IEEE ICC2016-Workshops: W14-Workshop on Massive Uncoordinated Access Protocols (MASSAP)

On the Sparsity for Random Access in Machine Type Communications under Frequency-Selective Fading

Jinho Choi
School of Information and Communications
Gwangju Institute of Science and Technology (GIST)
Email: jchoi0114@gist.ac.kr

***Abstract*—In this paper, we study random access for machine type communications (MTC). In particular, we consider a preamble detection method that can exploit the sparsity of active devices over frequency-selective fading channels to detect multiple preambles in random access. In order to understand the impact of various design parameters and channels, we analyze the performance in terms of the average number of successful preambles when a compressive sensing (CS) based approach is used for the preamble detection. For example, through the analysis, we can see that the performance of the CS-based estimation for the preamble detection can be limited by the length of the channel impulse response (CIR).**

***Index Terms*—random access, machine type communications, compressive sensing**

I. Introduction

There has been a growing interest in machine-type communications (MTC) or machine-to-machine (M2M) communications in order to support a number of devices that are to be connected to a network [1]. The applications of MTC are diverse from health care to wireless sensor networks and Internet of Things (IoT). For MTC, random access is usually considered as there are a number of devices with a low probability of activity [1]–[3]. In the long term evolution-advanced (LTE-A) system, a random access scheme, called random access (RACH) procedure, has also been proposed for MTC [4].

In the RACH procedure, devices that want to establish connections to an access point (AP) transmit randomly chosen preambles from a pool of preambles simultaneously. Thus, the AP needs to detect multiple preambles. It is possible to employ a conventional multiuser detection approach [5]. However, since only a fraction of devices are active, it would be efficient to exploit the notion of compressive sensing (CS) [6]–[9] to derive low-complexity detection methods in [10]–[14].

In this paper, we consider a low-complexity CS-based approach for the multiple preamble detection in random access. Unlike most existing approaches, we focus on the multiple preamble detection over frequency-selective fading channels. Note that in [15], [16], for random access, CS-based multiple preamble detection approaches over frequency-selective fading channels are also studied. In particular, the approach in [16] is similar to that in this paper as both the approaches are based on multicarrier systems, although there are some differences that will be explained later (in Section III-B). We also present performance analysis of the proposed CS-based multiple preamble detection method in order to see the impact of key parameters and channels on the performance.

Notation: Matrices and vectors are denoted by upper- and lower-case boldface letters, respectively. The superscripts T and H denote the transpose and complex conjugate, respectively. The p-norm of a vector $\mathbf{a}$ is denoted by $||\mathbf{a}||_p$ (If $p = 2$, the norm is denoted by $||\mathbf{a}||$ without the subscript). The superscript $\dagger$ denotes the pseudo-inverse. For a vector $\mathbf{a}$, $\mathrm{diag}(\mathbf{a})$ is the diagonal matrix with the diagonal elements from $\mathbf{a}$. For a matrix $\mathbf{X}$ (a vector $\mathbf{a}$), $[\mathbf{X}]_n$ ($[\mathbf{a}]_n$) represents the nth column (element, resp.). If n is a set of indices, $[\mathbf{X}]_n$ is a submatrix of $\mathbf{X}$ obtained by taking the corresponding columns. $\mathbb{E}[\cdot]$ and $\mathrm{Var}(\cdot)$ denote the statistical expectation and variance, respectively. $\mathcal{CN}(\mathbf{a}, \mathbf{R})$ ($\mathcal{N}(\mathbf{a}, \mathbf{R})$) represents the distribution of circularly symmetric complex Gaussian (CSCG) (resp., real-valued Gaussian) random vectors with mean vector $\mathbf{a}$ and covariance matrix $\mathbf{R}$.

II. System Model for Random Access

Suppose that there are a number of devices that might be connected to an AP. The total number of devices is denoted by K. While K can be very large, only a fraction of them can be active and send short messages to an AP at a time. Thus, random access can be used for communications from devices to the AP.

Throughout the paper, we assume a multicarrier system of N subcarriers for uplink transmissions from devices to the AP. The channel matrix from device k to the AP is denoted by $\mathbf{H}_k$ in the frequency domain, which is a diagonal matrix and given by

$$\mathbf{H}_k = \mathrm{diag}(H_{k,0}, \ \dots H_{k,N-1}), \tag{1}$$

where

$$H_{k,n} = \sum_{p=0}^{P-1} h_{k,p} e^{-\frac{2\pi pn}{N}}.$$

Here, $h_{k,p}$ is the channel impulse response (CIR) from device k to the AP and P is the length of CIR.

Suppose that a pool of preambles, denoted by $\mathcal{C}$, is used for random access as in [4]. To establish a connection to the AP, a device can randomly choose a preamble from $\mathcal{C}$ and transmit it to the AP. All the devices are synchronized and active devices can transmit their randomly selected preambles simultaneously. For convenience, denote by L the number of preambles in $\mathcal{C}$. In addition, let $\mathbf{c}_l$ represent the lth preamble in $\mathcal{C}$, i.e., $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_L\}$. We assume that the length of $\mathbf{c}_l$ is the same as the number of subcarriers, N (we ignore cyclic prefix for convenience). Consequently, we can see that multicarrier-code division multiple access (MC-CDMA) is adopted for random access.

We denote by $l(m)$ the index of the preamble chosen by the mth active device and define as $\mathcal{A} = \{l(1), \ldots, l(M)\}$, where M is the number of active devices among K devices. Then, the received signal vector at the AP is given by

$$\mathbf{y} = \sum_{m \in \mathcal{A}} \mathbf{H}_{k(m)} \mathbf{c}_{l(m)} + \mathbf{n}, \tag{2}$$

where $k(m)$ is the device index of the mth active device and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, N_0 \mathbf{I})$ is the background noise vector. For example, let $K = 64$, $L = 10$, and $M = 2$. If the device indices of the first and second active devices are 2 and 54, respectively, we have $k(1) = 2$ and $k(2) = 54$. If the first and second active devices choose preambles 4 and 10, respectively, we have $l(1) = 4$ and $l(2) = 10$.

In (2), we assume perfect synchronization. In addition, we will assume perfect channel estimation at the AP throughout the paper. Note that in [15], [17], the channel estimation is studied in random access with the sparsity of active devices.

III. Compressive Sensing Approach for Preamble Detection

In this section, we study the preamble detection by exploiting the sparsity of active devices [11], [15].

A. The Case of Frequency-Flat Fading

Suppose that

$$\mathbf{H}_k = h_k \mathbf{I}. \tag{3}$$

where $h_k \in \mathbb{C}$. This is the case that the channels from devices to the AP are modeled as flat-fading channels. Then, (2) is rewritten as

$$\mathbf{y} = \sum_{m \in \mathcal{A}} h_{k(m)} \mathbf{c}_{l(m)} + \mathbf{n}. \tag{4}$$

Define the following measurement matrix:

$$\mathbf{C} = [\mathbf{c}_1 \ \ldots \ \mathbf{c}_L] \in \mathbb{C}^{N \times L}. \tag{5}$$

Using $\mathbf{C}$, from (4), $\mathbf{y}$ becomes

$$\mathbf{y} = \mathbf{C}\mathbf{a} + \mathbf{n}, \tag{6}$$

where $\mathbf{a}$ is M-sparse (i.e., there are M non-zero elements in $\mathbf{a}$). In particular, we have

$$[\mathbf{a}]_l = \begin{cases} h_{k(m)}, & \text{if } l = l(m); \\ 0, & \text{otherwise.} \end{cases}$$

Let Σ_S denote the set of S-sparse vectors of length L. Then, the maximum likelihood (ML) estimation problem of $\mathbf{a}$ can be formulated as follows:

$$\hat{\mathbf{a}} = \underset{\mathbf{a} \in \Sigma_S}{\operatorname{argmin}} ||\mathbf{y} - \mathbf{C}\mathbf{a}||^2. \tag{7}$$

Since the computational complexity to solve (7) is prohibitively high, we need to exploit the sparsity of $\mathbf{a}$ for the case that $M \ll L$ in order to derive low-complexity approaches.

Suppose that $M \ll L$ and $M < N$. Then, from [7], [9], [18], to find an approximate solution to (7), we can have the following problem:

$$\min_{\mathbf{a}} ||\mathbf{y} - \mathbf{C}\mathbf{a}||^2 + \kappa ||\mathbf{a}||_1, \tag{8}$$

where $\kappa > 0$ is the Lagrange multiplier. Although (8) is a convex problem, the complexity to solve this problem can be still high. Thus, various suboptimal but low-complexity approaches have been proposed to find approximate solutions such as greedy algorithms [9, Chp. 8].

In addition to low-complexity suboptimal algorithms, the conditions for the measurement matrix are important. In (6), since we can determine the measurement matrix $\mathbf{C}$ as the column vectors are the preambles in $\mathcal{C}$, it might be possible to have $\mathbf{C}$ with the restricted isometry property (RIP) [6]. For example, if the elements of $\mathbf{C}$ are binary random variables, with a high probability, the RIP can be satisfied. Once the RIP is satisfied, the number of the measurements, N has to be [9], [19]

$$N \geq \alpha M \log \frac{L}{M} = \alpha \omega_L(M), \tag{9}$$

where α is constant, for successful sparse signal estimation in most CS-based approach.

B. The Case of Frequency-Selective Fading

The approach in Subsection III-A is limited and cannot be used for frequency-selective fading channels. Thus, in this subsection, we focus on the modification to be used for frequency-selective fading channels.

If device k transmits the lth preamble, the received signal vector without interference and background noise can be written as

$$\mathbf{H}_k \mathbf{c}_l = \mathbf{C}_l \mathbf{h}_k, \tag{10}$$

where $\mathbf{h}_k = [H_{k,0} \ \ldots \ H_{k,N-1}]^{\mathrm{T}}$ and $\mathbf{C}_l = \mathrm{diag}(\mathbf{c}_l)$. Furthermore, it can be shown that

$$\mathbf{h}_k = \mathbf{F}\mathbf{v}_k, \tag{11}$$

where $\mathbf{v}_k = [h_{k,0} \ \ldots \ h_{k,P-1}]^{\mathrm{T}}$ and $[\mathbf{F}]_{l,p} = e^{-\frac{j2\pi lp}{N}}$, $l \in \{0, \ldots, N-1\}$, $p \in \{0, \ldots, P-1\}$. Then, (2) is rewritten as

$$\begin{aligned} \mathbf{y} &= \sum_{m \in \mathcal{A}} \mathbf{H}_{k(m)} \mathbf{c}_{l(m)} + \mathbf{n} \\ &= \sum_{m \in \mathcal{A}} \mathbf{C}_{l(m)} \mathbf{F} \mathbf{v}_{k(m)} + \mathbf{n} \\ &= \mathbf{\Psi}\mathbf{b} + \mathbf{n}, \end{aligned} \tag{12}$$

where

$$\mathbf{\Psi} = [\mathbf{C}_1\mathbf{F} \ \ldots \ \mathbf{C}_L\mathbf{F}] \in \mathbb{C}^{N \times PL} \quad (13)$$

and $\mathbf{b}$ is a PM-sparse vector of length PL. Note that if $P = 1$ (i.e., flat fading channels), $\mathbf{F} = [1 \ \ldots \ 1]^{\mathrm{T}}$ and $\mathbf{C}_l\mathbf{F} = \mathbf{c}_l$, which implies that $\mathbf{\Psi} = \mathbf{C}$ (i.e., (12) becomes (6)).

With (12), the required number of measurements for the sparse signal estimation, N, becomes

$$N \geq \alpha PM \log \frac{PL}{PM} = \alpha P \omega_L(M). \quad (14)$$

This shows that the required number of measurements grows linearly with P.

In this paper, we modify the orthogonal matching pursuit (OMP) algorithm [20], [21] to estimate the sparse signal vector $\mathbf{b}$ in (12). While the conventional OMP algorithm can be used, the block sparsity of $\mathbf{b}$ could be taken into account for a better performance. That is, since every P elements of $\mathbf{b}$ would be either zeros or non-zeros, this block sparsity can be used as a constraint for a better performance. In Section V, we use this modified OMP algorithm for simulations.

As mentioned earlier, the random access scheme in this paper is similar to that in [16]. However, there is a key difference. In [16], each device has a unique[1] preamble. Thus, there should be K preambles. On the other hand, in this paper, we consider a pool of L preambles, $\mathcal{C}$, based on [4]. Due to this difference, the sizes of the measurement matrices are different. In [16], the size of the measurement matrix is $N \times KN$. In general, since $P \lll N$ and $L \lll K$, we can see that $\mathbf{\Psi}$ (of size $N \times LP$) has a much smaller number of columns than the measurement matrix in [16] has. This results in the difference in computational complexity. For example, as shown in [9, Chp. 8], the complexity of the OMP algorithm is approximately $O(mS + n)$, where m and n represent the number of rows and columns of the measurement matrix, respectively, and S is the sparsity. Thus, the complexities of the OMP algorithms for the approach in [16] and that in this paper become $O(NMP + KN)$ and $O(NMP + LP)$, respectively. If $K \ggg L$, the computational complexity of the proposed approach becomes a much lower than that in [16].

IV. Performance Analysis

A. Outage Probability

As shown in (14), the CS-based estimation may not provide a reasonable performance if the number of measurements, N, is not sufficiently large for a given M. Since the number of active devices is a random variable, we can define an outage event as the event that M does not satisfy the relation in (14) for given L, N, and P. Thus, the outage probability can be given by

$$P_{\mathrm{out}} = \Pr\left(N \leq \alpha P \omega_L(M)\right). \quad (15)$$

[1]In this case, i.e., each device has a unique preamble, the device identification becomes straightforward as the detection of preamble is equivalent to the device identification. However, in the RACH procedure [4] (and the approach in this paper as well) for the active device identification, there should be additional steps once a device is connected to the AP as a preamble is not associated with any device and is randomly chosen.

Since $\log(x) \geq 1 - \frac{1}{x}$, we have the following upper-bound on $\omega_L(M)$:

$$\omega_L(M) \leq \bar{\omega}_L(M) = M(\log(L) - 1) + 1. \quad (16)$$

Using this upper-bound, an upper-bound on the outage probability can be found as

$$\begin{aligned} P_{\mathrm{out}} &\leq \Pr\left(N \leq \alpha P \bar{\omega}_L(M)\right) \\ &= \Pr\left(M \geq \frac{\frac{N}{\alpha P} - 1}{\log(L) - 1}\right). \end{aligned} \quad (17)$$

For convenience, let

$$\tau(N, L) = \frac{\frac{N}{\alpha P} - 1}{\log(L) - 1}. \quad (18)$$

We assume that each device can be active independently with a probability p_a, which is referred to as the activity probability. Then, the probability that $M = m$ becomes

$$\begin{aligned} \Pr(M = m) &= \binom{K}{m} p_a^m (1 - p_a)^{K-m} \\ &\approx \frac{\lambda^m}{m!} e^{-\lambda}, \end{aligned} \quad (19)$$

where $\lambda = K p_a$. In (19), the approximation is the Poisson approximation [22]. From (19), the upper-bound on the outage probability becomes

$$\begin{aligned} P_{\mathrm{out}} &\leq 1 - \sum_{m=0}^{\lfloor \tau(N,L) \rfloor} \Pr(M = m) \\ &\approx 1 - \sum_{m=0}^{\lfloor \tau(N,L) \rfloor} \frac{\lambda^m}{m!} e^{-\lambda}, \end{aligned} \quad (20)$$

where the approximation is due to the Poisson approximation.

B. Average Number of Successful Preambles

While the outage probability can be used to see the performance, we may gain better insight into the performance by the average number of successful preambles. In this subsection we derive an expression for the average number of successful preambles conditioned on no outage.

The expected number of preambles without collisions when M active devices randomly choose preambles from $\mathcal{C}$ is given by

$$M_{\mathrm{s}} = M \left(1 - \frac{1}{L}\right)^{M-1}. \quad (21)$$

As mentioned earlier, if a CS-based approach is used to detect preambles, there is an outage event when M is greater than the threshold in (18). Thus, if $M \geq \lfloor \tau(N, L) \rfloor$, we can assume that the number of successful preambles (preambles without collisions) is[2] 0. From this, the average number of successful

[2]This assumption makes the analysis tractable. However, in practice, although the outage happens, there might be some preambles that are correctly detected. Thus, this assumption would lead to a pessimistic result.

preambles can be found as

$$\bar{M}_{\rm s}(\lambda, L, N) = \sum_{m=1}^{\lfloor \tau(N,L) \rfloor} m \left(1 - \frac{1}{L}\right)^{m-1} \Pr(M = m)$$
$$\approx \lambda e^{-\lambda} \sum_{m=0}^{\lfloor \tau(N,L) \rfloor - 1} \frac{\left(\lambda \left(1 - \frac{1}{L}\right)\right)^m}{m!}. \quad (22)$$

Again, the approximation in (22) is based on the Poisson approximation.

From (22), we can see the impact of key parameters on the performance in terms of the average number of successful preambles. For example, since $\tau(N, L)$ increases with N for a given L, it is easy to see that $\bar{M}_{\rm s}(\lambda, L, N)$ in (22) increases with N. This would be a natural consequence as a better performance is achieved when more subcarriers are used.

V. Simulation Results

In this section, we present simulation results with the modified OMP algorithm. For simulations, each coefficient of the CIR is assumed to be independent and has the following Gaussian distribution:

$$h_{k,p} \sim \mathcal{CN}\left(0, \frac{1}{P}\right), \ p = 0, \ldots, P - 1.$$

For preambles, $\{\mathbf{c}_l\}$, we consider random binary vectors with $||\mathbf{c}_l|| = 1$. The signal-to-noise ratio (SNR) is assumed to be $\frac{1}{N_0}$ as each preamble is normalized to have unit-norm.

Fig. 1 shows the average number of successful preambles for different values of the activity probability, p_a, when $K = 100$, $N = 64$, $L = 128$, $P = 4$, and SNR $= 10$ dB. It is shown that the theoretical results from (22) are in good agreement with the simulation results when p_a is small. However, as p_a increases, the average number of successful preambles from simulations becomes larger than that from (22). In deriving (22), we assume that no preambles are detected when the outage event happens. However, as mentioned earlier, there could be some preambles that can be detected even if the outage event happens, which results in a better performance in simulation results.

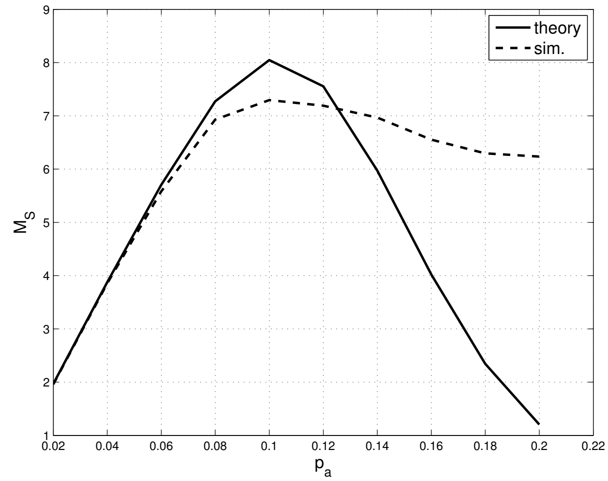

Fig. 1. The average number of successful preambles versus p_a when $K = 100$, $N = 64$, $L = 128$, $P = 4$, and SNR $= 10$ dB.

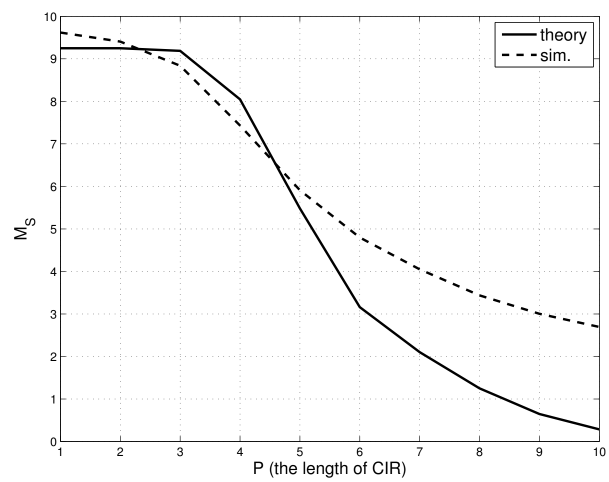

Fig. 2. The average number of successful preambles versus the length of CIR, P, when $K = 100$, $N = 64$, $L = 128$, $p_a = 0.1$, and SNR $= 10$ dB.

Fig. 2 shows the average number of successful preambles for different lengths of CIR when $K = 100$, $N = 64$, $L = 128$, $p_a = 0.1$, and SNR $= 10$ dB. It is shown that the average number of successful preambles decreases with P. Since the probability of outage event increases with P, the increase of P results in the decrease of the average number of successful preambles. Consequently, we can see that when the sparsity is exploited in random access for MTC, the frequency-selective fading environment is not desirable.

We may expect to have a better performance in random access if there are more preambles as there would be less preamble collisions. In Fig. 3, we show the average number of successful preambles for various values of L when $K = 100$, $N = 64$, $P = 4$, and SNR $= 10$ dB. When the activity probability, p_a, is low (i.e., $p_a = 0.05$), we can see that the average number of successful preambles can increase with L. However, for a larger p_a (i.e., $p_a = 0.1$), the average number of successful preambles can decrease with L once L is sufficiently large. This behavior results from the outage event in the CS-based estimation as the outage probability increases with L. Based on this observation, we can see that a large L is not desirable in terms of both performances and complexity (as the complexity increases with L).

Fig. 2 shows the average number of successful preambles for different values of N when $K = 100$, $P = 4$, $L = 128$, $p_a = 0.1$, and SNR $= 10$ dB. We expect that the performance can be improved if more subcarriers are used, i.e., the average number of successful preambles increases with N. Note that the average number of successful preambles cannot be larger than Kp_a. Thus, the average number of successful preambles increases, and then becomes saturated as N increases as shown in Fig. 4. This indicates that p_a might be adjusted for a given N.

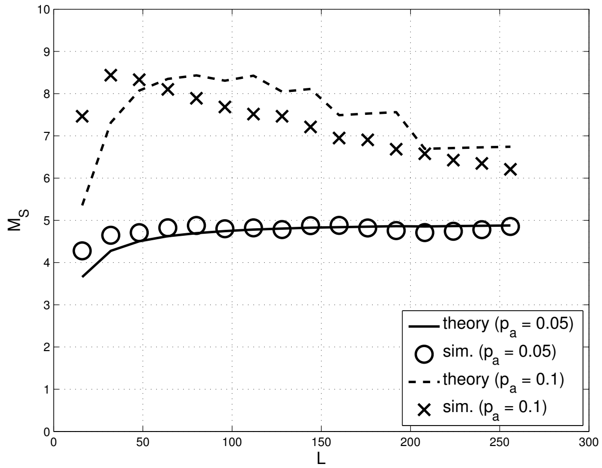

Fig. 3. The average number of successful preambles versus L when $K = 100$, $N = 64$, $P = 4$, and SNR = 10 dB.

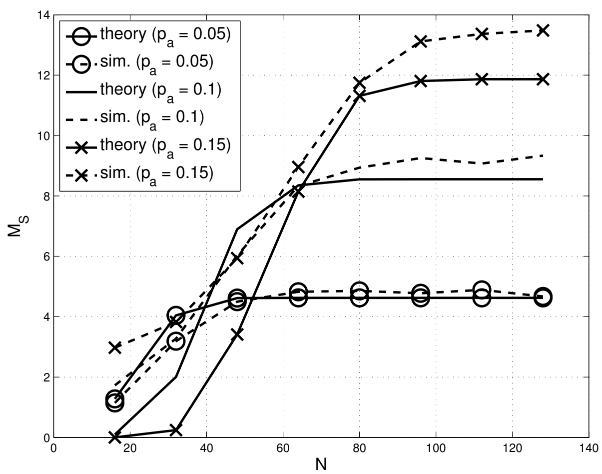

Fig. 4. The average number of successful preambles versus N when $K = 100$, $L = 64$, $P = 4$, and SNR = 10 dB.

VI. Concluding Remarks

In this paper, we studied random access for MTC over frequency-selective fading channels. For the random access scheme, we considered an approach based on the RACH procedure in [4] and exploited the sparsity of active devices to derive a low-complexity CS-based approach for detecting multiple preambles transmitted by a few active devices simultaneously. We analyzed the performance of the CS-based multiple preamble detection method in terms of the number of successful preambles in order to see the impact of key parameters on the performance. We had a few interesting observations. For example, the increase of the number of preambles did not help improve the performance unless the activity probability is sufficiently low. It was also shown that the performance becomes worse as the length of CIR increases. Thus, when the sparsity is exploited in random access for MTC, the frequency-selective fading environment seems not desirable.

References

[1] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in lte-advanced networks: issues and approaches," *IEEE Communications Magazine*, vol. 51, pp. 86–93, June 2013.

[2] K.-D. Lee, S. Kim, and B. Yi, "Throughput comparison of random access methods for m2m service over lte networks," in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pp. 373–377, Dec 2011.

[3] D. Niyato, P. Wang, and D. I. Kim, "Performance modeling and analysis of heterogeneous machine type communications," *IEEE Trans. Wireless Communications*, vol. 13, pp. 2836–2849, May 2014.

[4] 3GPP TR 37.868 V11.0, *Study on RAN improvments for machine-type communications*, October 2011.

[5] S. Verdu, *Multiuser Detection*. Cambridge University Press, 1998.

[6] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Information Theory*, vol. 51, pp. 4203–4215, Dec 2005.

[7] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, April 2006.

[8] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[9] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[10] A. K. Fletcher, S. Rangan, and V. K. Goyal, "On-off random access channels: A compressed sensing framework," *CoRR*, vol. abs/0903.1022, 2009.

[11] H. Zhu and G. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Trans. Communications*, vol. 59, pp. 454–465, February 2011.

[12] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, "Asynchronous code-division random access using convex optimization," *Physical Communication*, vol. 5, no. 2, pp. 129–147, 2012. Compressive Sensing in Communications.

[13] F. Fazel, M. Fazel, and M. Stojanovic, "Random access compressed sensing over fading and noisy communication channels," *IEEE Trans. Wireless Communications*, vol. 12, pp. 2114–2125, May 2013.

[14] J.-P. Hong, W. Choi, and B. Rao, "Sparsity controlled random multiple access with compressed sensing," *IEEE Trans. Wireless Communications*, vol. 14, pp. 998–1010, Feb 2015.

[15] H. F. Schepker, C. Bockelmann, and A. Dekorsy, "Exploiting sparsity in channel and data estimation for sporadic multi-user communication," in *Proc. International Symposium on Wireless Communication Systems (ISWCS 2013)*, pp. 1–5, Aug 2013.

[16] G. Wunder, P. Jung, and M. Ramadan, "Compressive random access using A common overloaded control channel," *CoRR*, vol. abs/1504.05318, 2015.

[17] G. Wunder, P. Jung, and C. Wang, "Compressive random access for post-LTE systems," in *Proc. IEEE ICC*, pp. 539–544, June 2014.

[18] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Information Theory*, vol. 52, pp. 489–509, Feb 2006.

[19] E. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[20] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, pp. 40–44 vol.1, Nov 1993.

[21] G. M. Davis, S. G. Mallat, and Z. Zhang, "Adaptive time-frequency decompositions," *Optical Engineering*, vol. 33, no. 7, pp. 2183–2191, 1994.

[22] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probability Analysis*. Cambridge University Press, 2005.