# A New Approach for Sparse Bayesian Channel Estimation in SCMA Uplink Systems

Kirill Struminsky *, Stanislav Kruglik †‡§, Dmitry Vetrov *†, Ivan Oseledets †¶
HSE *, Skoltech †, IITP RAS ‡, MIPT §, INM RAS ¶
Russian Federation
Email: kstruminskiy@hse.ru, kruglik@iitp.ru, vetrovd@yandex.ru, i.oseledets@skoltech.ru

***Abstract*—The rapid growth of traffic and number of simultaneously available devices leads to the new challenges in constructing fifth generation wireless networks (5G). To handle with them various schemes of non-orthogonal multiple access (NOMA) were proposed. One of these schemes is Sparse Code Multiple Access (SCMA), which is shown to achieve better link level performance. In order to support SCMA signal decoding channel estimation is needed and sparse Bayesian learning framework may be used to reduce the requirement of pilot overhead. In this paper we propose a modification of sparse Bayesian learning based channel estimation algorithm that is shown to achieve better accuracy of user detection and faster convergence in numerical simulations.**

***Index Terms*—5G; SCMA; channel estimation; sparse Bayesian learning; active user detection**

## I. Introduction

A fifth generation (5G) wireless communication standard, which expected to be commercially used in 2020, includes support of very diverse applications and tremendous number of users as a basic part of IoT concept. It must also support massive connectivity, achieve spectral efficiency and lower latency [1]. Unfortunately, newly launched Long Term Evolution Advanced (LTE-A) networks are not efficient enough to meet all requirements that are imposed to 5G systems, especially in uplink (UL). To deal with them a new Non-Orthogonal Multiple Access (NOMA) scheme – Sparse Code Multiple Access (SCMA) was introduced in [2]. As other NOMA schemes SCMA brings some controllable interference to implement overloading at the cost of increased receiver complexity in order to achieve higher spectral efficiency and massive connectivity [3]. In SCMA systems incoming data from different streams are directly mapped to the codewords from multi-dimensional codebooks. Multiple users select their codebooks and pilots and then transmit their data in preconfigured resource blocks without preliminary request procedures. The main advantage of SCMA over another NOMA schemes is some potential gain of multi-dimensional constellation shaping [4].

The main problem with coherent signal detection is necessity of channel estimation, because demodulation of received signal is possible only after obtaining channel state information. In SCMA, this question is being studied unlike traditional digital telecommunication systems. In [5] blind active user detection with joint message passing algorithm was introduced. Unfortunately, it makes an assumption that SCMA layers share the same time-frequency resource block which isn't true in general. The other approaches – compressive sensing detectors with orthogonal matching pursuit [6] and compressive sampling matching pursuit [7] suffer from severe performance loss in case of further decreasing of number of received pilot resources. This problem arises due to convex relaxation from $l_0$ minimization to $l_1$ minimization [8].

To handle with these problems and to increase overall frequency efficiency a robust active UE detector based on Sparse Bayesian Learning was proposed in [9]. In this article, we modify this approach and achieve better convergence rate which is important in case of tremendous number of simultaneously active users as one of the factors that can increase total decoding speed and as a result decrease latency. Our numerical simulation results show a five time increase of convergence rate in case of 36 potential users with 6 active users and 20 pilots in one fading block. Our algorithm employs a modified iterative scheme for approximate Bayesian channel parameter estimation. This modification was first mentioned in [10] as the one that shows better convergence results in machine learning problems.

The rest of this paper is organized as follows. The system model is presented in Section II. Our modification of SBL detector is presented in Section III. Comparison of different user detectors based on SBL framework performed by numerical simulation is presented in Section IV. This paper is concluded in Section V.

## II. System model

### A. SCMA encoding

An SCMA encoding procedure is defined as mapping from $log_2(M)$ bits to a $K$-dimensional complex codebook of size $M$ [2]. In the uplink transmission each user has its own codebook and its data bits are mapped into $K$-dimensional complex codeword with $N < K$ non-zero entries selected from corresponding codebook. After it these bits are transmitted through $K$ resource elements (REs) (for example Orthogonal Frequency-Division Multiple Access resource elements). In this case one RE is the resource of one subcarrier in Orthogonal Frequency-Division Multiplexing and multiple overlapped SCMA blocks may fit within assigned time-frequency resources [11]. In the uplink transmission scheme with SCMA multiple access is achieved through the sharing of the same time-frequency resources among SCMA layers of multiple active users [12].

Signal from $U$ users that arrived through different channels after the synchronous level multiplexing can be expressed as [2]

$$\vec{y} = \sum_{u=1}^{U} \sqrt{P_u} diag(\vec{h_u}) \vec{x_u} + \vec{n}$$

where $x_u = [x_{u,1}, x_{u,2}, ..., x_{u,K}]^T$ is SCMA codeword transmitted by user $u$, $P_u$ – power of received signal, $\vec{h_u} = [h_{u,1}, h_{u,2}, \dots, h_{u,K}]^T$ – channel vector of user $u$, $\vec{n}$ – Gaussian ambient noise and cell-off interference.

### B. Channel model

In this article we consider system with allocation of SCMA codewords from $N$ users into one resource block. We assume time-invariant channel response within range $t_1$ to $t_2$ and block Rayleigh fading with size of each block equal to $B$. Total amount of fading blocks is $Q$. Within each fading block we assign $N_p$ REs of pilots to identify different users and $N_d$ REs to transmit SCMA codewords ($B = N_p + N_d$). An active user pick its pilot sequence and its corresponding codebook during SCMA UL access.

In one resource block the received pilot vector $\vec{y}$ can be expressed as follows:

$$\vec{y} = [P_1 \ P_2 \ \dots \ P_N] \begin{bmatrix} \vec{c_1} a_1 \\ \vec{c_2} a_2 \\ \vdots \\ \vec{c_N} a_N \end{bmatrix} + \vec{n}$$

where $P_n = \begin{bmatrix} \vec{p}_{n,1} & 0 & \dots & 0 \\ 0 & \vec{p}_{n,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \vec{p}_{n,Q} \end{bmatrix}$ $(1 \leq n \leq N)$

$$\vec{c_n} = [c_{n1} \ c_{n2} \ \dots \ c_{nQ}]^T$$

and $p_{n,q}$–$n$-th pilot sequence of $q$-th fading block, $a_n \in \{0, 1\}$–$n$-th pilot sequence activity indicator, $c_{n,q}$–q-th fading block channel response of active user who picked $n$-th pilot-code, $\vec{n}$ – additive white Gaussian noise.

During transmission we make following additional assumptions:

- Each user is synchronized in symbol level as well as in block level
- Only small part of all available users transmit information simultaneously
- During transmission active user picks a pilot sequence and its corresponding SCMA codebook according to its user index or randomly
- Pilot sequences are assigned without repetitions
- For each user channel response within one fading block is constant
- The size of fading block depends on the channel condition
- Each resourse block is competitively grant-free accessed by multiple users

It should be mentioned that the number of available pilot codes must exceed the number of active users in order to ensure low error rate for user separation.

## III. Channel estimation algorithm

### A. Real-valued formulation and notation

In matrix notatin received pilot vector may be expressed as

$$\vec{y} = P\vec{\theta} + \vec{n}. \tag{1}$$

Here we have to estimate vector $\vec{\theta} = [\vec{c_1} a_1, \vec{c_2} a_2, \dots, \vec{c_N} a_n]^T$ given $\vec{y}$ and $P$.

All the entries of (1) are complex-valued and for the sake of clarity we will reformulate the problem of $\vec{\theta}$ estimation in terms of real-valued vectors and matrices. We use standard embedding to the real vector spaces:

$$\begin{aligned}
\vec{y}_{\mathbb{R}} \leftarrow & [\mathrm{Re}(y_1),\ \mathrm{Im}(y_1),\ \dots,\ \mathrm{Re}(y_D),\ \mathrm{Im}(y_D)]^T \\
\vec{n}_{\mathbb{R}} \leftarrow & [\mathrm{Re}(n_1),\ \mathrm{Im}(n_1),\ \dots,\ \mathrm{Re}(n_D),\ \mathrm{Im}(n_D)]^T \\
\vec{\theta}_{\mathbb{R}} \leftarrow & [\mathrm{Re}(c_{(1,1)} a_1),\ \mathrm{Im}(c_{(1,1)} a_1),\ \dots,\ \mathrm{Re}(c_{(1,Q)} a_1),\ \mathrm{Im}(c_{(1,Q)} a_1), \\
& \dots \\
& \mathrm{Re}(c_{(N,1)} a_N), \mathrm{Im}(c_{(N,1)} a_N), \dots, \mathrm{Re}(c_{(N,Q)} a_N), \mathrm{Im}(c_{(N,Q)} a_N)]^T
\end{aligned}$$

Vector $\vec{\theta}$ contains information about channel coefficients for $Q$ fading blocks for $N$ users. We used double indices to denote its components, the dimensionality of embedded vector $\vec{\theta}_{\mathbb{R}}$ is twice the dimensionality of $\vec{\theta}$. The embedding replaces each complex fading block coefficient with its real and imaginary parts, so we will use double indices to denote $\theta_{\mathbb{R}}$ components where the second integer in index now varies from 1 to $2Q$.

Each component of $P$ is mapped into $2 \times 2$ block in real-valued matrix by the following rule:

$$\begin{aligned}
P_{\mathbb{R},2i,2j}, P_{\mathbb{R},2i+1,2j+1} &\leftarrow \mathrm{Re}(P_{ij}) \\
P_{\mathbb{R},2i+1,2j} &\leftarrow \mathrm{Im}(P_{ij}) \\
P_{\mathbb{R},2i,2j+1} &\leftarrow -\mathrm{Im}(P_{ij})
\end{aligned}$$

From now we will operate only with real-valued vectors, thus with a slight abuse of notation we will use the original symbols to denote real-valued representations of vectors and matrices (e.g. $\vec{y}$ instead of $\vec{y}_{\mathbb{R}}$).

### B. Probabilistic model

We consider the following probabilistic model of pilots transmission for SCMA UL detection:

$$p(\vec{y}, \vec{\theta} | \vec{\gamma}) = p(\vec{y} | \vec{\theta}) p(\vec{\theta} | \vec{\gamma}) \tag{2}$$

This model inherets distribution $p(\vec{y}|\vec{\theta})$ from the channel design and introduces a prior distribution $p(\vec{\theta}|\vec{\gamma})$ in order to perform Bayesian inference in the model.

Here the received vector $\vec{y}$ depends on channel parameter vector $\vec{\theta}$ as follows:

$$p(\vec{y}|\vec{\theta}) = \mathcal{N}(\vec{y} | P\vec{\theta}, diag(\rho, ..., \rho))$$

This distribution both captures linear dependence between $\vec{\theta}$ and $\vec{y}$, and additive Gaussian noise with variance $\rho$. We assume that $\rho$ is a fixed parameter that is known to the receiver.

Prior distribution over $\vec{\theta}$ is parametrized by vector $\vec{\gamma}$ and defined as follows:

$$p(\vec{\theta}|\vec{\gamma}) = \prod_{n=1}^{N}\prod_{q=1}^{2Q} \frac{1}{\sqrt{2\pi\gamma_n}} \exp\left\{\left(-\frac{\theta_{(n,q)}^2}{2\gamma_n}\right)\right\} = \mathcal{N}(\vec{\theta}|0,\Gamma),$$

were $\Gamma = diag(\underbrace{\gamma_1,...,\gamma_1}_{2Q},...,\underbrace{\gamma_N,...,\gamma_N}_{2Q})$.

Note that for each user $n$ channel components $(\theta_{(n,1)},...,\theta_{(n,2Q)})$ have normal distribution with variance parametrized by $\gamma_n$. If $\gamma_n \to 0$, then the components of distribution concentrates at zero. On the other hand, as $\gamma_n$ becomes larger, the distribution becomes less restrictive. Thus $p(\vec{\theta}|\vec{\gamma})$ is capable to model sparse vectors $\vec{\theta}$ with certain parameters $\vec{\gamma}$.

### C. Parameter estimation

A standard approach to infer a model parameter or an unobserved variable with respect to observed data is to compute it's posterior distribution via Bayes rule. Given $\vec{\gamma}$, one can easily compute posterior distribution $p(\vec{\theta}|\vec{y},\vec{\gamma})$ to estimate $\vec{\theta}$. However, parameters $\vec{\gamma}$ are not known, but the estimation performance dramatically depends on their values.

Bayesian approach for point estimates suggests to select parameters that maximize model evidence:

$$\vec{\gamma}_* = arg\max p(\vec{y}|\vec{\gamma})$$

where evidence is defined by the summation rule:

$$p(\vec{y}|\vec{\gamma}) = \int p(\vec{y},\vec{\theta}|\vec{\gamma})p(\vec{\theta}|\vec{\gamma})\mathrm{d}\vec{\theta}.$$

Since the density function $p(\vec{y}|\vec{\gamma})$ is a convolution of two normal distributions, it has the following form:

$$p(\vec{y}|\vec{\gamma}) = \frac{1}{(2\pi)^D}|\Sigma_t|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\vec{y}^T\Sigma_t^{-1}\vec{y}\right),$$

where $\Sigma_t = \frac{1}{2\rho}I + P\Gamma P^T$.

Analytical model evidence maximization leads to the system of non-linear equations, and it is more practical to maximize a lower bound for evidence via EM-algorithm. Consider the following log-evidence representation:

$$\log p(\vec{y}|\gamma) = \mathbb{E}_{q(\vec{\theta})}\log p(\vec{y},\vec{\theta}|\vec{\gamma}) + H(q(\vec{\theta})) + KL(q(\vec{\theta})||p(\vec{\theta}|\vec{y},\vec{\gamma})). \tag{3}$$

Kullback-Leibler divergence $KL(q(\vec{\theta})||p(\vec{\theta}|\vec{y},\vec{\gamma}))$ is always non-negative, thus terms $\mathcal{L}(q(\theta),\gamma) = \mathbb{E}_{q(\vec{\theta})}\log p(\vec{y},\vec{\theta}|\vec{\gamma}) + H(q(\vec{\theta}))$ form a lower bound for log-evidence. EM-algorithm iteratevly maximizes the evidence lower bound by minimizing the third term in 3 and maximizing the first term in 3 is known as EM-algorithm.

On $k$-th E-step the algorithm minimizes the third term in (3) with respect to distribution $q(\vec{\theta})$ by setting it to be equal to

$$\begin{aligned} p(\vec{\theta}|\vec{y},\vec{\gamma}^{(k)}) &= \mathcal{N}(\mu_\theta,\Sigma_\theta) \\ \vec{\mu}_\theta &= \Gamma^{(k)}P^T(P\Gamma^{(k)}P^T + \frac{1}{2\rho}I)^{-1}\vec{y} \\ \Sigma_\theta &= \Gamma^{(k)} - \Gamma^{(k)}P^T(P\Gamma^{(k)}P^T + \frac{1}{2\rho}I)^{-1}P\Gamma^{(k)} \end{aligned}$$

for $\Gamma^{(k)} = diag(\underbrace{\gamma_1^{(k)},...,\gamma_1^{(k)}}_{2Q},...,\underbrace{\gamma_N^{(k)},...,\gamma_N^{(k)}}_{2Q})$.

Indices for $\vec{\mu}_\theta$ and $\Sigma_\theta$ are inherited from indices for $\vec{\theta}$. On the M-step the first term in (3) is maximized with respect to $\vec{\gamma}$. Finding exact maximum of the first term of (3) give us the following set of equations:

$$0 = 2Q - \frac{1}{\gamma_n}\sum_{q=1}^{2Q}(\Sigma_{\theta(n,q)(n,q)} + \mu_{\theta(n,q)}^2), \quad n = 1,...,N. \tag{4}$$

Solving each equation with respect to $\gamma_n$ leads to the following M-step:

$$\gamma_n^{(k+1)} = \frac{\sum_{q=1}^{2Q}\left(\Sigma_{\theta(n,q)(n,q)} + \mu_{\theta(n,q)}^2\right)}{2Q}. \tag{5}$$

This scheme was proposed [9], yet in the original paper a different prior $p(\vec{\theta}|\vec{\gamma})$ was used and the summation in (5) was introduced as a heuristic. It turns out that the original scheme can be represented as EM-algorithm in our probabilistic model.

In this paper we propose to follow the approach of McKay [10] [13] for the M-step, since it is known to converge faster in practice. Firstly, we rearrange the terms in (4) as follows:

$$\frac{1}{\gamma_n}\sum_{q=1}^{2Q}\mu_{\theta(n,q)}^2 = 2Q - \frac{1}{\gamma_n}\sum_{q=1}^{2Q}\Sigma_{\theta(n,q)(n,q)}. \tag{6}$$

Let us consider an iterative scheme where new values of $\gamma_n^{(k+1)}$ are evaluated by substituting all entries of $\gamma_n$ with $\gamma_n^{(k)}$ on the right hand side of 6, substituting $\gamma_n$ with $\gamma_n^{(k+1)}$ on the left hand side of the equation and solving it with respect to $\gamma_n^{(k+1)}$. We obtain the following iterative scheme:

$$\gamma_n^{(k+1)} = \frac{\gamma_n^{(k)}\sum_{q=1}^{2Q}\mu_{\theta(n,q)}^2}{\gamma_n^{(k)}2Q - \sum_{q=1}^{2Q}\Sigma_{\theta(n,q)(n,q)}}. \tag{7}$$

It is clear that fixed points of this scheme are extreme points of the evidence lower bound. This M-step defines EM-algorithm summarized in **Algorithm 1**.

Complexity of the first algorithm was analyzed in [9], our detector has the same complexity $O(K(D^3+(QN)^3)$. Here $K$ is the maximal number of iterations, $D$ is the dimensionality of $y$, $Q$ is number of fading blocks and $N$ is number of potential users.

**Algorithm 1** SBL-Based Active User Detector and Channel Estimator for SCMA UL

**Require:** $\vec{y}$ is a real-valued received vector, $P$ is a real-valued matrix, $\rho$ defines AWGN variance, $\vec{\gamma}^{initial}$ is an initial prior parameter, $\delta$ defines decision threshold, $K$ defines maximal number of iterations

**function** ESTIMATOR($\vec{y}, P, \rho, \vec{\gamma}^{initial}, \delta, K$)
  $\vec{\gamma}^{(1)} \leftarrow \vec{\gamma}^{initial}$
  **for** $k \leftarrow 1$ to $K$ **do**
    $\Gamma^{(k)} \leftarrow diag(\gamma_1^{(k)}, ...\gamma_1^{(k)}, ..., \gamma_N^{(k)}, ..., \gamma_N^{(k)})$
    $\vec{\mu}_\theta \leftarrow \Gamma^{(k)} P^T (P\Gamma^{(k)} P^T + \frac{1}{2\rho} I)^{-1} \vec{y}$
    $\Sigma_\theta \leftarrow \Gamma^{(k)} - \Gamma^{(k)} P^T (P\Gamma^{(k)} P^T + \frac{1}{2\rho} I)^{-1} P\Gamma^{(k)}$
    **for** $n \leftarrow 1$ to $N$ **do**
      $\gamma_n^{(k+1)} \leftarrow \frac{\sum_{q=1}^{2Q} \mu^2_{\theta(n,q)}}{\gamma_n^{(k)} 2Q + \sum_{q=1}^{2Q} \Sigma_{\theta(n,q)(n,q)}}$
      **if** $\gamma_n^{(k+1)} < \delta$ **then**
        $\gamma_n^{(k+1)} \leftarrow 0$
      **end if**
    **end for**
  **end for**
  $\Gamma \leftarrow diag(\gamma_1^{(K)}, ...\gamma_1^{(K)}, ..., \gamma_N^{(K)}, ..., \gamma_N^{(K)})$
  $\hat{\theta} \leftarrow \Gamma P^T (P\Gamma P^T + \frac{1}{2\rho} I)^{-1} \vec{y}$
  **return** $\hat{\theta}$
**end function**

## IV. SIMULATION RESULTS

Finally, simulation results are obtained to compare channel estimation and user detection performance of the proposed algorithm and the sparse Bayesian estimator from [9].

In our experiments we perform link-level simulation for uplink transmission in Rayleigh fading channel. We simulate a SCMA UL system with $N = 36$ potential and 6 active users. The total resource block was divided into $Q = 5$ fading blocks, within each fading block pilot sequences of 20 elemets were assigned to each user.

We used the pilot sequences that are used in demodulated reference signal in LTE systems [14]. We set 6 base pilot sequences to be Zadoff-Chu sequences and then obtain new sequences with cyclic shift of base pilot sequences.

We used Mean Square Error ($MSE = \mathbb{E}\left[\frac{(\vec{\theta}-\theta^*)^T(\vec{\theta}-\theta^*)}{N}\right]$) to measure the quality of channel estimation and User Detection Error Rate ($UDER = \mathbb{E}\left[\frac{\sum_{n=1}^{N}[\hat{a}_n \neq a_{n^*}]}{N}\right]$) to measure the quality of user detection. Also we analyzed the convergence of the algorithm with respect to the maximal number of iterations $K$.

Figure 1 shows the convergence of mean squared error for diffefent channel conditions. The simulation results show that the proposed scheme almost reaches the minimum of mean squared error several times faster than the original algorithm. On the other hand, the proposed scheme is less precise at low signal-to-noise ratios. Figure 2 presents the convergence of user detection error rate for different channel conditions. It can be seen that the proposed scheme has both lower detection error rate and converges faster.

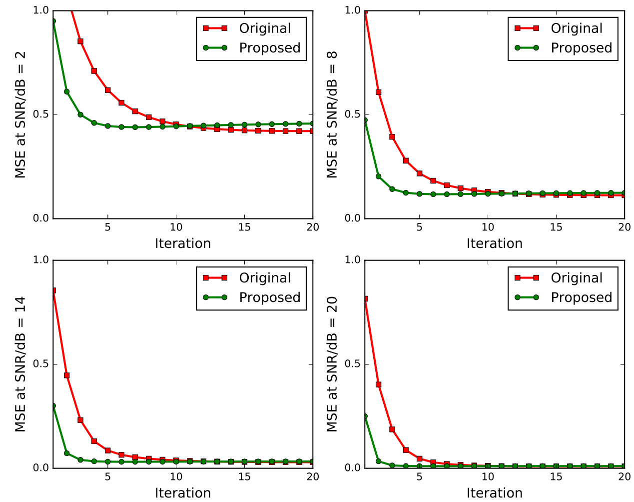


Fig. 1. Convergence of the Mean Square Error of channel estimation for different noise ratios

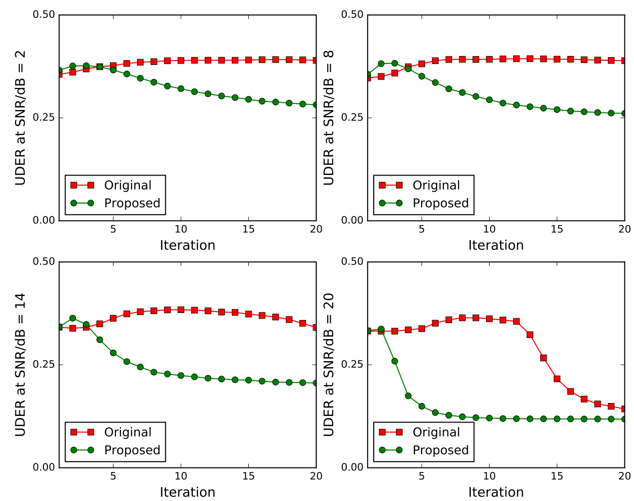


Fig. 2. Convergence of the user detection error rate for different noise ratios

Figure 3 demonstrates error rate after $K = 20$ iteration steps for the original scheme and $K = 10, 20$ iteration steps for the proposed scheme with different channel conditions. The simulation results show that the proposed scheme maintains user detection improvement even for fewer number of iterations.

## V. CONCLUSION

In this paper the improvement of active user detector based on the theory of sparse Bayesian learning is proposed. This improvement is inspired by an empirical fact about EM-algorithm performance in the probabilistic model that we use. Simulation results are provided to substantiate the performance improvement of the detector and practical value of the proposed scheme. Although no theoretical explanations for this behavior are presented in literature, such explanations

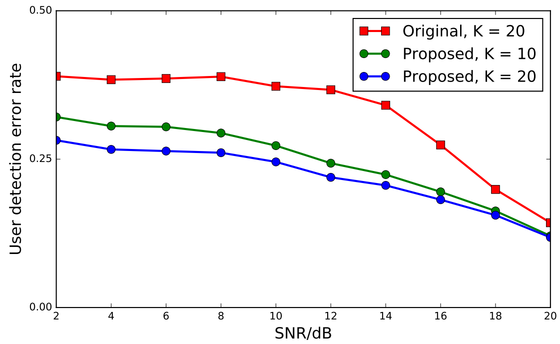


Fig. 3. User detection error rate after $K$ iterations

and adaptation of faster or approximate procedures to replace matrix inversion in the algorithm are the directions for future research.

## Acknowledgment

The research of Stanislav Kruglik was supported by the Russian Foundation for Basic Research (project 16-01-00716). The research of Dmitry Vetrov was supported by the Russian Foundation for Basic Research (project 15-31-20596 (mol-a-ved)). The research of Ivan Oseledets was supported by the Russian Science Foundation (project 14-11-00659-a)

## References

[1] Chih-Lin I et all., *New Paradigm of 5G Wireless Internet*. IEEE Journal on Selected Areas in Communications, volume 34, issue 3, pp. 474 -482, 2016.
[2] Hosein Nikopour and Hadi Baligh, *Sparse code multiple access*. IEEE 24th PIMRC, pp. 332-336, 2013.
[3] B. Wang, K. Wang, Z. Lu, T. Xie and J. Quan, *Comparation study of non-orthogonal multiple access schemes for 5G*. BMSB, pp. 1-5, 2015.
[4] L. Dai, B. Wang, Y. Yuan, S. Han, C. I and Z. Wang, *Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research thrends*. IEEE Commun. Mag., volume 53, issue 9, pp. 74-81, 2015.
[5] Nikopour H. ; Baligh H. Bayesteh, A; Yi. *Blind detection of scma for uplink grant-free multiple-access*. In Wireless Communications Systems (ISWCS), 2014 11th International Symposium on, pages 853-857. IEEE, 2014.
[6] D.L. Donoho. *Compressed sensing*. Information Theory, IEEE Transactions on, 52(4):1289-1306, 2006.
[7] Tropp, J.A. ; Gilbert. *Signal recovery from random measurements via orthogonal matching pursuit*. In Information Theory, IEEE Transactions on, pages 4665-4666. IEEE, 2007.
[8] D. Needell; J.A. Tropp. *Cosamp: Iterative signal recovery from incomplete and inaccurate samples*. Applied and Computational Harmonic Analysis, 26(3); 301-321, 2008.
[9] Yufeng Wang; Shidong Zhou ; Limin Xiao ; Xiujun Zhang ; Jin Lian. *Sparse Bayesian learning based user detection and channel estimation for SCMA uplink systems*.Wireless Communications & Signal Processing (WCSP), 2015 International Conference on, pages 1-5, 2015.
[10] David J. C. MacKay. *Bayesian Interpolation*. Neural Computation 4, 415-447, 1991.
[11] H. ; Yi E. Au, K. ;Liqing Zhang ; Nikopour. *Uplink contention based scma for 5g radio access*. In Globecom Workshops (GC Wkshps), 2014, pages 900-905. IEEE, 2014.
[12] H. ; Bayesteh A. ; Baligh H. Taherzadeh, M. ; Nikopour. *Scma codebook design*. In Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th, pages 1-5. IEEE, 2014.
[13] Tipping, M. E. *Sparse Bayesian learning and the relevance vector machine*. Journal of Machine Learning Research 1, 211-244, 2001.
[14] 3GPP official website, LTE official specification, Web: http://www.3gpp.org/specifications.