slowly fading coefficients (the norms). Note that this is only possible under the conditions given by a massive MIMO scenario, where the norms are stable. When the linear combinations have slowly fading coefficients, it is possible to apply SIC across multiple time slots in order to potentially resolve collisions and improve throughput. More specifically, SIC can be applied across δ time slots, since the norms are assumed constant in this period. We can thus see the filtered signals, $\boldsymbol{f}_{nj}$ and $\boldsymbol{g}_{nj}$ $\forall\, j$ and $n = 1, 2, ..., \delta$, as a code on which iterative belief propagation can be performed. See Fig. 5 for a graph showing the inter-dependencies between $\boldsymbol{f}_{nj}$ and $\boldsymbol{g}_{nj}$.

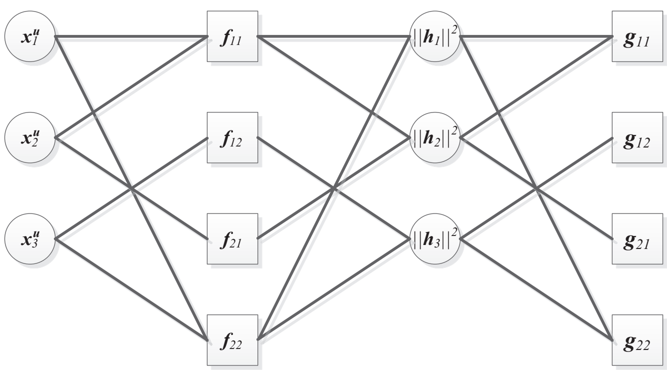

Fig. 5. A graph representation of data and pilot collisions.

Example: Consider the simple example already introduced in Fig. 2. We assume $\delta = 2$, such that the resulting graphs after matched filtering are found in Fig. 3 and Fig. 4. Note, that since the norms are assumed constant, we can omit the time index, such that $||\boldsymbol{h}_k||^2 = ||\boldsymbol{h}_{nk}||^2 \;\forall\; n$. We then have

$$\begin{aligned}
\boldsymbol{f}_{11} &= ||\boldsymbol{h}_1||^2 \boldsymbol{x}_1^u + ||\boldsymbol{h}_2||^2 \boldsymbol{x}_2^u + \boldsymbol{z}_{11}^{u'}, \\
\boldsymbol{f}_{12} &= ||\boldsymbol{h}_3||^2 \boldsymbol{x}_3^u + \boldsymbol{z}_{12}^{u'}, \\
\boldsymbol{f}_{21} &= ||\boldsymbol{h}_2||^2 \boldsymbol{x}_2^u + \boldsymbol{z}_{21}^{u'}, \\
\boldsymbol{f}_{22} &= ||\boldsymbol{h}_1||^2 \boldsymbol{x}_1^u + ||\boldsymbol{h}_3||^2 \boldsymbol{x}_3^u + \boldsymbol{z}_{22}^{u'}, \\
\boldsymbol{g}_{11} &= ||\boldsymbol{h}_1||^2 \boldsymbol{s}_1 + ||\boldsymbol{h}_2||^2 \boldsymbol{s}_1 + \boldsymbol{z}_{11}^{pu'}, \\
\boldsymbol{g}_{12} &= ||\boldsymbol{h}_3||^2 \boldsymbol{s}_2 + \boldsymbol{z}_{12}^{pu'}, \\
\boldsymbol{g}_{21} &= ||\boldsymbol{h}_2||^2 \boldsymbol{s}_1 + \boldsymbol{z}_{21}^{pu'}, \\
\boldsymbol{g}_{22} &= ||\boldsymbol{h}_1||^2 \boldsymbol{s}_2 + ||\boldsymbol{h}_3||^2 \boldsymbol{s}_2 + \boldsymbol{z}_{22}^{pu'}.
\end{aligned} \tag{10}$$

We introduce the variable $\boldsymbol{c}$ which accounts for accumulated noise components and estimation errors. Note, that the magnitude of the elements in $\boldsymbol{c}$ increases as processing progresses. This will be discussed in further detail in section IV.

Initially, we can isolate the contribution from user 1 to the filtered signal, giving us

$$||\boldsymbol{h}_1||^2 \boldsymbol{x}_1^u + \boldsymbol{c} = \boldsymbol{f}_{11} - \boldsymbol{f}_{21}. \tag{11}$$

Since the applied pilot sequence is known a priori by the base station, we can find the norm as

$$||\boldsymbol{h}_1||^2 + \boldsymbol{c} = \tau^{-1} (\boldsymbol{g}_{11} - \boldsymbol{g}_{21})^T \boldsymbol{s}_1^{-1}. \tag{12}$$

Finally, we can express the estimate of the message from user 1, $\hat{\boldsymbol{x}}_1^u$, as

$$\hat{\boldsymbol{x}}_1^u = \tau (\boldsymbol{f}_{11} - \boldsymbol{f}_{21}) ((\boldsymbol{g}_{11} - \boldsymbol{g}_{21})^T \boldsymbol{s}_1^{-1})^{-1}. \tag{13}$$

Similar operations can be performed in order to find $\hat{\boldsymbol{x}}_2^u$ and $\hat{\boldsymbol{x}}_3^u$.

B. Downlink

In the downlink phase we assume channel reciprocity, such that the user does not need to estimate each coefficient of $\boldsymbol{h}_{nk}$, which would require time divided pilot signals on all M antennas. Instead, we let the receiver estimate the concatenated "channel" consisting of both the downlink precoder, $\boldsymbol{w}_{nk}$, and the actual channel. Denoting the concatenated channel, q_{nk}, we have

$$q_{nk} = \boldsymbol{h}_{nk}^T \boldsymbol{w}_{nk}. \tag{14}$$

Estimating q_{nk} can be realized with a single pilot sequence, as expressed in (3).

In order to choose an appropriate precoder, the base station must have an estimate of the current channel, $\boldsymbol{h}_{nk}$. The coded operation applied in uplink does not guarantee that such an estimate is available. Uplink operation relies on SIC based only on knowledge of the norm. Hence, downlink transmission to a particular user is only possible if that user avoided collision during the previous uplink pilot phase, such that an uncontaminated channel estimate is available. This incurs a delay in downlink transmissions, whose magnitude is analyzed in section III-C.

C. Analysis

The performances of the reference scheme and the proposed scheme are tightly connected with the factor node degree distribution of the resulting graph. Here a factor degree, denoted as d_{nj}, refers to the number of users occupying the same resource block, i.e. applies the j'th pilot sequence in the n'th time slot. A user is active and applying pilot sequence j with probability p_a/τ, such that the degree probability distribution is

$$Pr(d_{nj} = d) = \binom{N}{d} \left(\frac{p_a}{\tau}\right)^d \left(1 - \frac{p_a}{\tau}\right)^{N-d}. \tag{15}$$

For the ALOHA scheme, we found that the optimal performance is achieved when the probability of having $d_{nj} = 1$ is maximized. Differentiating $Pr(d_{nj} = 1)$ with respect to p_a and finding the roots satisfying our conditions, we get that $p_a = \frac{\tau}{N}$ maximizes the performance of the ALOHA scheme.

We can not use the same approach for optimizing the proposed scheme, since resource slots with $d_{nj} > 1$ may be useful. Instead we must seek a well performing degree distribution which favors the iterative belief propagation. Several works have studied this, e.g. in [10, 11], however, in this work we are not able to freely tailor the degree distribution. We are limited to the binomial distribution as expressed in (15), with only the freedom to choose a proper p_a. Similar limitations were considered in [9] with a focus on choosing a proper average degree, $\bar{d}$, which was optimized numerically. In our context, we have

$$\bar{d} = \frac{p_a N}{\tau}. \tag{16}$$

A numerical optimization of $\bar{d}$ and thereby in turn p_a for a specific pair of N and τ will be performed in section IV.

As described in section III-B, downlink transmissions experience a delay due to lack of channel knowledge. We denote the delay for user k, Δ_k. This delay is equal to the number of time slots until user k is active and avoids a collision during the uplink pilot phase. Denoting the probability of a user being active and avoiding collision, p_a^*, we have

$$p_a^* = p_a \left(1 - \frac{p_a}{\tau}\right)^{N-1}. \tag{17}$$

The probability distribution of Δ_k follows the negative binomial and is therefore given by

$$Pr(\Delta_k = \delta) = p_a^*(1 - p_a^*)^{\delta-1}. \tag{18}$$

The expected value, $\mathrm{E}[\Delta_k]$, of the delay is then found as

$$\mathrm{E}[\Delta_k] = \frac{(1 - p_a^*)}{p_a^*}. \tag{19}$$

There is a natural tradeoff between optimizing p_a for high throughput in the uplink phase and optimizing it for limiting the delay in the downlink phase. Such a joint optimization is outside the scope of this work. In the numerical evaluations in section IV, we will solely be concerned with the uplink throughput.

IV. Numerical Results

The proposed scheme is simulated and compared to framed slotted ALOHA in terms of uplink throughput and block error rate. The proposed scheme is based on an assumption that the channel coefficients in different time slots are i.i.d, while their two-norms remain constant within a period of β time slots. In the numerical evaluations, we will challenge these assumptions by simulating with fading channels. A rich scattering environment is assumed, such that $h_{nk}(m)$ can be modeled using Clarke's model [12], hence

$$h_{nk}(m) = \frac{1}{\sqrt{N_s}} \sum_{i=1}^{N_s} \mathrm{e}^{j2\pi f_d n t_s \cos\alpha_i + \phi_i}, \tag{20}$$

where N_s is the number of scatterers, f_d is the maximum Doppler shift, α_i and ϕ_i is the angle of arrival and initial phase, respectively, of the wave from the i'th scatterer. Both α_i and ϕ_i are i.i.d. in the interval $[-\pi, \pi)$ and $f_d = \frac{v}{c} f_c$, where v is the speed of the UE, c is the speed of light and f_c is the carrier frequency. An overview of the parameters, which are common for all simulations, is given in Table I. Note that $\beta = 1.2N/\tau$ is chosen in order to ensure a 20% surplus of resource blocks relative to N, such that the iterative belief propagation performs well. All simulation results are averages over $10,000$ iterations.

Initially, in Fig. 6 we present results for the normalized throughput of the proposed scheme as a function of the average degree of a resource block, which is directly related to the activation probability, as seen in (16). We define normalized throughput as the total number of successfully decoded messages divided by $\beta\tau$, i.e. the total amount of resource blocks. It is evident that an average degree of approximately 2.5 should be aimed for, which is confirmed by the results from [9]. All other simulations are performed using an average degree of 2.5.

Fig. 7 shows the normalized goodput, i.e. throughput with erroneous messages discarded, as a function of the number of users accessing the base station. It is seen that the proposed scheme clearly outperforms the conventional method of framed slotted ALOHA. The gap between the two schemes increases with N, since the proposed scheme benefits from a larger number of messages to code across. An increase in N can be viewed as an increase in the block length, which improves the coding efficiency.

The coding gain comes at the price of an increased error rate. As mentioned in section III, whenever SIC is performed, noise and estimation errors are accumulated, which may lead to errors. At higher N, it is more common to see high degrees in the code graph, even if the average degree remains constant. Moreover, SIC is performed across a larger time span, which leads to larger errors in the norm estimation. As a result, we experience an increased error rate for increasing N, which is illustrated in Fig. 8. The figure also shows that the error rate drops significantly, when the number of base station antennas is increased. The reason is that the norm stabilizes for increasing M, such that the assumption of a constant norm is increasingly valid.

V. Conclusions

We have presented a solution for the pilot contamination problem in crowded scenarios, where users within a single cell must share a small set of pilot sequences. We view the problem of intra-cell pilot contamination as a random access problem and draw on newly developed ideas from this area of research. The massive MIMO setting provides two essential properties; near-orthogonality between user channels and near-stability of channel powers. These properties allow us to view a set of contaminated pilot signals as a graph code on which iterative belief propagation can be performed. The proposed solution proves highly efficient, comfortably outperforming the

TABLE I
SIMULATION PARAMETERS

Parameter	Value	Description
f_c	1.8 GHz	Carrier frequency
v	3 km/h	User mobility
N_s	20	Number of scatterers
σ_n^2	0.1	Relative noise power
τ	5 bits	Length and number of pilot sequences
t_s	0.01 s	Length of a time slot
L	1000 bits	Length of uplink data messages
β	$1.2N/\tau$	Number of time slots

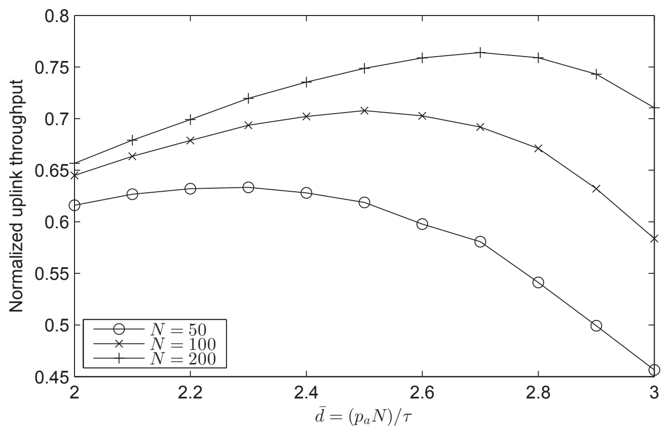

Fig. 6. Throughput as a function of the average degree of the resource blocks.

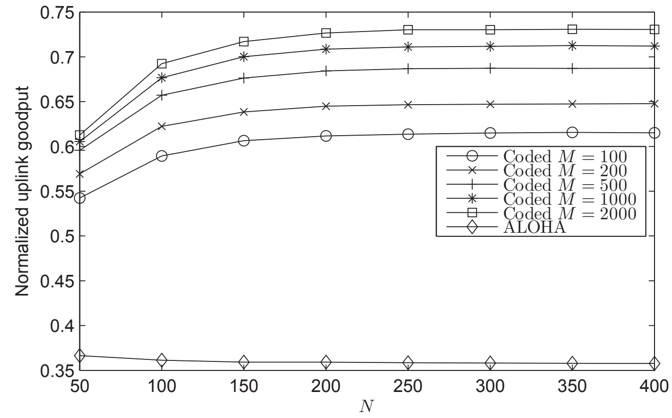

Fig. 7. Throughput as a function of the number of users.

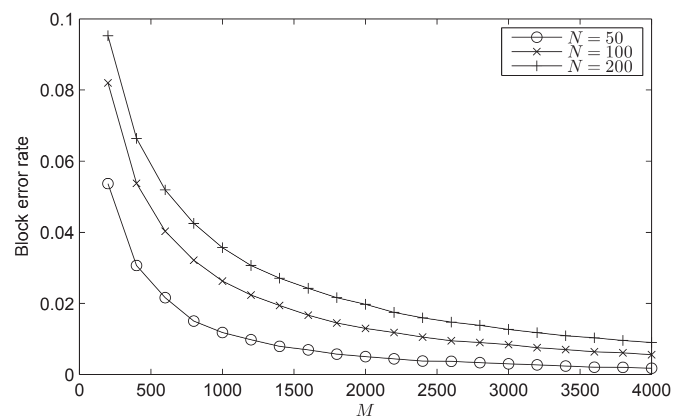

Fig. 8. Block error rate as a function of the number of base station antennas.

conventional ALOHA approach to random access. In terms of pilot contamination, we demonstrate that hundreds of users can be served using only five orthogonal pilot sequences. The price to pay is an increased error rate, due to accumulation of estimation errors in the belief propagation algorithm. However, this downside is shown to significantly diminish as the number of base station antennas increases.

REFERENCES

[1] T. Marzetta, "How much training is required for multiuser mimo?," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, pp. 359–363, Oct 2006.

[2] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5g," *Communications Magazine, IEEE*, vol. 52, pp. 74–80, February 2014.

[3] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up mimo: Opportunities and challenges with very large arrays," *Signal Processing Magazine, IEEE*, vol. 30, pp. 40–60, Jan 2013.

[4] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *Selected Areas in Communications, IEEE Journal on*, vol. 31, pp. 264–273, February 2013.

[5] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pp. 1137–1141, July 2012.

[6] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell tdd systems," *Wireless Communications, IEEE Transactions on*, vol. 10, pp. 2640–2651, August 2011.

[7] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5g mobile and wireless communications: the vision of the metis project," *Communications Magazine, IEEE*, vol. 52, pp. 26–35, May 2014.

[8] G. Liva, "Graph-Based Analysis and Optimization of Contention Resolution Diversity Slotted ALOHA," *Communications, IEEE Transactions on*, vol. 59, pp. 477 –487, february 2011.

[9] C. Stefanovic, P. Popovski, and D. Vukobratovic, "Frameless ALOHA protocol for Wireless Networks," *IEEE Comm. Letters*, vol. 16, pp. 2087–2090, Dec. 2012.

[10] M. Luby, "LT Codes," in *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pp. 271–280, 2002.

[11] J. Sørensen, P. Popovski, and J. Østergaard, "Design and analysis of lt codes with decreasing ripple size," *Communications, IEEE Transactions on*, vol. 60, pp. 3191 –3197, november 2012.

[12] R. Clarke, "A statistical theory of mobile-radio reception," *Bell system technical journal*, vol. 47, no. 6, pp. 957–1000, 1968.