RSMA Receiver

Sergei Semenov
HiSilicon/Huawei
Lund, Sweden
e-mail: sergei.semenov@huawei.com

***Abstract*—In this paper, we propose a new algorithm for the detection of Resource Spread Multiple-Access (RSMA) signal. The proposed algorithm provides a significant performance gain in comparison with other algorithms for the high spectral efficiency case with affordable complexity.**

Keywords-RSMA; NOMA; OLMA; message passing algorithm; projection based IC.

I. INTRODUCTION

Resource spread multiple access (RSMA) is an overloaded multiple-access (OLMA) strategy, which is closely related to code-division multiple access (CDMA) [7] and interleave-division multiple-access (IDMA) [8]. RSMA was proposed in [1][2] as a candidate for new radio (NR) uplink (UL) multiple access. Current assumption for RSMA receiver implementation is that very low rate channel coding is used in the system. In this case, quite simple approaches can be used for the RSMA receiver implementation. In [3], the use of two types of receivers was proposed:

1. Match filter (MF): Each layer descrambles and de-spreads the signal before passing it to the decoder. Detection is done by the Hermitian transpose of spreading/scrambling sequence matrix, which can be viewed as a match filter.
2. MF+successive interference cancellation (SIC): Once a packet is decoded, then its waveform can be cancelled from the received waveform. The receiver re-attempts to decode unsuccessful packets. The iteration stops when no new packet needs to be decoded.

This type of receiver can be used in Ultra Reliable Low Latency Communications (URLLC) scenario and to some extent in Massive Machine Type Communications (mMTC) scenario but it hardly can be used for Enhanced Mobile Broadband (eMBB) transmission since in this case, low spectral efficiency (SE) caused by usage of very low code rate is a drawback. Joint maximum likelihood (ML) detection provides the optimum performance but the complexity of this type of the receiver grows exponentially with the number of users.

The message passing algorithm (MPA) can be used in the receiver for non-orthogonal MA employing low-density signatures like LDSMA [4] or SCMA and it provides good performance with affordable complexity. However, it cannot be applied directly to RSMA detection since RSMA does not use low-density signatures, i.e., the number of users colliding over one resource element (RE) is equal to the number of all users.

In this paper, we propose to use a combination of MPA and projection based interference cancellation (IC) for RSMA signal detection.

The rest of this paper is organized as follows. Section II describes the application of MPA for OLMA detection and the projection based IC. Section III describes the hybrid receiver combining MPA and projection based IC for RSMA detection. Simulation results are represented in Section IV. The conclusions close the article.

II. MPA FOR OLMA DETECTION AND PROJECTION BASED IC

A. MPA for LDSMA signal

The application of MPA for the detection of the LDSMA signal is described in [4]. This application is based on the structure of the LDSMA signal. Users in LDSMA share the available REs in such a way that only limited number of users can transmit over a particular RE.

The structure of this system can be described with the help of indicator matrix $\mathbf{F}_{N\times K}$ defining the signals of K users spreading their signals over N REs. An example of indicator matrix for N = 12 and K = 16 is presented in Figure 1.

$$\mathbf{F}_{12\times16} = \begin{bmatrix} 0&0&1&0&0&0&0&1&0&1&0&0&1&0&0&0 \\ 0&0&0&1&0&0&1&0&0&0&0&0&1&0&1&0 \\ 0&1&0&0&1&0&1&0&0&1&0&0&0&0&0&0 \\ 0&0&0&1&0&1&0&0&0&0&1&0&0&1&0&0 \\ 0&0&1&0&0&0&0&0&1&0&0&0&0&0&1&1 \\ 1&0&0&0&0&1&0&0&0&1&0&0&0&0&1&0 \\ 0&0&0&1&0&0&0&1&1&0&0&1&0&0&0&0 \\ 0&1&0&0&0&1&0&0&0&0&0&1&0&0&0&1 \\ 1&0&0&0&0&0&1&0&0&0&0&0&0&1&0&1 \\ 0&0&1&0&1&0&0&0&0&0&0&1&0&1&0&0 \\ 0&1&0&0&0&0&0&0&1&0&1&0&1&0&0&0 \\ 1&0&0&0&1&0&0&1&0&0&1&0&0&0&0&0 \end{bmatrix}$$

Figure 1. Indicator matrix, $\boldsymbol{K = 16, N = 12, d_u = 3}$ and $\boldsymbol{d_c = 4}$. [4].

The position of ones in the n^{th} row of the indicator matrix $\mathbf{F}_{N\times K}$ denotes the set of users who contribute their data at the nth symbol, while its k^{th} column represents the set of symbols over which the user spreads his/her data. The maximum number of ones in each column d_u indicates the maximum number of nonzero spread symbols, which can be located for each user among the N possible time-frequency resources. It is also clearly seen that each spread symbol will collide in the

 ISBN: 978-1-61208-562-3

channel with d_c (maximum number of ones in row) symbols from other users.

If the indicator matrix has the same number of ones in each column, i.e., d_u and also the same number of ones in each row, i.e., d_c, but is not necessarily equal to d_u, then the structure is called regular indicator matrix, otherwise, it is called irregular indicator matrix.

As can be seen from the description of the indicator matrix it corresponds to the description of the parity-check matrix of the LDPC code. Due to this fact, the idea of MPA used for decoding LDPC codes is applicable for LDSMA signal detection.

An indicator matrix can be represented by a bidirectional bipartite factor graph. In Figure 2, the graph for the LDPC matrix shown in Figure 1 is illustrated.

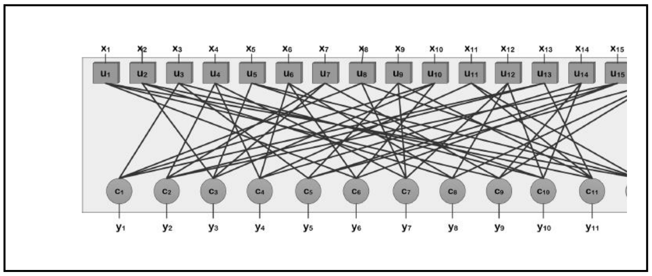

Figure 2. Factor graph representation of the indicator matrix shown in Figure 1 [4].

In this graph, the upper (variable) nodes $\{u_k\}, k = 1, \ldots, K$ are connected to the K user transmitted symbols; the lower (function) nodes $\{c_n\}, n = 1, \ldots, N$ represent N REs carrying the encoded information and being connected to the observations at these REs $\{y_n\}, n = 1, \ldots, N$; the edges between indicate which REs are occupied by a user. From this graph, we find that there is an edge between pair (c_i, u_j) if and only if the matrix element F_{ij} is nonzero. We see also that each node c_n is connected to d_c nodes and each node u_k is connected to d_u nodes.

Let edge $e_{n,k}$ be the edge that connects a function node c_n to a variable node u_k. At the function node c_n, the local channel observation at the corresponding RE y_n is made and is given by

$$p(y_n|\mathbf{x}^{[n]}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2\sigma^2}\left\|y_n - \mathbf{g}^{[n]^T}\mathbf{x}^{[n]}\right\|^2\right), \tag{1}$$

where $\mathbf{g}^{[n]}$ and $\mathbf{x}^{[n]}$ is the set of channel coefficients and the set of transmitted symbols corresponding to non-zero elements in the nth row of the indicator matrix, i.e., corresponding to the users contributing to the nth RE.

Each message being exchanged must be in the form of a vector of size $|\mathcal{A}|$ comprising the reliability values for each of the possible values taken from the symbol constellation alphabet $\mathcal{A}$. Note that, the message must be normalized such that the sum of all probability values for all alphabet symbols is one. LLRs can be used instead of probabilities. Then, the size of message is $|\mathcal{A}| - 1$.

Let symbol y_n be the symbol of observation. The message being sent from function node c_n onto edge $e_{n,k}$ is the product of the messages received from edges $e_{l,n}$ ($l \in \varepsilon_n \backslash k$, where ε_n is the set of d_c variable nodes connected to the function node c_n defined by the nth row of index matrix $\mathbf{F}_{N\times K}$ (variable node u_k must be excluded from this set)) with the local function at c_n and being summarized for the variable associated with the edge, i.e., x_k. Similarly, the variable node u_k will send onto edge $e_{n,k}$ a message, which is the product of the messages received from edges $e_{m,k}$ ($m \in \xi_k \backslash n$, where ξ_k is the set of d_u function nodes connected to the variable node u_k defined by the k^{th} column of index matrix $\mathbf{F}_{N\times K}$ (function node c_n must be excluded from this set)).

Let $\mu_{c_n \leftarrow u_k}$ and $\mu_{c_n \rightarrow u_k}$ be the message sent along edge $e_{n,k}$ from variable node u_k and function node c_n, respectively. The message $\mu_{c_n \leftarrow u_k}$ gives an updated inference of x_k based on the observation taken at symbols y_m, $m \in \xi_k \backslash n$:

$$\mu_{c_n \leftarrow u_k}(j) = \log\frac{P_{ext,n}(x_k = a_j)}{P_{ext,n}(x_k = a_0)} = \sum_{m \in \xi_k \backslash n} \mu_{c_m \rightarrow u_k}(j), \quad j = 1, \ldots, |\mathcal{A}| - 1, \tag{2}$$

where $a_j \in \mathcal{A}$ is the corresponding element of the constellation alphabet $\mathcal{A}$.

Appending the set of equations (2) with the additional restriction

$$\sum_{j=0}^{|\mathcal{A}|-1} \lambda_{n,k} P_{ext,n}(x_k = a_j) = 1, \tag{3}$$

where $\lambda_{n,k}$ is a normalizing coefficient, and solving this set of equations, we obtain

$$P_{ext,n}(x_k = a_j) = \frac{\exp\left(\mu_{c_n \leftarrow u_k}(j)\right)}{\lambda_{n,k}\left(\sum_{i=0}^{|\mathcal{A}|-1}\exp\left(\mu_{c_n \leftarrow u_k}(i)\right)\right)}, \quad \mu_{c_n \leftarrow u_k}(0) = 0. \tag{4}$$

The denominator in (4) is constant for all $P_{ext,n}(x_k = a_j)$, $j = 0, \ldots, |\mathcal{A}| - 1$. Then (4) can be simplified to

$$P_{ext,n}(x_k = a_j) = \exp\left(\lambda'_{n,k}\mu_{c_n \leftarrow u_k}(j)\right), \quad \mu_{c_n \leftarrow u_k}(0) = 0. \tag{5}$$

where $\lambda'_{n,k}$ is the normalization coefficient and is chosen to satisfy (3), $\lambda'_{n,k} = -\log\lambda_{n,k}$.

At the function node c_n, the inference of x_k is updated and is given by

$$\mu_{c_n \rightarrow u_k}(j) = \log\frac{p_{ext,n}(x_k = a_j | y_n, \mathbf{x}^{[n]} \backslash x_k)}{p_{ext,n}(x_k = a_0 | y_n, \mathbf{x}^{[n]} \backslash x_k)}, \quad j = 1, \ldots, |\mathcal{A}| - 1. \tag{6}$$

Applying Bayes' rule to the nominator and the denominator of (6) and taking into account that the a priori pmf of x_k should not be included in the computation of a posteriori pmf of x_k, we obtain

$$\begin{aligned} &\mu_{c_n \to u_k}(j) \\ &= \log \frac{p_{ext,n}(x_k = a_j | y_n, \mathbf{x}^{[n]} \backslash x_k)}{p_{ext,n}(x_k = a_0 | y_n, \mathbf{x}^{[n]} \backslash x_k)} \\ &= \log \frac{p_{ext,n}(y_n | \mathbf{x}^{[n]}, x_k = a_j) P(\mathbf{x}^{[n]} \backslash x_k)}{p_{ext,n}(y_n | \mathbf{x}^{[n]}, x_k = a_0) P(\mathbf{x}^{[n]} \backslash x_k)}, \\ &j = 1, \dots, |\mathcal{A}| - 1. \end{aligned} \tag{7}$$

Combining (1), (2), and (5) into (7), we can write a complete message being sent from the function node c_n to the variable node u_k onto the edge $e_{n,k}$ as follows:

$$\begin{aligned} &\mu_{c_n \to u_k}(j) \\ &= \log \frac{\sum_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_{c}}, \\ x_k = a_j}} \exp\left(\sum_{l \in \varepsilon_n \backslash k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) - \frac{1}{2\sigma^2} \left\| y_n - {\mathbf{g}^{[n]}}^T \mathbf{x}^{[n]} \right\|^2 \right)}{\sum_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_{c}}, \\ x_k = a_0}} \exp\left(\sum_{l \in \varepsilon_n \backslash k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) - \frac{1}{2\sigma^2} \left\| y_n - {\mathbf{g}^{[n]}}^T \mathbf{x}^{[n]} \right\|^2 \right)} \\ &= \max_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_{c}}, \\ x_k = a_j}}{}^{*} \left(\sum_{l \in \varepsilon_n \backslash k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) - \frac{1}{2\sigma^2} \left\| y_n - {\mathbf{g}^{[n]}}^T \mathbf{x}^{[n]} \right\|^2 \right) \\ &- \max_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_{c}}, \\ x_k = a_0}}{}^{*} \left(\sum_{l \in \varepsilon_n \backslash k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) - \frac{1}{2\sigma^2} \left\| y_n - {\mathbf{g}^{[n]}}^T \mathbf{x}^{[n]} \right\|^2 \right), \\ &j = 1, \dots, |\mathcal{A}| - 1, \end{aligned} \tag{8}$$

where $\mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}$ denotes the message from variable node u_l to function node c_n corresponding to vector $\mathbf{x}^{[n]}$ and function $\max^*$ is defined as

$$\max{}^*(a, b) = \log(e^a + e^b) = \max(a, b) + \log\left(1 + e^{-|a-b|}\right). \tag{9}$$

After the message arriving to variable nodes u_k, $k = 1, \dots, K$ have converged or the maximum iteration number has been reached, the variable nodes will use all messages received from all connected edges to calculate the final estimated inference for symbol x_k, because now we are calculating the symbol estimate rather than an extrinsic information, and, except this detail, this is done in the same way as in (2)

$$\begin{aligned} P(x_k = a_j) &= \exp\left(\sum_{n \in \xi_k} \lambda'_{n,k} \mu_{c_n \to u_k}(j) \right), \\ \mu_{c_n \to u_k}(0) &= 0. \end{aligned} \tag{10}$$

B. Projection based IC

Another method used in hybrid RSMA receiver is the interference cancellation based on projection techniques.

Consider the system model

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n}, \tag{11}$$

where $\mathbf{y}$ is the received signal, $\mathbf{x}$ is the vector, composed of the transmitted symbols of all users and $\mathbf{n}$ is the noise vector. Matrix $\mathbf{G}$ is the generalized channel matrix including both channel coefficients and user signatures.

We denote the part of vector x corresponding to signal of interest by $\mathbf{x}_\mathrm{T}$ and the other part of vector $\mathbf{x}$ corresponding to the interference by $\mathbf{x}_\mathrm{Q}$. Then, the matrix $\mathbf{G}$ can also be split into two parts corresponding to signal of interest and interference. Without loss of generality, we can assume that the first Δ elements of vector $\mathbf{x}$ and correspondingly first Δ columns of matrix $\mathbf{G}$ correspond to the signal of interest. Then, the matrix $\mathbf{G}$ can be represented as follows:

$$\mathbf{G} = [\mathbf{T}, \mathbf{Q}]. \tag{12}$$

And (11) can be represented as

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n} = \mathbf{T}\mathbf{x}_T + \mathbf{Q}\mathbf{x}_Q + \mathbf{n}. \tag{13}$$

The simplest solution to (13) wrt $\mathbf{x}_T$ is to apply a ZF type receiver. It can be done with the help of the following correlation operation

$$\begin{aligned} (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{y} &= (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H(\mathbf{T}\mathbf{x}_T + \mathbf{Q}\mathbf{x}_Q + \mathbf{n}) \\ &= (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{T}\mathbf{x}_T \\ &+ (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{Q}\mathbf{x}_Q + (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{n} \\ &= \mathbf{x}_T + (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{Q}\mathbf{x}_Q \\ &+ (\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{n}. \end{aligned} \tag{14}$$

The main problem with the ZF type receiver solution (14) is the term corresponding to interference $(\mathbf{T}^H\mathbf{T})^{-1}\mathbf{T}^H\mathbf{Q}\mathbf{x}_Q$.

The main idea of interference mitigation is to project the received signal y onto a vector subspace that is orthogonal to the interference vector subspace (the subspace spanned by the columns of matrix $\mathbf{Q}$).

The projection is a linear transformation $\mathbf{P}$ from a vector space to itself such that it is idempotent, i.e., $\mathbf{P}^2 = \mathbf{P}$.

Let W be an underlying vector space. Suppose the subspaces U and V are the range and null space of $\mathbf{P}$ respectively. Then, the projection has these basic properties:

1. $\mathbf{P}$ is the identity operator $\mathbf{I}$ on U: $\forall \mathbf{x} \in U: \quad \mathbf{P}\mathbf{x} = \mathbf{x}$.
2. W is a direct sum $W = U \oplus V$. This means that every vector $\mathbf{x} \in W$ may be decomposed uniquely in the manner $\mathbf{x} = \mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in U$ and $\mathbf{v} \in V$. The decomposition is given by
   $\mathbf{u} = \mathbf{P}\mathbf{x}, \ \ \mathbf{v} = \mathbf{x} - \mathbf{P}\mathbf{x} = (\mathbf{I} - \mathbf{P})\mathbf{x}.$

If the range U and the null space V are orthogonal subspaces, the projection $\mathbf{P}$ is an orthogonal projection. For orthogonal projection matrix $\mathbf{P}$ is Hermitian matrix, i.e., $\mathbf{P} = \mathbf{P}^H$. If $\mathbf{u}_1, \dots, \mathbf{u}_k$ is a basis of U, and $\mathbf{A}$ is the matrix with these vectors as columns, then the projection is

$$\mathbf{P}_\mathbf{A} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H. \tag{15}$$

Then the projection onto interference vector space, i.e., vector space spanned by columns of $\mathbf{Q}$ is defined by the matrix

$$\mathbf{P}_\mathbf{Q} = \mathbf{Q}(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H, \tag{16}$$

and, correspondingly, the projection onto null vector space is defined by matrix

 ISBN: 978-1-61208-562-3

$$\mathbf{P}_{\mathbf{Q}}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{Q}} = \mathbf{I} - \mathbf{Q}(\mathbf{Q}^{H}\mathbf{Q})^{-1}\mathbf{Q}^{H}. \tag{17}$$

After projecting the received signal $\mathbf{y}$ onto the null space of $\mathbf{Q}$, we obtain:

$$\begin{aligned} \mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{y} &= \left(\mathbf{I} - \mathbf{Q}(\mathbf{Q}^{H}\mathbf{Q})^{-1}\mathbf{Q}^{H}\right)\mathbf{y} \\ &= \mathbf{T}\mathbf{x}_{T} + \mathbf{Q}\mathbf{x}_{Q} + \mathbf{n} - \mathbf{Q}(\mathbf{Q}^{H}\mathbf{Q})^{-1}\mathbf{Q}^{H}\mathbf{T}\mathbf{x}_{T} - \underline{\underline{\mathbf{Q}(\mathbf{Q}^{H}\mathbf{Q})^{-1}\mathbf{Q}^{H}\mathbf{Q}}}\mathbf{x}_{Q} \\ &- \mathbf{Q}(\mathbf{Q}^{H}\mathbf{Q})^{-1}\mathbf{Q}^{H}\mathbf{n} = \mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{T}\mathbf{x}_{T} + \mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{n}. \end{aligned} \tag{18}$$

Denote vector $\mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{y}$ as $\bar{\mathbf{y}}$, vector $\mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{n}$ as $\bar{\mathbf{n}}$ and matrix $\mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{T}$ as $\bar{\mathbf{G}}$. Then (18) can be written as

$$\bar{\mathbf{y}} = \bar{\mathbf{G}}\mathbf{x}_{T} + \bar{\mathbf{n}}. \tag{19}$$

Then, any appropriate method can be used to solve (19). For example, in [5] [6], RAKE receiver is used to solve (19).

III. HYBRID RSMA RECEIVER

It is not possible to apply MPA directly to RSMA signal detection since the indicator matrix corresponding to RSMA would be a unit matrix (matrix consisting of ones). In this case, the MPA cannot provide good results since corresponding graph contains cycles of length 4. Moreover, the complexity of the algorithm grows exponentially with the number of users.

The drawback of the interference cancellation based on orthogonal projection is that if the number of interferers is high the attempt to map the received signal onto the null space of the interference leads to the distortion of the signal of interest as well.

In this paper, we propose to combine these two methods to create a hybrid receiver of the RSMA signal.

As it was mentioned above, since in RSMA all users collide over each RE, the indicator matrix for RSMA signal comprises unit matrix of dimension *N*x*K*, where *N* is the number of REs over which users are spreading their signals, and *K* is the number of users ($N < K$).

We propose to choose some sparse matrix of size *N*x*K* with good properties (e.g., the corresponding graph should not contain cycles of short lengths and stopping sets) and use it as an indicator matrix. A good way to construct such a matrix could be to use a combinatorial design.

Then, the general MPA described in equations (1)-(10) can be applied. However, before applying calculations corresponding to function nodes (8) the interference coming from interferers designated in the corresponding rows of the indicator matrix **F** must be nullified.

For example, if the RSMA system with 16 users spreading the signals over 12 REs is considered, and matrix **F** represented in Figure 1 is chosen as an indicator matrix, before calculating messages from the function node 1 corresponding to 1st row of matrix **F**, the interference from users 1, 2, 4, 5, 6, 7, 9, 11, 12, 14, 15, 16 should be nullified. And before calculating messages from the function node 2 the interference from users 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 14, 16 must be cancelled.

This step can be done with the help of projection based interference cancellation. We propose before calculating messages from the function node *n*, to collect the signatures of users corresponding to zeros in n^{th} row of matrix **F**, in matrix of interference $\mathbf{Q}$ and projecting the received signal $\mathbf{y}$ onto the null space of $\mathbf{Q}$ like it is done in (18). However, we propose to avoid excessive noise enhancement and take the noise amplification into account when projecting the received signal onto the null space of $\mathbf{Q}$. Due to this, the projection is slightly modified in accordance with MMSE solution:

$$\mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{y} = \left(\mathbf{I} - \mathbf{Q}(\mathbf{Q}^{H}\mathbf{Q} + \sigma_{0}^{2}\mathbf{I})^{-1}\mathbf{Q}^{H}\right)\mathbf{y}\ , \tag{20}$$

where σ_{0}^{2} is the noise variance.

Then calculation of message from the function node *n* can be done in accordance with (7)-(8) but in space skewed by the projection $\mathbf{P}_{\mathbf{Q}}^{\perp}$, i.e., vector $\mathbf{y}$ in (7)-(8) should be substituted by skewed vector $\bar{\mathbf{y}} = \mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{y}$ and matrix $\mathbf{G}$ in (7)-(8) should be substituted by skewed matrix $\bar{\mathbf{G}} = \mathbf{P}_{\mathbf{Q}}^{\perp}\mathbf{T}$.

The number of users to be cancelled should not exceed *N*. Otherwise matrix $(\mathbf{Q}^{H}\mathbf{Q} + \sigma_{0}^{2}\mathbf{I})$ can be singular.

Taking into account the fact that in RSMA signal all users collide over all REs we can improve the calculation of (7)-(8) by calculating (7) simultaneously for a few components of vector $\bar{\mathbf{y}}$ rather than for just one component with index *n* like it is done in (7). Now, the calculations (7)-(8) are done for all components of vector $\bar{\mathbf{y}}$ corresponding to non-zero elements of column *k* in indicator matrix **F** except the component with index *n*. Actually, it means that the component of vector $\bar{\mathbf{y}}$ with index *n* in (7) should be substituted by vector $\bar{\mathbf{y}}^{[n]}$, where vector $\bar{\mathbf{y}}^{[n]}$ comprises vector $\bar{\mathbf{y}}$ with zeros at the same positions where zeros are located in column *k* in indicator matrix **F** , and one more zero is located in the component with index *n*. And vector $\mathbf{g}^{[n]}$ in (8) should be substituted by matrix $\bar{\mathbf{G}}^{[n]}$. The corresponding LLRs should be summed up. Then this can be written as it is shown in (21).

In fact, calculations in (21) can be considered as a joint ML detection of users corresponding to non-zero elements in row *n* of indicator matrix **F**. If we consider the implementation of the receiver in vector processor, the complexity of calculation (21) is the same as the complexity of calculation (8). And the calculation of message going from the function node in the form shown in (21) provides performance gain. The main complexity of the detector is contributed by calculation of (21), i.e., by the joint ML detection of users. As it was mentioned above, the complexity of the receiver grows exponentially with number of users to be detected jointly with the help of ML algorithm. Due to the sparsity of the indicator matrix **F**, the number of users to be detected jointly can be set to be constant or can grow very slowly. If the number of users to be detected jointly is set to be constant, the complexity of the receiver can be considered to be polynomial.

$$
\begin{aligned}
&\mu_{c_n \to u_k}(j) \\
&= \log \frac{\sum_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_c},\\ x_k = a_j}} \exp\left(\sum_{l \in \varepsilon_n \setminus k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) - \frac{1}{2\sigma^2} \left\| \bar{\mathbf{y}}^{[n]} - \bar{\mathbf{G}}^{[n]^T}\right.\right.}{\sum_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_c},\\ x_k = a_0}} \exp\left(\sum_{l \in \varepsilon_n \setminus k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) - \frac{1}{2\sigma^2} \left\| \bar{\mathbf{y}}^{[n]} - \bar{\mathbf{G}}^{[n]^T}\right.\right.} \\
&= \max_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_c},\\ x_k = a_j}}{}^{*} \left(\sum_{l \in \varepsilon_n \setminus k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) \right. \\
&\left. - \frac{1}{2\sigma^2} \left\| \bar{\mathbf{y}}^{[n]} - \bar{\mathbf{G}}^{[n]^T} \mathbf{x}^{[n]} \right\|^2 \right) \\
&- \max_{\substack{\mathbf{x}^{[n]} \in \mathcal{A}^{d_c},\\ x_k = a_0}}{}^{*} \left(\sum_{l \in \varepsilon_n \setminus k} \lambda'_{n,l} \mu_{c_n \leftarrow u_l}^{[\mathbf{x}^{[n]}]}(j) \right. \\
&\left. - \frac{1}{2\sigma^2} \left\| \bar{\mathbf{y}}^{[n]} - \bar{\mathbf{G}}^{[n]^T} \mathbf{x}^{[n]} \right\|^2 \right), \qquad j = 1, \dots, |\mathcal{A}| - 1,
\end{aligned}
\tag{21}
$$

The indicator matrix **F** should be chosen in such a way that at least in one row the set of zero elements corresponds to the weakest users if the information about signal power of different users is available at the receiver.

If turbo-equalization processing is used, i.e., detector and decoder exchange the extrinsic information in a few iterations, it is better to change the indicator matrix **F** for each iteration allowing different users being cancelled with the help of projection based method for the same row of indicator matrix in different iterations. This step increases the channel diversity for different iterations. The simple way to obtain the new indicator matrix with the same good properties as the initial indicator matrix is to generate the new indicator matrix by permutation of columns of the initial indicator matrix.

IV. SIMULATION RESULTS

The simulation results are represented in Figure 3. **Simulation results for 6 users over 4REs. QPSK, Code Rate = 1/3.**. Simulations were done for 6 users occupying 4 RBs, meaning the overloading factor was 150%. The proposed algorithm, designated in plots by MPA+Proj is compared with Successive Interference Cancellation (SIC), and partial ML algorithms designated by ML3 and ML1. In partial ML algorithm, M users ($M < K$) are detected jointly while interference from other users is considered as a noise. This procedure is repeated a few times until all users are detected. In simulation, each user was detected at least twice. Then, the overlapping estimates were combined. In partial ML algorithm, the decoders' output is not used for the interference cancellation. SIC receiver choses the most powerful user and decodes it first. Then, the bit estimates of this user are mapped to symbol estimates and canceled from the signal taking into account the channel estimates. Next, the same procedure is applied to the next user with highest power. At first iteration, SIC receiver uses partial ML detection, i.e., the most powerful user is detected jointly with other M-1 users ($M < K$). Starting from the second iteration, the interference cancellation procedure relies on decoders' output only. Since the ML detection is used at first iteration only the SIC receiver is vulnerable to the error propagation. Due to this reason only extrinsic information is used for IC.

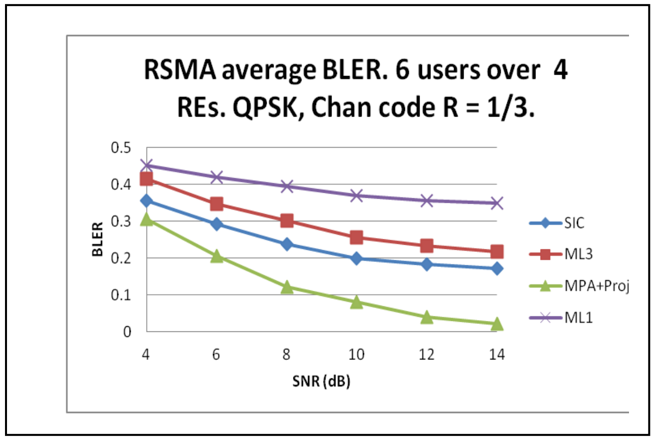

Figure 3. Simulation results for 6 users over 4REs. QPSK, Code Rate = 1/3.

As can be seen from the plots in Figures 3-4, the proposed algorithm provides very high gain for the case of high spectral efficiency. For low code rates (low SE) the gain provided by the hybrid receiver decreases since in this case, all other algorithms start to work quite well.

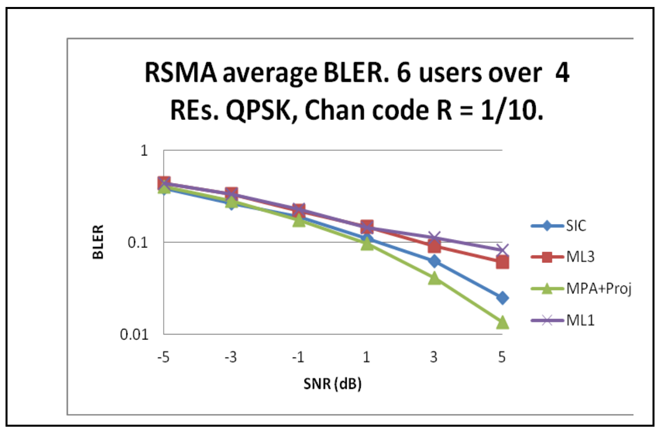

Figure 4. Simulation results for 6 users over 4REs. QPSK, Code Rate = 1/10.

V. CONCLUSION

In this paper, a new hybrid algorithm to detect the RSMA signal is proposed. The proposed algorithm combines MPA and projection based IC and provides significant performance gain in comparison with other algorithms in the case of high SE.

REFERENCES

[1] 3GPP R1-163510 "Candidate NR Multiple Access Schemes," Qualcomm Inc., Busan, Korea, 11th – 15th April 2016.

[2] 3GPP R1-164688 "RSMA," Qualcomm Inc., Nanjing,China, 23rd – 27th May 2016.

[3] 3GPP R1-164689 “RSMA and SCMA comparison,” Qualcomm Inc., Nanjing,China, 23rd – 27th May 2016.

[4] R. Hoshyar, F. P. Wathan, “Novel Low-Density Signature for Synchronous CDMA System Over AWGN Channel,” IEE Trans. On Signal Proc., Vol 56, No 4, April 2008.

[5] J. Thomas, W. Kober, E. Olson, R. Krumvieda, "Interference cancellation in a signal", WO 03/030440 A1, 2001.

[6] W. Kober, J. Thomas, M. Vis, "Rake receiver for spread spectrum signal demodulation", US 2002/0090025 A1, 2001.

[7] G. Miao, J. Zander, K. W. Sung, B. Slimane, “Fundamentals of Mobile Data Networks,” Cambridge University Press, 2016.

[8] Y. Chen, F. Schaich, T. Wild, "Multiple Access and Waveforms for 5G: IDMA and Universal Filtered Multi-Carrier", Vehicular Technology Conference (VTC Spring) 2014 IEEE 79th, pp. 1-5, 2014.

 ISBN: 978-1-61208-562-3