A Simple Message Passing Detector Based on QR-Decomposition for MIMO Systems

Akihide David Shigyo* and Koji Ishibashi†
Advanced Wireless & Communication Research Center (AWCC), The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan
Email: *shigyo@awcc.uec.ac.jp, †koji@ieee.org

***Abstract*—This paper proposes a simple message passing detector with QR-decomposition for multi-input multi-output (MIMO) systems. Our message passing algorithm exploits the structure of QR-decomposed received signals and achieves optimum maximum likelihood (ML) detection with less complexity. Computer simulations confirm the superior performance of our proposed approach to conventional message passing techniques such as belief propagation (BP).**

I. Introduction

To deal with the recent exponential growth of wireless data traffic, multi-input multi-output (MIMO) systems have gained much attention from not only academia but also industry. Its high spectral efficiency is achieved by simultaneously transmitting multiple independent signals, i.e., *streams*, from multiple antennas. This simultaneous transmission however causes strong interference between streams at the receiver side, and the receiver has to treat this interference to detect the data.

Maximum likelihood detection (MLD) is known as an optimal detection technique while its computational complexity is exponentially high. To reduce the complexity, linear spatial filter based on zero-forcing (ZF) or minimum mean square error (MMSE) is an attractive alternative. Although these spatial filters significantly reduce the receiver complexity, their detection performance is even worse than MLD especially when the number of streams becomes large. Moreover, if powerful error-correcting codes approaching to the channel capacity such as low-density parity check (LDPC) codes or turbo codes are utilized in transmission, their channel decoders require soft-outputs of the detector, i.e., *marginal probabilities*, while these spatial filters cannot output such soft information basically.

Belief propagation (BP) is a practical and powerful method to calculate marginal probabilities; this efficient calculation is performed via a parallel computation called *message passing* on a factor graph composed of variable and factor (observation) nodes [1]. In BP algorithm, messages (beliefs) are iteratively sent from one node to its neighboring nodes. These beliefs become identical to exact marginal probabilities if and only if the graph is a tree. Nevertheless, BP algorithm exhibits good approximation results of marginal probabilities even if the graph has cycles [2].

BP-based detection for MIMO systems has been studied in [3] where factor graphs are defined by channel matrices between transmitting and receiving antennas. Corresponding graphs are always fully-connected, namely *complete graphs*, since signals transmitted from one antenna arrive at all receiving antennas. Thus, the performance of BP-based detection is remarkably degraded compared with MLD in terms of bit error rate (BER) [4] when the number of antennas is not so large. To overcome this *loopy* graph problem, QR-decomposed BP has been proposed in [5]. After QR-decomposition, the graph has less number of cycles, and this transformation leads to the better convergence of BP while the BER performance is still worse than MLD. As another approach, QR-decomposed *generalized belief propagation* (GBP) for MIMO systems has been proposed in [6]. Using a region graph instead of the factor graph after QR-decomposition, QR-GBP can achieve the near-optimal performance close to MLD. However its computational complexity increases exponentially as the number of antennas and modulation level increase. Thus, a complexity reduction technique for QR-GBP has been studied in [7]. Also, it is worth noting that the paper also proposed the log-domain calculation of GBP algorithm so as to make it computationally stable.

In this paper, we propose a *simple message passing* (SMP) detector for MIMO systems. This detector is also based on QR-decomposition and exploits the structure of QR-decomposed received signals. Numerical results confirm that our proposed detector achieves the identical performance to MLD while its complexity is even less than MLD.

II. System model

Let us consider a MIMO system with N_t transmitting and N_r receiving antennas as illustrated in Fig. 1. Throughout the paper, we assume a spatial multiplexing system; each antenna transmits an independent stream. Let $\mathbf{x} = [x_1, x_2, \ldots, x_{N_t}]^{\mathrm{T}}$ denote a transmitted real signal vector, and $\mathbf{H}$ denote a real channel matrix whose i-th row and j-th column element h_{ij} is an independent and identically distributed (i.i.d.) Gaussian random variable with zero mean and half variance, where $(\cdot)^{\mathrm{T}}$ denotes the transpose of vector or matrix. Without loss of generality, we assume that binary phase shift keying (BPSK) is used as a modulation. The channel matrix $\mathbf{H}$ is assumed to be ideally available only at the receiver side. Then, a received real signal vector $\mathbf{y} = [y_1, y_2, \ldots, y_{N_r}]^{\mathrm{T}}$ can be written by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{1}$$

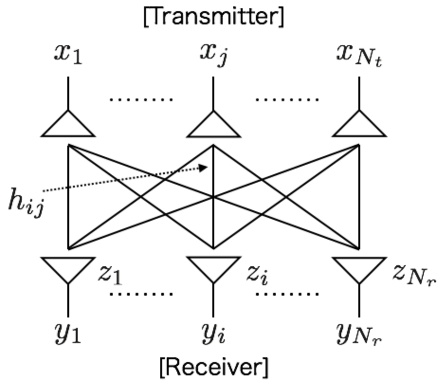

Fig. 1. System model with N_t transmitting and N_r receiving antennas.

where $\mathbf{z} = [z_1, z_2, \ldots, z_{N_r}]^{\mathrm{T}}$ is a noise vector whose elements are i.i.d. Gaussian random variables with zero mean and variance $N_0/2$.

III. Simple Message Passing Algorithm and Generalized Belief Propagation

A. Simple Message Passing Algorithm

A major idea of our proposed SMP algorithm is to reduce the number of variables in marginalization. SMP algorithm is performed on the graph named *regional graph*[1]. This graph is constructed from the factor graph defined by MIMO channel matrix. In the following, we explain how to construct this regional graph.

Applying QR-decomposition to the channel matrix $\mathbf{H}$, (1) can be rewritten as

$$\mathbf{y} = \mathbf{QRx} + \mathbf{z}, \tag{2}$$

where $\mathbf{Q}$ is a unitary matrix and $\mathbf{R}$ is an upper triangular matrix. Multiplying $\mathbf{Q}^{\mathrm{T}}$ from left, (2) can be rewritten as

$$\tilde{\mathbf{y}} = \mathbf{Rx} + \tilde{\mathbf{z}}, \tag{3}$$

where $\tilde{\mathbf{y}} = \mathbf{Q}^{\mathrm{T}}\mathbf{y} = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{N_r}]^{\mathrm{T}}$, and $\tilde{\mathbf{z}} = \mathbf{Q}^{\mathrm{T}}\mathbf{z} = [\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_{N_r}]^{\mathrm{T}}$ is i.i.d. Gaussian random variables with zero mean and variance $N_0/2$. In the rest of the paper, $(N_t, N_r) = (4, 4)$ is assumed for ease of explanation. Then, QR-decomposed factor graph can be represented as Fig. 2.

A *region* is defined as follows. A region of a factor graph is a set of variable and observation nodes. A single observation node and all the neighboring variable nodes must be in each region. From Fig. 2, four regions A, B, C, and D can be obtained. Then, those regions are connected by arrows according to a following rule; a region with k variable nodes must be a tail of an arrow and a region with $(k-1)$ variable nodes must be a head of an arrow where $k = 2, \ldots, N_t$. Then, the regional graph corresponding to the factor graph in Fig. 2 can be illustrated as Fig. 3. In this case, region A is a parent region of region B, region B is a parent region of region C, and region C is a parent region of region D.

[1]Note that our regional graph is different from *region graph* proposed in [2].

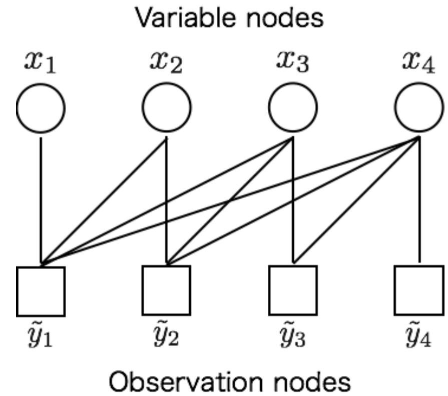

Fig. 2. QR-decomposed factor graph with $(N_t, N_r) = (4, 4)$.

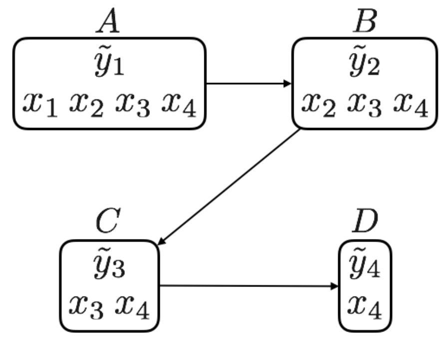

Fig. 3. Degenerated region graph for SMP algorithm with $(N_t, N_r) = (4, 4)$.

Based on this regional graph, messages are calculated and passed along arrows. A message from region P to region R is given by

$$m_{\mathrm{P}\to\mathrm{R}}(\mathbf{x}_{\mathrm{R}}) = \sum_{\mathbf{x}_{\mathrm{P}} \setminus \mathbf{x}_{\mathrm{R}}} p(\tilde{y}_{\mathrm{P}} | \mathbf{x}_{\mathrm{P}}) m_{\mathrm{P}'\to\mathrm{P}}(\mathbf{x}_{\mathrm{P}}) \tag{4}$$

where $\mathbf{x}_{\mathrm{R}}$ and $\mathbf{x}_{\mathrm{P}}$ are all the variables in the region R and region P respectively, $\mathbf{x}_{\mathrm{P}} \setminus \mathbf{x}_{\mathrm{R}}$ is a single variable in region P but not in R, and P$'$ is parent region of P.

In SMP, these messages must be calculated from the largest region to the smallest region. For example, a message from A to B is calculated at first. Namely,

$$m_{\mathrm{A}\to\mathrm{B}}(x_2, x_3, x_4) = \sum_{x_1} p(\tilde{y}_1 | x_1, x_2, x_3, x_4). \tag{5}$$

In this equation, any messages are not used since region A does not have a parent region. Subsequently, $m_{\mathrm{A}\to\mathrm{B}}(x_2, x_3, x_4)$ is calculated as

$$m_{\mathrm{B}\to\mathrm{C}}(x_3, x_4) = \sum_{x_2} p(\tilde{y}_2 | x_2, x_3, x_4) m_{\mathrm{A}\to\mathrm{B}}(x_2, x_3, x_4). \tag{6}$$

Finally, the message $m_{\mathrm{C}\to\mathrm{D}}(x_4)$ is calculated as

$$m_{\mathrm{C}\to\mathrm{D}}(x_4) = \sum_{x_3} p(\tilde{y}_3|x_3, x_4) m_{\mathrm{B}\to\mathrm{C}}(x_3, x_4). \tag{7}$$

Using these messages, every transmitted symbol is successively decided. An estimate of transmitted signal x_i, $i = 1, 2, \ldots, N_t$, is given by

$$\hat{x}_i = \arg\max_{x_i \in \{\pm 1\}} \left[p(\tilde{y}_i|x_i, \hat{x}_{i+1} \ldots \hat{x}_{N_t}) m_{\mathrm{I}'\to\mathrm{I}}(x_i, \hat{x}_{i+1} \ldots \hat{x}_{N_t}) \right], \tag{8}$$

where region I is the region with $x_i, \ldots, x_{N_t}$, and I$'$ is the parent region of region I. As obvious in (8), in order to decide x_i, estimates $\hat{x}_{i+1} \ldots \hat{x}_{N_t}$ from children are necessary. This is the reason why we have to start the decision from x_{N_t}. Then all the estimated values are respectively given by

$$\hat{x}_4 = \arg\max_{x_4 \in \{\pm 1\}} \left[p(\tilde{y}_4|x_4) m_{\mathrm{C}\to\mathrm{D}}(x_4) \right], \tag{9}$$

$$\hat{x}_3 = \arg\max_{x_3 \in \{\pm 1\}} \left[p(\tilde{y}_3|x_3, \hat{x}_4) m_{\mathrm{B}\to\mathrm{C}}(x_3, \hat{x}_4) \right], \tag{10}$$

$$\hat{x}_2 = \arg\max_{x_2 \in \{\pm 1\}} \left[p(\tilde{y}_2|x_2, \hat{x}_3, \hat{x}_4) m_{\mathrm{A}\to\mathrm{B}}(x_2, \hat{x}_3, \hat{x}_4) \right], \tag{11}$$

$$\hat{x}_1 = \arg\max_{x_1 \in \{\pm 1\}} \left[p(\tilde{y}_1|x_1, \hat{x}_2, \hat{x}_3, \hat{x}_4) \right]. \tag{12}$$

As obvious from above equations, SMP is identical with standard MLD while its complexity is less than MLD because of successive decisions of transmitted signals.

B. Interpretation of SMP as Generalized Belief Propagation in Degenerated Region Graph

In this section, we further describe that our proposal SMP algorithm can be interpreted as *parent-to-child* (PtC) algorithm on *degenerated* but still *valid* region graph. PtC algorithm is performed on a graph called *region graph* [2]. Different from regional graph representation, region graph has two different regions: largest regions and child regions. Each largest region includes a single observation node and neighboring variable nodes while child regions consist of common variable nodes of their parent regions. In the following, if there exists a directed path from region v_a to v_d, we say that v_a is an *ancestor* of v_d and v_d is a *descendant* of v_a. Let us define the *counting number* c_v for every region v as

$$\mathrm{c}_v = 1 - \textstyle\sum_{\mathrm{R} \in A(v)} c_{\mathrm{R}}, \tag{13}$$

where $A(v)$ is the set of regions that are ancestors of region v. When the counting number meets the following *region graph condition*: $\sum_{\mathrm{R}} c_{\mathrm{R}} = 1$, the region graph is valid. If this condition does not hold, the accuracy of PtC algorithm is not guaranteed. Note that the construction of region graphs is not unique, and different region graphs would lead to different message passing algorithm and performance.

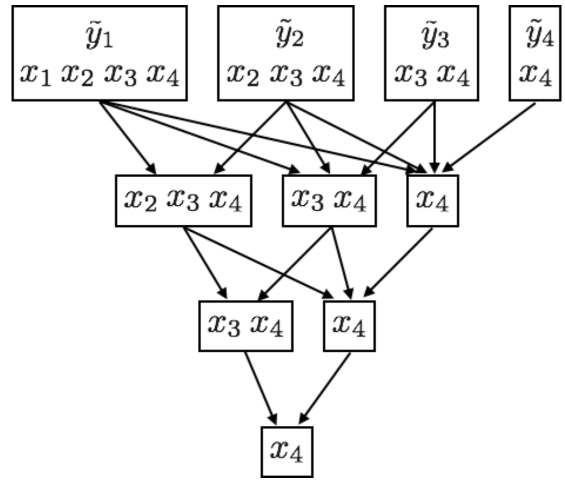

Fig. 4. Example of original region graph for PtC algorithm with $(N_t, N_r) = (4, 4)$.

1) Parent-to-Child Algorithm: In PtC algorithm, each region R has a belief $b_{\mathrm{R}}(\mathbf{x}_{\mathrm{R}})$ given by

$$b_{\mathrm{R}}(\mathbf{x}_{\mathrm{R}}) \propto p(\tilde{y}_{\mathrm{R}}|\mathbf{x}_{\mathrm{R}}) \left(\prod_{\mathrm{P}\in\mathcal{P}(\mathrm{R})} m_{\mathrm{P}\to\mathrm{R}}(\mathbf{x}_{\mathrm{R}}) \right) \left(\prod_{\mathrm{D}\in\mathcal{D}(\mathrm{R})} \prod_{\mathrm{P}'\in\mathcal{P}(\mathrm{D})\setminus\varepsilon(\mathrm{R})} m_{\mathrm{P}'\to\mathrm{D}}(\mathbf{x}_{\mathrm{D}}) \right), \tag{14}$$

where $m_{\mathrm{P}\to\mathrm{R}}(\mathbf{x}_{\mathrm{R}})$ is a message passed from region P to region R with transmitted signals included in region R which are denoted by $\mathbf{x}_{\mathrm{R}}$. Also, $\mathcal{P}(\mathrm{R})$ is a set of regions which are parents of region R, and $\mathcal{D}(\mathrm{R})$ is a set of all regions that are descendants of region R. $\varepsilon(\mathrm{R}) \equiv \mathrm{R} \cup \mathcal{D}(\mathrm{R})$ is a set of all regions that are descendants of R and also R itself.

Also, the message-update rule in GBP algorithm is given by

$$m'_{\mathrm{P}\to\mathrm{R}}(\mathbf{x}_{\mathrm{R}}) := \sum_{\mathbf{x}_{\mathrm{P}}\setminus\mathbf{x}_{\mathrm{R}}} p(\tilde{y}_{\mathrm{P}}|\mathbf{x}_{\mathrm{P}}) \frac{\prod_{(\mathrm{I},\mathrm{J})\in N(\mathrm{P},\mathrm{R})} m_{\mathrm{I}\to\mathrm{J}}(\mathbf{x}_{\mathrm{J}})}{\prod_{(\mathrm{I},\mathrm{J})\in D(\mathrm{P},\mathrm{R})} m_{\mathrm{I}\to\mathrm{J}}(\mathbf{x}_{\mathrm{J}})}, \tag{15}$$

where $N(\mathrm{P},\mathrm{R})$ is the set of all connected pairs of region (I,J) such that J is in $\varepsilon(\mathrm{P})$ but not in $\varepsilon(\mathrm{R})$ while I is not in $\varepsilon(\mathrm{P})$. $D(\mathrm{P},\mathrm{R})$ is the set of all connected pairs of regions (I,J) such that J is in $\varepsilon(\mathrm{R})$ while I is in $\mathcal{D}(\mathrm{P})$ but not in $\varepsilon(\mathrm{R})$. After the sufficient number of message updates, beliefs of the smallest regions are used for the decision.

2) GBP on Degenerated Region Graph: Figure 4 shows the valid region graph constructed from the factor graph in Fig. 2. This region graph contains cycles, and thus the performance of PtC algorithm does not achieve the optimum performance of MLD [2]. Eliminating some arrows and regions, we obtain *degenerated* region graph without cycles as shown in Fig. 5. Note that this graph is still valid.

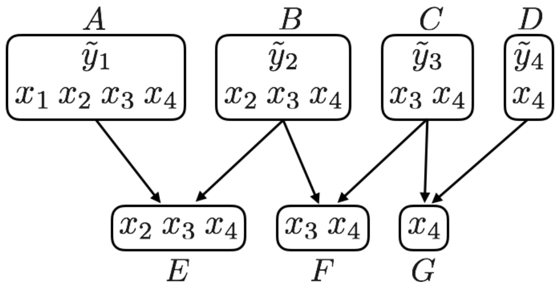

Fig. 5. Example of degenerated but valid region graph for parent-to-child algorithm with $(N_t, N_r) = (4, 4)$.

Applying PtC algorithm to this degenerated but valid graph, corresponding messages are given by

$$m'_{\mathrm{A}\to\mathrm{E}}(x_2, x_3, x_4) = \sum_{x_1} p(\tilde{y}_1|x_1, x_2, x_3, x_4), \tag{16}$$

$$m'_{\mathrm{B}\to\mathrm{F}}(x_3, x_4) = \sum_{x_2} p(\tilde{y}_2|x_2, x_3, x_4) m_{\mathrm{A}\to\mathrm{E}}(x_2, x_3, x_4), \tag{17}$$

$$m'_{\mathrm{C}\to\mathrm{G}}(x_4) = \sum_{x_3} p(\tilde{y}_3|x_3, x_4) m_{\mathrm{B}\to\mathrm{F}}(x_3, x_4). \tag{18}$$

These messages are calculated by (5)–(7). Then the rest of messages are also calculated by

$$m'_{\mathrm{D}\to\mathrm{G}}(x_4) = p(\tilde{y}_4|x_4), \tag{19}$$

$$\begin{aligned} m'_{\mathrm{C}\to\mathrm{F}}(x_3, x_4) &= p(\tilde{y}_3|x_3, x_4) m_{\mathrm{D}\to\mathrm{G}}(x_4) \\ &= p(\tilde{y}_3, \tilde{y}_4|x_3, x_4), \end{aligned} \tag{20}$$

$$\begin{aligned} m'_{\mathrm{B}\to\mathrm{E}}(x_2, x_3, x_4) &= p(\tilde{y}_2|x_2, x_3, x_4) m_{\mathrm{C}\to\mathrm{F}}(x_3, x_4) \\ &= p(\tilde{y}_2, \tilde{y}_3, \tilde{y}_4|x_2, x_3, x_4). \end{aligned} \tag{21}$$

Then, beliefs of region E, F, G are respectively given by

$$\begin{aligned} b_G(x_4) &= m'_{C\to G}(x_4) m'_{\mathrm{D}\to\mathrm{G}}(x_4) \\ &= m'_{\mathrm{C}\to\mathrm{G}}(x_4) p(\tilde{y}_4|x_4). \end{aligned} \tag{22}$$

$$\begin{aligned} b_F(x_3, x_4) &= m'_{\mathrm{B}\to\mathrm{F}}(x_3, x_4) m'_{\mathrm{C}\to\mathrm{F}}(x_3, x_4) \\ &= m'_{\mathrm{B}\to\mathrm{F}}(x_3, x_4) p(\tilde{y}_3, \tilde{y}_4|x_3, x_4), \end{aligned} \tag{23}$$

$$\begin{aligned} b_E(x_2, x_3, x_4) &= m'_{\mathrm{A}\to\mathrm{E}}(x_2, x_3, x_4) m'_{\mathrm{B}\to\mathrm{E}}(x_2, x_3, x_4) \\ &= m'_{\mathrm{A}\to\mathrm{E}}(x_2, x_3, x_4) p(\tilde{y}_2, \tilde{y}_3, \tilde{y}_4|x_2, x_3, x_4), \end{aligned} \tag{24}$$

These messages are identical to the right hand side of (9)–(11).

Finally, estimates of PtC algorithm are respectively given by

$$\hat{x}_4 = \arg\max_{x_4\in\{\pm 1\}} \left[b_G(x_4) \right], \tag{25}$$

$$\hat{x}_3 = \arg\max_{x_3\in\{\pm 1\}} \left[b_F(x_3, \hat{x}_4) \right], \tag{26}$$

$$\hat{x}_2 = \arg\max_{x_2\in\{\pm 1\}} \left[b_E(x_2, \hat{x}_3, \hat{x}_4) \right], \tag{27}$$

$$\hat{x}_1 = \arg\max_{x_1\in\{\pm 1\}} \left[p(\tilde{y}_1|x_1, \hat{x}_2, \hat{x}_3, \hat{x}_4) \right]. \tag{28}$$

Equations (12) and (28) are identical, and the other equations are almost same. The difference between corresponding equations is the number of observation nodes used for the decision. Namely, PtC algorithm utilizes more observations. For example, to estimate transmitted signal x_2, PtC algorithm uses all the observation nodes. However x_2 is not received at $\tilde{y}_3$ and $\tilde{y}_4$. Therefore, $\tilde{y}_3$ and $\tilde{y}_4$ are not necessary to estimate x_2, and that is the reason why SMP algorithm only uses $\tilde{y}_1$ and $\tilde{y}_2$.

Summarizing above discussions, PtC algorithm propagates all the observation values, but SMP algorithm propagates only indispensable messages. This leads to the complexity reduction. Moreover, as described above, SMP algorithm is identical to MLD, and thus PtC algorithm on degenerated but valid graph shown in Fig. 5 provides the better performance than one on the original graph shown in Fig. 4. This interesting observation suggests that eliminating cycles in given graph may result in better convergence (performance) when message passing algorithm based on free energy approximation is used [2].

IV. Numerical Results

Finally, we evaluate the BER performance of proposed SMP detection and its computational complexity via computer simulations.

Figure 6 shows the BER performances of BP [3], QR-decomposed BP (QR-BP) [5], PtC algorithm [2] on original region graphs, PtC on degenerated region graphs, our proposed SMP, and MLD when the number of both N_t and N_r is assumed to be 8. The number of iterations of BP and QR-BP is assumed to be 15. The number of iteration of GBP with PtC algorithm is assumed to be 7. Moreover, numerically-optimized damping is used for BP and QR-BP. As observed from the figure, our proposed SMP achieves the identical performance to MLD while BP and QR-BP exhibit large gap from the MLD. PtC algorithm on the degenerated graph also achieves the same performance while PtC algorithm on the original graph exhibits a gap from the MLD.

Figure 7 shows complexity comparison of BP, QR-BP, PtC algorithm on degenerated but valid region graph, proposed SMP detection, and MLD in terms of the number of multiplications. Vertical axis shows the number of multiplication and horizontal axis shows the number of antennas where $N_t = N_r$ is assumed. Clearly, SMP algorithm requires even less complexity than the other algorithms especially when the number of antennas becomes large.

From these observations, we can conclude that SMP algorithm can achieve the optimal performance with even lower complexity than PtC algorithm and MLD.

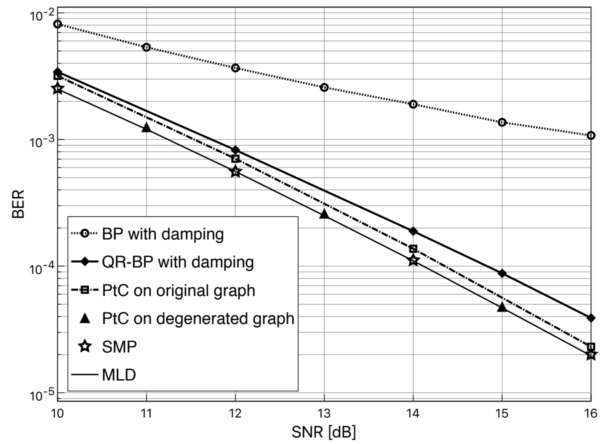

Fig. 6. BER performances of conventional BP with damping, QR-decomposed BP with damping , PtC algorithm on original region graph, PtC algorithm on degenerated but valid region graph, proposed SMP detection, and MLD where $(N_t, N_r) = (8, 8)$.

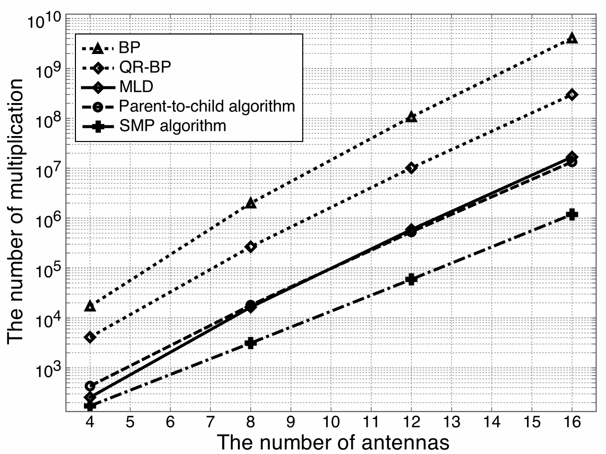

Fig. 7. Complexity comparison of conventional BP, QR-BP, PtC algorithm on degenerated but valid graph, proposed SMP detection, and MLD.

V. Conclusion

In this paper, we have proposed SMP detection. SMP detection achieves the MLD performance while the complexity is even less than the MLD.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP15H02252 and JP18H03765. We sincerely thank Prof. T. Ohgane, Prof. K. Hayashi, Prof. S. Ibi, and Prof. T. Nishimura for their insightful comments and discussions.

References

[1] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[2] J. S. Yedidia, W. T. Freeman, and Y. Weiss, “Constructing free energy approximations and generalized belief propagation algorithms,” *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282-2312, July 2005.

[3] J. Hu and T. M. Duman, “Graph-based detection algorithms for layered space-time architectures,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 269-280, Feb. 2008.

[4] J. Soler-Garrido, R. J. Piechocki, K. Maharatna, and D. McNamara, “Analog MIMO detection on the basis of Belief Propagation,” in 49th IEEE International Midwest Symposium on Circuits and Systems, pp. 50–54, Aug. 2006.

[5] S. Park and S. Choi, “QR decomposition aided belief propagation detector for MIMO systems,” *Electron. Lett.*, vol.51, no.11, pp.873–874, May 2015.

[6] A.D. Shigyo, S. Tanabe, and K. Ishibashi, “QR-decomposed generalized belief propagation for MIMO detection,” in The International Symposium on Information Theory and Its Applications (ISITA’16), Monterey, CA, pp. 468–472, Nov. 2016.

[7] A.D. Shigyo and K. Ishibashi, “QR-decomposed generalized belief propagation with smart message reduction for low-complexity MIMO signal detection,” Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1795–1799, Dec. 2017.

[8] T. Takahashi, S. Ibi, T. Ohogane and S. Sampei, “On normalized belief of Gaussian BP in correlated large MIMO channels,” in The International Symposium on Information Theory and Its Applications (ISITA’16), Monterey, CA, pp. 458–462, Nov. 2016.

[9] R. Kikuchi, “A theory of cooperative phenomena,” *Phys. Rev.*, vol. 81, no. 6, pp. 988–1003, 1951.