Belief-Selective Propagation Detection for MIMO Systems

Wenyue Zhou, Yifei Shen, Student Member, IEEE, Liping Li, Yongming Huang, Senior Member, IEEE, Chuan Zhang, Senior Member, IEEE, and Xiaohu You, Fellow, IEEE

Abstract—Compared to the linear MIMO detectors, the Belief Propagation (BP) detector has shown greater capabilities in achieving near-optimal performance and better nature to iteratively cooperate with channel decoders. Aiming at real applications, recent works mainly fall into the category of reducing the complexity by simplified calculations, at the expense of performance sacrifice. However, the complexity is still unsatisfactory with exponentially increasing complexity or required exponentiation operations. Furthermore, the state-of-the-art (SOA) BP detectors persistently encounter error floor in high signal-to-noise ratio (SNR) region, which becomes even worse with calculation approximation. This work aims at a revised BP detector, named Belief-selective Propagation (BsP) detector by selectively utilizing the trusted incoming messages with sufficiently large a priori probabilities for updates. Two proposed strategies: symbol-based truncation (ST) and edge-based simplification (ES) squeeze the complexity (orders lower than the BP detector), while greatly relieving the error floor issue over a wide range of antenna and modulation combinations. For the 256-QAM 128 × 64 uplink massive multiuser MIMO (MU-MIMO) system, the BsP detector achieves more than 1 dB performance gain (@BER = 10^{-3}) with lower complexity than the state-of-the-art (SOA) BP detector. Trade-off between performance and complexity towards different application requirements can be conveniently obtained by tuning the parameters of the ST and ES strategies.

Index Terms—Multiple-input multiple-output (MIMO), belief-selective propagation (BsP), error floor, complexity.

I. INTRODUCTION

MULTIPLE-INPUT Multiple-Output (mimo) ranks among the key enabling technologies for nowadays wireless systems by offering high spectrum efficiency (SE) and energy efficiency (EE) [1], [2]. Therefore, designing an efficient MIMO detection algorithm balancing both performance and complexity becomes one important topic [3]. Although the maximum a posteriori (MAP) detection or the maximum likelihood (ML) detection leads to minimum error probability, the exponentially increasing complexity with modulation order and MIMO scale hinders real applications. To this end, the sphere decoder (SD) [4] has been proposed to reduce complexity by restricting candidate symbols to ones within a sphere (depth-first) [5] or a list of size K (width-first) [6]. Providing sub-optimal performance with acceptable complexity, SD enjoys a very successful application for small-scale MIMO. However, for large-scale MIMO, it still suffers from high complexity penalty, which increases in cubic (often roughly) with system scale [7]. For large-scale MIMO, one compromise choice is to employ linear detectors such as zero-forcing (ZF) detector or more often the linear minimum mean square error (LMMSE) detector [3]. Though the linear detector is much simpler and more hardware friendly, its performance is much worse than that of the MAP/ML detector or SD [8]. Furthermore, for large-scale MIMO, the excessive complexity resulting from matrix inversion becomes a problem. By making use of the channel sparsity [9], approaches such as Neumann series approximation (NSA) [10] and iterative methods [11] have been considered to address the matrix inversion issue. However, the performance of these linear variants is still bounded by that of the LMMSE detector, and cannot fully exploit the MIMO benefits.

With the fast development of wireless communications, especially the emergence of new service scenarios, such as the enhanced mobile broadband (eMBB) service and the low-latency communications (URLLC) service [12], [13], MIMO detectors which can deliver performance beyond the LMMSE curve with good implementation feasibility are highly expected.1 Among all candidates, the Belief Propagation (BP) detector [14], which is based on the Bayes’ rule [15], [16], has recently drawn increasing attentions from both academia and industry. Although the BP detector, at least for now is not an “all-around winner” for different performance/complexity requirements, it has shown considerable potential as a well-balanced solution with the following merits. First, the BP detector can deliver near-optimal

1Admittedly, the satisfactory system performance can still be achieved by employing more advanced channel codes with LMMSE detector. This topic is out of the scope of this paper and will be discussed in another paper of ours.
performance and outperforms the linear ones. Second, the BP detector inherently holds on a soft-input soft-output character, which can iteratively work with channel decoders and other baseband modules in nature. Third, thanks to the regularity of the factor graph (FG), the BP detector is hardware-friendly and can be easily parallelized and scaled for applications.

The MIMO BP detector with the FG representation is originally proposed in [14] for the Bell labs layered space-time (BLAST) architecture [17]. It can achieve near-optimal performance in small-to-medium scale MIMO, with the cost of exponentially increasing complexity. The following works mainly focus on reducing the complexity while maintaining the performance advantage of BP detector. The approaches can be roughly categorized into two groups: simplifying the edges of the FG, and simplifying the nodes of the FG.

Edge reduction: To lower the complexity, the regular-\(d_f\) BP (EBRD-BP) detector is proposed in the same paper [14], which is in fact an edge-pruning approach by remaining the edges corresponding to the \(d_f\) largest channel coefficients whereas disconnecting the others. Simulation results have shown the complexity reduction with acceptable performance degradation.

Node reduction: A major part of related literatures apply the node reduction. In [18], the BP detector with the Gaussian approximation of interference (GAI-BP) has been proposed to reduce the node calculation. In [19], the single edge-based BP detector with the pseudo prior (PP) information initialization and the Gaussian approximation feedback (GF) information has been proposed. This PP-GF-BP detector further reduces the computational complexity. However, since both the GAI-BP and the PP-GF-BP detectors are bit-based, their performance with high-order modulations has been sacrificed. In [20], the GAI-BP detector is extended to the real domain (RD-GAI-BP), which successfully outperforms both the GAI-BP and the PP-GF-BP detectors with high-order modulations. The Gaussian sum-product algorithm over a wireless network (SPAWN) is proposed in [21], which reduces the complexity by replacing the messages with Gaussian means and variances. The convergence properties of Gaussian SPAWN are further analyzed in [22].

So far, the design of efficient MIMO BP detectors remains challenging. First, the existing BP detectors suffer an aggravated performance error floor in relatively high signal-to-noise ratio (SNR) region, due in part to the well-known impact of the inherent loopy structure for the full-connected FG model [23], in part to the approximations such as GAI. Second, the complexity of existing BP detectors is still high and not flexible enough for various applications. These multiple challenges motivate us to think of an approach which can reduce the detection complexity while mitigating the error floor towards better performance.

In this paper, a Belief-selective Propagation (BsP) detector is proposed, which only utilizes the incoming messages with relatively large \textit{a priori} probabilities to update the output messages. Avoiding to propagate the messages with “low beliefs”, the belief-selective method can not only mitigate the error floor performance of the state-of-the-art (SOA) BP detector in high SNR region, but also lower the computation complexity. It is noted that a similar idea of the BsP was previously proposed for decoding non-binary low-density parity-check (NB-LDPC) codes [24], [25], [26], as extended min-sum (EMS). The initial motivation of EMS was only for lower complexity but not for better performance. Second, its application for MIMO detection is not straightforward. Compared to the EMS decoding, the proposed BsP detector differs in the message update for factor nodes (FNs). The BsP detector updates the beliefs by exhaustively searching for the most likely transmitted symbol vector in the given set, whereas the EMS decoding directly computes the messages associated with symbols using the check-node constraints.

Contributions: The contributions of this work are listed as follows:

- The MIMO BsP detector has been first proposed. Compared with the SOA BP detector, it can offer better performance with lower complexity. The persistent error-floor issue of the SOA BP detector has been greatly relieved.
- The symbol-based truncation (ST) strategy and the edge-based simplification (ES) strategy are proposed to make the BsP detector feasible for applications. With both strategies, the transmitted symbols corresponding to relatively small probabilities are removed from the search space in the node message update. Therefore, the complexity of the BsP detector has been further reduced. Both strategies are equipped with adjustable parameters to flexibly meet different performance/complexity requirements.
- The results of an LMMSE detector is utilized to provided the PP information of the transmitted symbols for the proposed BsP detector. Activating by the results of the LMMSE detector can not only solve the inherent initialization requirement of the BsP detector with low complexity cost, but also accelerate the BsP detector’s convergence.
- The convergence analysis of the proposed BsP detector is implemented by means of the extrinsic information transfer chart (EXIT chart). Furthermore, a case study together with theoretical derivations are provided to demonstrate the effectiveness of the proposed BsP. In particular, the error floor performance of the SOA BP detector is first illustrated. Then, the reason that the belief-selective method relieves the error floor is derived theoretically.

In addition, numerical results confirm that the advantages of BsP detector over its counterparts hold in various MIMO scenarios. This robustness guarantees its wide applications.

The reminder of this paper is organized as follows. In Section II, the system model and the MIMO BP detector are reviewed. In Section III, the BsP detector is proposed with details. The computational complexity issues of the BsP detector are analyzed and discussed in Section IV. The convergence analysis is provided in Section V. The numerical results are given in Section VI. Finally, Section VII concludes the entire paper.
II. PRELIMINARIES

A. System Model and the Optimal Solution

Consider a MIMO system with \( N_r \) receive antennas and \( N_t \) transmit antennas, then the transmitted symbols is a \( N_t \times 1 \) i.i.d. vector \( s = [s_1, s_2, \ldots, s_{N_t}]^T \). Assume all of the transmitted symbols are chosen from the same QAM modulation of order \( M \), then the cardinality of the constellation is \( |A| = 2^M \). Let the \( N_r \times N_t \) matrix \( \mathbf{H} = [h_{11}, h_{12}, \ldots, h_{N_rN_t}]^T \) denote the flat-fading complex MIMO channel, where \( h_{ij} \) represents the \( i \)-th row vector of \( \mathbf{H} \), and each component of \( h_{ij} \) is a complex fading channel coefficient following the zero-mean and unit-variance Gaussian distribution. The received signal \( y = [y_1, y_2, \ldots, y_{N_r}]^T \) is given by

\[
y = \mathbf{H} s + n,
\]

where \( n = [n_1, n_2, \ldots, n_{N_r}]^T \) is a noise vector. The components of \( n \) are complex independent white Gaussian variables with zero mean and \( \sigma^2 \) variance.

The goal of the MIMO detection is to recover the transmitted symbol vector \( s \) according to the received signal \( y \) and the channel matrix \( \mathbf{H} \). The optimal solution to achieve this goal is the MAP detector, which requires an exhaustive search to find the transmitted symbols. This MAP detector is illustrated by

\[
\hat{s} = \arg \max_{s \in A^{N_t}} P(s|y, \mathbf{H}),
\]

where the a posteriori probability \( P(s|y, \mathbf{H}) \) can be rewritten as

\[
P(s|y, \mathbf{H}) = P(y|s, \mathbf{H})P(s),
\]

employing the Bayes’ rule. Although MAP detector has the optimal error performance, it suffers from an exponentially increasing computational complexity dominated by the number of transmit antennas and the order of the modulations.

B. BP Detection With the FG Model

Apart from the optimal solution, BP algorithm based on FG model is an alternative solution for the MIMO detection. Fig. 1 shows FG model of BP detector. There are two types of nodes in a FG model, namely FN and symbol node (SN), denoted as

\[
\begin{align*}
f_i & \leftrightarrow \text{FN}, \quad i \in \{1, 2, \ldots, N_r\}, \\
S_j & \leftrightarrow \text{SN}, \quad j \in \{1, 2, \ldots, N_t\}.
\end{align*}
\]

This model is related to the channel matrix \( \mathbf{H} \). In particular, each \( f_i \) corresponds to the \( i \)-th row of the channel matrix, and each \( S_j \) relates to the \( j \)-th column of the channel matrix. Every pair of \( f_i \) and \( S_j \) are connected with an edge corresponding to the channel coefficient \( h_{ij} \).

With the received signal \( y \), BP detector begins with the update of message \( \beta \). In particular, the message \( \beta_{ij} \) will be delivered from \( f_i \) to \( S_j \). Then the BP detector computes message \( \alpha \) with the received \( \beta \) messages for each SN, and transmits them in a reverse direction. BP detector iteratively updates \( \beta \) and \( \alpha \) messages, until it achieves the maximum number of the iterations. The BP detection ends up with computing the output soft messages of the coded bits as shown in Fig. 1.

In this paper, log-likelihood ratios (LLRs) are used to represent the soft messages. The \( \alpha \) message passing from the \( j \)-th SN to the \( i \)-th FN at the \( l \)-th iteration is denoted as

\[
\alpha_{ji}(l)(k) = \log \frac{p_i(s_j | \mu_k)}{p_i(s_j | \mu_1)},
\]

where \( \mu_k \) (\( k \in \{1, 2, \ldots, |A|\} \)) denotes \( k \)-th modulation symbol in the constellation. The details of the message update in the BP detection are as follows.

1) \( \beta \) Message Update: To update message \( \beta \), FNs receive message \( \alpha \) in the preceding iteration from SNs, then operate with the following formula

\[
\beta_{ij}(l)(k) = \max_{s_j : s_j = \mu_k} \left\{ \frac{-1}{2\sigma^2} \| y_i - h_{ij}s \|^2 + \sum_{t=1, t \neq j}^{N_t} \alpha_{ti}(l-1) \right\}
\]

\[
- \max_{s_j : s_j = \mu_1} \left\{ \frac{-1}{2\sigma^2} \| y_i - h_{ij}s \|^2 + \sum_{t=1, t \neq j}^{N_t} \alpha_{ti}(l-1) \right\},
\]

where \( \beta_{ij}(l) \) denotes message \( \beta \) passing from the \( i \)-th FN to the \( j \)-th SN at the \( l \)-th iteration. The detailed derivation of (6) can be referred to [14].

2) \( \alpha \) Message Update: In \( l \)-th iteration, SNs compute message \( \alpha \) by

\[
\alpha_{ji}(l)(k) = \sum_{t=1, t \neq i}^{N_r} \beta_{ij}(l)(k),
\]

which only involves addition operations. Therefore, message \( \alpha \) updating has much less computational complexity compared with message \( \beta \) updating.

BP detector terminates the message update when reaching the maximum iteration number, and then computes symbol
Given a group of symbol LLRs, its associated coded bit LLRs can be derived according to the FG of FNs \( M_j \) (\( j \in \{1, 2, \ldots, N_t\} \)) and related bit nodes \( c_m \) (\( m \in \{1, 2, \ldots, M\} \)) in Fig. 1. With the sum-product algorithm and the max-log approximation, the coded bit LLRs can be computed by

\[
\gamma_j(k) = \max_{\mu_k, \gamma_m = 1} \gamma_j(k) - \max_{\mu_k, \gamma_m = 0} \gamma_j(k),
\]

where \( r_j(m) \) is the \( m \)-th bit LLR associated with \( j \)-th SN. More details about (9) can be found in [27].

III. THE PROPOSED BS\( P \) DETECTION

The BP detector requires an exhaustive search involving all of possible choices of the transmitted symbols to update the message \( \beta_{ij} \) for each FN. Such an exhaustive search causes an exponential computational complexity dominated by the constellation cardinality (\(|A|\)) of the MIMO systems and the degree (\( N_i \)) of the FNs. Therefore, the BP detector suffers from a prohibitive computational complexity in high-order high-dimensional MIMO systems. Fortunately, it is observed that the transmitted symbol vector with relatively low reliability has limited contribution to the message \( \beta_{ij} \). Therefore, squeezing the search space by removing low-reliability transmitted symbols is an intuitively reasonable method to bring down the computational complexity. In addition, it is expected that circumventing the propagation of the messages with low reliability will also mitigate the error propagation problem, thus relieving the error floor performance.

Based on this idea, a low-complexity BS\( P \) detector is proposed in this paper. The BS\( P \) detector is equipped with the following two strategies: 1) ST strategy to reduce the cardinality of the transmitted symbol (from \(|A|\) to \( d_m, d_m \ll |A| \)); 2) ES strategy to simplified the connected edges of the FNs (from \( N_i \) to \( d_f, d_f \ll N_i \)). In addition, before the iterations of the message passing, the PP information for each transmitted symbol is acquired by exploiting the results of an LMMSE detector, which accelerates the convergence of the proposed BS\( P \). In this section, the ST and ES strategies are proposed first. Then the method for acquiring the PP information of the transmitted symbol is introduced. Finally, the BS\( P \) algorithm is summarized.

A. Symbol-Based Truncation Strategy

The basic idea of the ST strategy is that the cardinality of the transmitted symbol can be reduced by discarding symbols with relatively low probability, which make negligible contributions to the results of the detector. In particular, each incoming \( \alpha_{ki} \) is truncated to acquire \( \alpha_{ki}' \) by eliminating the LLR with relatively small values. Consequently, the transmitted symbol corresponding to \( \alpha_{ki}' \) has less possible choices, contributing to a smaller search space. Fig. 2(a) illustrates the procedure of updating \( \beta_{kj} \) with the ST strategy. Each \( s_k (k \neq j) \) is connected with a rectangle box involving the incoming \( \alpha_{ki} \), which consists of \(|A|\) symbol LLRs. The incoming message \( \alpha_{ki}^t \) contains \( d_m, d_m \ll |A| \) largest LLRs inheriting from \( \alpha_{ki} \). The message \( \beta_{ij} \) is computed by the following steps. First, each incoming \( \alpha_{ki} \) is sorted in a descend order, expressed as

\[
\alpha_{ki} = \{\alpha_{ki}(1), \alpha_{ki}(2), \ldots, \alpha_{ki}(|A|)\}.
\]

Second, \( \alpha_{ki}' \) is acquired by saving the first \( d_m \) symbol LLRs in \( \alpha_{ki} \), denoted as

\[
\alpha_{ki}' = \{\alpha_{ki}(1), \alpha_{ki}(2), \ldots, \alpha_{ki}(d_m)\}.
\]

Finally, \( \beta_{ij} \) is computed with the truncated message \( \alpha_{ki}' \). Let \( s_{k/j} \) denote the vector involving all of the elements in \( s \) except the \( s_j \), written as

\[
s_{k/j} = \{s_{k1}, s_{k2}, \ldots, s_{kj-1}, s_{kj+1}, \ldots, s_{kN_t}\}.
\]

Define \( B(d_m) \) as a configuration set consisting of all of the possible choices of \( s_{k/j} \), illustrated as

\[
B(d_m) = \left\{ s_{k/j} = \left[ \mu_{k_1}^1, \mu_{k_2}^2, \ldots, \mu_{k_{N_t-1}}^{N_t-1} \right]^T : \forall k = [k_1, k_2, \ldots, k_{N_t-1}]^T \in \{1, 2, \ldots, d_m\}^{N_t-1} \right\}.
\]

With \( B(d_m) \), \( \beta_{ij} \) is computed as (14), shown at the bottom of the next page. It is noted that the cardinality of \( B(d_m) \) is

\[
|B(d_m)| = (d_m)^{N_t-1}.
\]

With this cardinality, the number of possible choices of transmitted symbols in (14) is denoted as \( \psi_{ST} \)

\[
\psi_{ST} = |A| \times |B(d_m)|,
\]

which is smaller than that in (6). Therefore, the ST strategy reduces the computational complexity of the BP detector.

B. Edge-Based Simplification Strategy

Although the ST strategy relieves the computational complexity by reducing the cardinality of the transmitted symbols, the BP detector still suffers from very high computational complexity due to its exponential distribution, especially in high-order high-dimensional MIMO systems. In order to further reduce the computational complexity of the BP detector, the ES strategy is proposed here. With the ES strategy, some transmitted symbols in \( s_{k/j} \) are directly decided as the symbol with the largest reliability. Fig. 2 shows details of the ES strategy for computing \( \beta_{ij} \). Given \( N_t - 1 \) incoming edges associated with the messages \( \alpha_{ki}, k \neq j \), suppose that there are \( d_f - 1 \) chosen edges and \( N_t - d_f \) simplified edges. The incoming messages of the chosen edges remain unchanging (the black boxes), whereas the incoming messages corresponding to the simplified edges are directly truncated into \( \alpha_{ki}' \) (the blue boxes), which only involves the maximum LLR of the message \( \alpha_{ki} \). Consequently, the transmitted symbol associated with the simplified edge is the symbol with the maximum reliability.

The configuration set with this ES strategy is denoted as \( B(|A|, d_f) \), which is illustrated as

\[
B(|A|, d_f) = \left\{ s_{k/j} = \left[ s_{k/j,d_f}, \tilde{s}_{k/j,d_f}^T \right]^T \right\},
\]

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
C. Initialization With PP Information

The BS-P detector requires the a priori information for each transmitted symbol to build up the configuration set. However, this a priori information is unavailable (set as the zero information) when the BS-P detection starts to update the message β. A natural way to obtain the initial a priori information is to perform one iteration of the BP detector, but this will cause an unacceptable computational complexity, which is inconsistent with the motivation of the proposed BS-P detector. Recently, many works [19], [28], [29] utilize the results of LMMSE detector to acquire the PP information of transmitted symbols, which improves the error rate performance of the LMMSE detector to acquire the PP information of transmitted symbols.

The BS-P detector requires the a priori information for each transmitted symbol to build up the configuration set. However, this a priori information is unavailable (set as the zero information) when the BS-P detection starts to update the message β. A natural way to obtain the initial a priori information is to perform one iteration of the BP detector, but this will cause an unacceptable computational complexity, which is inconsistent with the motivation of the proposed BS-P detector. Recently, many works [19], [28], [29] utilize the results of LMMSE detector to acquire the PP information of transmitted symbols, which improves the error rate performance of the LMMSE detector to acquire the PP information of transmitted symbols.

The BS-P detector requires the a priori information for each transmitted symbol to build up the configuration set. However, this a priori information is unavailable (set as the zero information) when the BS-P detection starts to update the message β. A natural way to obtain the initial a priori information is to perform one iteration of the BP detector, but this will cause an unacceptable computational complexity, which is inconsistent with the motivation of the proposed BS-P detector. Recently, many works [19], [28], [29] utilize the results of LMMSE detector to acquire the PP information of transmitted symbols, which improves the error rate performance of the LMMSE detector to acquire the PP information of transmitted symbols.

![Fig. 2. (a) Updating message βij with the ST strategy; (b) Updating message βij with the ES strategy.](image)

where $s_{k,j}^{t}$ represents transmitted symbols corresponding to the chosen edges, and $s_{k,j}^{t}$ denotes transmitted symbols linked with the simplified edges. Note that the chosen edges are arbitrary $d_f - 1$ edges selected from $N_t - 1$ incoming edges, therefore the cardinality of $B(|A|, d_f)$ is

$$|B(|A|, d_f)| = \frac{(N_t - 1)}{d_f - 1} |A|^{d_f - 1}, \quad (18)$$

and the number of possible choices of transmitted symbols with this ES strategy is

$$\psi_{ES} = |A| \times \left(\frac{(N_t - 1)}{d_f - 1}\right) |A|^{d_f - 1}. \quad (19)$$

C. Initialization With PP Information

The BS-P detector requires the a priori information for each transmitted symbol to build up the configuration set. However, this a priori information is unavailable (set as the zero information) when the BS-P detection starts to update the message β. A natural way to obtain the initial a priori information is to perform one iteration of the BP detector, but this will cause an unacceptable computational complexity, which is inconsistent with the motivation of the proposed BS-P detector. Recently, many works [19], [28], [29] utilize the results of LMMSE detector to acquire the PP information of transmitted symbols, which improves the error rate performance of the data detection. Inspired by this, in this work, the proposed BS-P detector is also activated by using the results of LMMSE detector. The details are as below.

The LMMSE detector estimates transmitted symbols by

$$\hat{s}_{MMSE} = (H^H H + \sigma^2 \mathbf{I})^{-1} H^H y. \quad (20)$$

With the results of LMMSE detector, a priori probabilities of transmitted symbols are computed by

$$p_j(\mu_k) = \exp\left\{-\frac{||\mu_k - \hat{s}_{MMSE,j}||^2}{2\sigma^2_{MMSE,j}}\right\}, \quad (21)$$

where $\hat{s}_{MMSE,j}$ is the estimates of $j$-th ($j \in \{1, 2, \ldots, N_t\}$) symbol, and $\sigma^2_{MMSE,j}$ is the element of $j$-th row and $j$-th column in the noise covariance matrix $K$. The matrix $K$ is calculated by

$$K = (H^H H + \sigma^2 \mathbf{I})^{-1}. \quad (22)$$

Given the a priori probability for each transmitted symbol, the initial $\alpha$ messages are computed according to (5).

D. The Proposed BS-P Detector

Equipping with both ST and ES strategies, the BS-P detector is proposed here. Fig. 3 shows the message updating of $\beta_{ij}$ in the BS-P detector. Given $N_t - 1$ incoming edges, the blue lines denote the $N_t - d_f$ simplified edges, which are connected to the blue boxes containing the message $\alpha^m_{k,i}$, and the black lines represent the $d_f - 1$ chosen edges, which are linked with

![Fig. 3. Updating message βij in the BS-P detector.](image)
the black boxes involving the message $\alpha_{ji}^t$. Denote $\mathcal{B}(d_m, d_f)$ as the configuration set in the BsP detector, $\mathcal{B}(d_m, d_f)$ is computed by

$$\mathcal{B}(d_m, d_f) = \left\{ s_{k/d} = \left[ s_{k/d}^T, s_{k/d}^T \right]^T \right\}, \quad (23)$$

where $s_{k/d}^T$ denotes transmitted symbols of the chosen edges, calculated by

$$s_{k/d} = \left[ \mu^1_k, \mu^2_k, \ldots, \mu^{d_f-1}_k \right]^T; \quad \forall k \in \{1, 2, \ldots, m\} d_f-1. \quad (24)$$

With $\mathcal{B}(d_m, d_f)$, $\beta_{ij}$ in the BsP detector is updated as (25), shown at the bottom of the next page. The cardinality of $\mathcal{B}(d_m, d_f)$ is

$$|\mathcal{B}(d_m, d_f)| = \frac{(N_t - 1)}{d_f - 1}; \quad (26)$$

Therefore, the number of possible choices of transmitted symbols in the proposed BsP detector is

$$\psi_{\text{BsP}} = |A| \left( \frac{N_t - 1}{d_f - 1} \right) d_m, \quad (27)$$

Algorithm 1 BsP Detection

```
Input : received signal y, channel matrix H
Output: soft information of coded bits r
1 Initialization: $\alpha \leftarrow \text{InitMMSE}(y, H)${
2 for Iter $\leftarrow 1$ to MaxIterNum do {
3     for $i \leftarrow 1$ to $N_t$ do // compute $\beta$ messages {
4         for $j \leftarrow 1$ to $N_c$ do // compute \textbf{truncated} $\alpha$ messages {
5             sort symbol LLRs in $\alpha_{ji}$;
6             reserve the largest $d_m$ symbol LLRs to form $\alpha_{ji}^t$;
7             for $j \leftarrow 1$ to $N_t$ do {
8                 build $\mathcal{B}(d_m, d_f)$ as (23);\n9                 update $\beta_{ij}$ as (25);\n10            }
11         }
12     for $j \leftarrow 1$ to $N_t$ do // compute $\alpha$ messages {
13         update $\gamma_{ji}$ as (8);\n14         for $i \leftarrow 1$ to $N_c$ do {
15             $\alpha_{ji} \leftarrow \gamma_{ji} - \beta_{ij}$;\n16         }
17     }
18     compute r as (9);\n19 }
```

Algorithm 1 shows the BsP detection algorithm with the configuration set $\mathcal{B}(d_m, d_f)$. The \textit{a priori} information is initialized with an LMMSE detector. The details are as follows. Given the received signal $y$ and the channel matrix $H$, first, the LMMSE detector is utilized to initialize \textit{a priori} probabilities of transmitted symbols as (21). Then the initial $\alpha$ messages are acquired as (5). In order to update $\beta$ messages, for each $f_i$, symbol LLRs in each incoming $\alpha_{ji}$ ($j \in \{1, 2, \ldots, N_t\}$) are sorted to build up $\alpha_{ji}^t$. Note that $\alpha_{ji}^t$ in the message update of $\beta_{ij}$ ($j \in \{1, 2, \ldots, N_t\}$) for $f_i$ can be reused, therefore symbol LLRs of each incoming $\alpha$ only needs to be sorted once in an iteration. With $\alpha_{ji}^t$, $\mathcal{B}(d_m, d_f)$ is built up as (23), and $\beta_{ij}$ is then updated as (25). To update $\alpha_{ji}$, the incoming $\beta_{ij}$ ($i \in \{1, 2, \ldots, N_c\}$) for each $S_j$ are added together to acquire the overall symbol LLRs. Then $\alpha_{ji}$ is updated with

$$\alpha_{ji}(k) = \gamma_{ji}(k) - \beta_{ij}(k). \quad (28)$$

The BsP detection stops the message update when the number of the maximum iterations is achieved. Finally, it outputs bit LLRs computed as (9).

Empirically, since the BsP detector reduces the computational complexity by shrinking the search space of the transmitted symbol vector, it may suffer from performance degradation due to possibly losing optimal transmitted symbols in the search space. However, experiments in Section VI show different results. Simulation results demonstrate that the proposed BsP detector outperforms the BP detector, especially in the MIMO systems with high-order modulations. The reason is comprised of two parts. On the one hand, the proposed BsP detector only discards the transmitted symbols with relatively low reliability, which have little influence to results. On the other hand, the proposed BsP detector circumvents the propagation of the message with relatively small probability, mitigating the error propagation in BP detector, thus relieving the error floor performance.

It is noted that the truncation process in the proposed BsP is very similar to the procedure of sorting paths and preserving the paths with relatively large metrics in the $K$-best LSD [6]. But in fact there are several differences between the BsP and the $K$-best LSD. First, the $K$-best LSD is a tree search method and can not iteratively update the results, while the BsP detector is a message passing method based on the factor graph. Therefore, compared with the $K$-best LSD, the BsP detector supports soft-in/soft-out data detection and can constitute iterative receivers (e.g., IDD receiver [30]) in a natural. Second, the metrics together with the sorted objects are different. Specifically, the $K$-best LSD sorts all of the extended paths according to the path metrics and preserves the $K$-best paths, while the BsP detector sorts each incoming $\alpha$ messages according to the LLR values. Finally, the metrics for the sorting process have different meanings. The path metric in the $K$-best LSD denotes the Euclidean distance of the received signal and the estimated transmitted signal, while the LLR value in the BsP detector denotes the reliability of each transmitted symbols.

IV. COMPUTATIONAL COMPLEXITY

In this section, the computational complexity of the proposed BsP detector is analyzed. Then this complexity is compared with the complexity of the existing BP detectors, involving the BP detector, the EBRDF-BP detector [14], the PP-GF-BP detector [19], and the RD-GAI-BP detector [20].

A. Complexity Analysis of the BsP Detector

For fair comparison, the computational complexity for each detection algorithm is measured with the principle in [14].
This principle instructs that the computational complexity of the BP detection is dominated by the number of multiplication operations per channel use (deliver $MN_t$ transmission bits), because the multiplication operation demands much higher computational complexity than the addition and comparison operations. Therefore, the computational complexity of the proposed BsP detection is represented with the number of the multiplication operation in this section. For clarity, the computational complexity is analyzed with the following rules:

1) The multiplication operation results corresponding to $Hs$ are reused in different iterations.
2) One complex number multiplication operation requires four real number multiplication operations.

Since the proposed BsP detection with the configuration set $B(|A|, N_t)$ shares the same computational complexity with BP detector, the computational complexity of the proposed BsP detector can be derived from that of BP detector. According to [14], possible choices of $s$ in (6) is

$$\psi_{BP} = |A|^N_r.$$  \hspace{1cm} (29)

Let $\Psi_{BP}$ denote the number of the multiplications for BP detector, expressed as

$$\Psi_{BP} = 4|A|^{N_r}N_r N_t.$$  \hspace{1cm} (30)

Note that $\Psi_{BP}$ is irrelevant to the iteration number $Q_L$, because the computational results of $h_s$ are reused in different iterations. In the same way, the proposed BsP detector with $B(d_m, d_f)$ requires to search $\psi_{BsP}$ possible choices of $s$ to compute $h_s$ in (25). Therefore, the overall number of the multiplications of the proposed BsP is

$$\Psi_{BsP} = 4|A|^{N_t - 1} d_m^{d_f - 1} N_r N_t.$$  \hspace{1cm} (31)

For simplicity, the additional complexity of the initialization is ignored here. Note that, although the BsP detection requires the additional $d_m|A|N_r N_t Q_L$ comparison operations to sort symbol LLRs for each incoming $a_{jt}$, the computational complexity of comparison operations is still insignificant compared with the complexity of multiplication operations.

### B. Complexity Comparisons With The SOA BP Detectors

Table I shows the computational complexity of the proposed BsP detector and existing BP detectors. For simplicity, big $O$ representation is used in Table I. Several observations can be seen from the table. First, BP detector suffers from the highest computational complexity. This computational complexity is even comparable to that of the MAP detector. The computational complexity of EBRDF-BP detector is dominated by $d_f$. Second, PP-GF-BP detector and RD-GAI-BP detector circumvent the number of multiplications, but requiring exponentiation operations due to the GAI. Finally, the computational complexity of the BsP detector is

$$\chi_{BsP} = O(|A|^{N_t - 1} d_m^{d_f - 1} N_r N_t) + O(N_t^3),$$  \hspace{1cm} (32)

where $O(N_t^3)$ is the additional complexity of LMMSE detector. It is observed that $\chi_{BsP}$ is dominated by the parameters $d_m$ and $d_f$. Empirically, the proposed BsP detector with $B(2, 2)$ exhibits a good enough error rate performance, which is demonstrated in Section VI-A. Therefore, compared with BP detector and EBRDF-BP detector, the proposed BsP detector achieves a lower computational complexity. Compared with PP-GF-BP detector and RD-GAI-BP detector, the proposed BsP detector shows a comparable computational complexity.

Fig. 4 shows comparison results of the computational complexity vs. error rate performance for the proposed BsP detector and existing BP detectors in the MIMO system with $N_r = 8$, $N_t = 4$ and 16-QAM modulation. It is observed that BP detector suffers from the highest computational complexity while EBRDF-BP detector suffers from the worst BER performance. The proposed BsP detector with configuration set $B(2, 2)$ shows the best BER performance, although its computational complexity is slightly higher than that of RD-GAI-BP detector. The proposed BsP detector with configuration set $B(1, 1)$ enjoys the lowest computational complexity, and achieves a better error rate performance than RD-GAI-BP detector. In addition, PP-GF-BP detector is absent from the figure, due to its poor error rate performance. Fig. 4 shows that the proposed BsP detector can realize the trade-off between the computational complexity and the error rate performance by tuning parameters of the configuration set.

$$\beta_{ij}^{(l)}(k) = \max_{s_{k,j} \in B(d_m, d_f)} \left\{ -\frac{1}{2\sigma^2} \|y_i - h_s\|^2 + \sum_{t=1, t \neq j}^{N_t} \alpha_{yt}^{(l-1)} \right\} - \max_{s_{k,j} \in B(d_m, d_f)} \left\{ \frac{1}{2\sigma^2} \|y_i - h_s\|^2 + \sum_{t=1, t \neq j}^{N_t} \alpha_{yt}^{(l-1)} \right\}$$  \hspace{1cm} (25)
codes \[31\], \[32\] and polar codes \[33\], exhibit good results employed in error correction codes (ECC), e.g., LDPC on LLR messages. For instance, the BP decoding algorithm ically derive the converged results for BP algorithms based

**TABLE I**

| Operations     | Multiplications | Additions                                                                 | Comparisons                        | Exponentiations |
|----------------|-----------------|---------------------------------------------------------------------------|------------------------------------|-----------------|
| BP [14]        | \(O(N^2N_1N_2)\) | \(O([4]^N_1N_2)+[4]^N_1N_2Q_L\)                                       | \(O([4]^N_1N_2Q_L)\)               | 0               |
| EBRDF-BP [14]  | \(O([4]^2d_1N_2)\) | \(O([4]^2d_1N_2)+[4]^2d_1N_2Q_L\)                                      | \(O([4]^2d_1N_2Q_L)\)             | 0               |
| PP-GF-BP [19]  | \(O([4]^N_1N_2Q_L)+O(N_f^2)\) | ★ \(O([4]^N_1N_2Q_L)+O(N_f^2)\)                                       | 0                                  | \(O([4]^N_1N_2Q_L)\) |
| RD-GAI-BP [20] | \(O(\sqrt{4N_1N_2Q_L})\) | \(O(\sqrt[4]{4N_1N_2Q_L})\)                                           | 0                                  | \(O(\sqrt[4]{4N_1N_2Q_L})\) |
| Proposed BsP†  | \(O(\sqrt{4N_1N_2Q_L})\) | ★ \(O(\sqrt[4]{4N_1N_2Q_L})\)                                       | 0                                  | \(O(\sqrt[4]{4N_1N_2Q_L})\) |

†: Initialized with an LMMSE detector.

**V. CONVERGENCE ANALYSIS**

To the authors’ best knowledge, it is intractable to theoretically derive the converged results for BP algorithms based on LLR messages. For instance, the BP decoding algorithm employed in error correction codes (ECC), e.g., LDPC and polar codes [33], exhibit good results but their theoretical convergence analysis is still unavailable. One approximation for this issue is the EXIT chart proposed by Prof. Stephan ten Brink [34], which can empirically track the convergence for iterative systems. In this paper, the EXIT chart is utilized to characterize the convergence behavior of the proposed BsP algorithm and other existing detection algorithms. It is noted that the motivation for employing the EXIT chart is to predict the performance of different detection algorithms. The details of the EXIT chat analysis are provided below.

**A. Convergence Performance Comparisons**

The EXIT chart is usually employed to track the extrinsic mutual information in different modules (e.g., detection module and decoding module) for coded systems. For the proposed BsP served in uncoded systems, the feedback from the decoder is unavailable, but the EXIT chart can still be employed here by computing the ergodic average mutual information (AMI) of the proposed BsP. The AMI between the a source bit and the corresponding transmitted symbols are recovered by the proposed BsP detector, and the corresponding \(I_{out}^C\) is available. This method is also suitable for the compared detectors to compute their AMI results. In the same scenario, a higher AMI result indicates that the detector enjoys better error performance.

Fig. 5 shows the EXIT chart of the different detectors in 8×4 MIMO system with 256-QAM modulation. The number of iterations for all of the BP detectors is set as \(Q_L = 10\). It can be seen that the proposed BsP with \(B(1, 1)\) enjoys a higher AMI than the LMMSE throughout the simulated \(E_b/N_0\) regions, which predicts that the proposed BsP with \(B(1, 1)\) would outperform the LMMSE in terms of the BER performance. Compared with the EBRDF, \(d_f = 2\) [14] and the RD-GAI-BP [20], the proposed BsP with \(B(1, 1)\) suffers from the lower AMI in small \(E_b/N_0\) regions, but achieves the higher AMI in relatively large \(E_b/N_0\) regions. Therefore, it is predicted that the proposed BsP with \(B(1, 1)\) would have performance loss in small \(E_b/N_0\) regions compared with the EBRDF, \(d_f = 2\) and the RD-GAI-BP, but outperform them in large \(E_b/N_0\) regions. In addition, with a larger configuration set \(B(2, 2)\), the proposed BsP achieves a higher AMI than the BsP with \(B(1, 1)\), which predicts that the proposed BsP with \(B(2, 2)\) can enjoy better performance than the proposed BsP with \(B(1, 1)\). In the following section, numerical results in the same MIMO scenario (Fig. 9) will be provided to demonstrate the convergence of the proposed BsP.

Fig. 6 further shows the numerical results with respect to the convergence process of BER vs. \(Q_L\) for the proposed BsP and the other detectors in the 8×4, 256-QAM MIMO scenario. For a fair comparison, the RD-GAI-BP is initialized with LMMSE as well. The target \(E_b/N_0\) is set as 22 dB to reach the BER of \(10^{-3}\). As shown in Fig. 6, the proposed BsP detection algorithms with both \(B(1, 1)\) and \(B(2, 2)\) show convergence after \(Q_L = 10\). Compared to the RD-GAI-BP, the proposed BsP detection algorithms show great performance advances at \(Q_L = 1\). This is because the RD-GAI-BP suffers from aggravated error floor here. With \(Q_L\) increasing, the performance gaps degrade and finally remain stable till \(Q_L = 10\).

**B. Case Study: Belief-Selective Technique Relieves the Error Floor**

The goal of this subsection is to discuss the reason why the proposed BsP could mitigate the error performance compared with the RD-GAI-BP. To this end, a theoretical analysis of the error performance for the RD-GAI-BP is first derived. Then, statistical error cases are provided to demonstrate the error floor performance of the RD-GAI-BP and illustrate the major components leading to the error floor. Finally, a corollary is proposed to prove that the proposed BsP can circumvent the error cases above, thus relieving the error floor performance.

1) **Error Performance Analysis of RD-GAI-BP:** To analyze the error performance of the RD-GAI-BP, the expressions of RD-GAI-BP are provided first. For RD-GAI-BP, the received
signal \( y_i \) can be written as

\[
y_i = h_{ij}s_j + \sum_{k=1, k \neq j}^{2N_t} h_{ik}s_k + n_i = h_{ij}s_j + z_{ij} + n_i,
\]

where \( z_{ij} \) is the interference and can be approximated by a Gaussian distribution \( \mathcal{N}(\epsilon_{ij}, \sigma_{ij}^2) \) with

\[
\begin{align*}
\epsilon_{ij} &= \sum_{k=1, k \neq j}^{2N_t} h_{ik}E[s_k], \\
\sigma_{ij}^2 &= \sum_{k=1, k \neq j}^{2N_t} h_{ik}^2 \text{Var}(s_k) + \sigma_{n_i}^2.
\end{align*}
\]

\( E[s_k] \) and \( \text{Var}(s_k) \) denote the mean and the variance of the \( k \)-th transmitted symbol. The likelihood with respect to symbol \( s_j \) can be written as

\[
P_{s_j}(y_i|h_i) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left\{ -\frac{(y_i - \epsilon_{ij} - h_{ij}s_j)^2}{2\sigma_{ij}^2} \right\},
\]

In order to achieve the maximum likelihood probability in (36), the estimated transmitted symbol \( \hat{s}_j \) must satisfy

\[
\hat{s}_j = \arg\min_{s_j \in \mathcal{A}} |s_{e_{ij}} - s_j|,
\]

where \( s_{e_{ij}} = \frac{y_i - \epsilon_{ij} - n_i}{h_{ij}} \) denotes the calculated point through (36). Define \( \xi_j \) as the deviation between \( s_{e_{ij}} \) and \( s_j \), and it can be written as

\[
\xi_j = s_{e_{ij}} - s_j = \frac{z_{ij} - \epsilon_{ij}}{h_{ij}}.
\]

For the MIMO scenario employing 16-QAM modulation, \( s_j \) is chosen from the real-domain constellation with \( \sqrt{|A|} = \{-0.3162, -0.9487, +0.3162, +0.9487\} \). Fig. 7 shows the cases in which the incorrect \( \hat{s}_j \) is determined from the calculated point \( s_{e_{ij}} \) according to (39). It can be seen that given \( s_j = -0.9487, \hat{s}_j \) would be determined to be an incorrect transmitted symbol if \( \xi_j > 0.3162 \) (error case 01). Given \( s_j = 0.9487, \hat{s}_j \) would be determined to be an incorrect transmitted symbol if \( \xi_j < 0.3162 \) (error case 02). Given \( s_j = (\pm)0.3162, \hat{s}_j \) would be determined to be an incorrect transmitted symbol and the symbol LLR of RD-GAI-BP passing from the \( i \)-th FN to the \( j \)-th SN is computed as

\[
\beta_{ij}(m) = \ln \frac{P_{s_j=m}(y_j|s)}{P_{s_j=m}(y_j|\bar{s})} = \frac{-(y_i - \epsilon_{ij} - h_{ij}s_j)^2}{2\sigma_{ij}^2} - \frac{-(y_i - \epsilon_{ij} - h_{ij}m)^2}{2\sigma_{ij}^2},
\]

(37)

where \( \mu_m \) denotes the \( m \)-th symbol in the real-domain constellation.

It can be derived from (34) that the transmitted symbol \( s_j \) is

\[
s_j = \frac{y_i - \epsilon_{ij} - n_i}{h_{ij}}.
\]

(38)
if detector. A fair comparison, the PP initialization is also employed for the RD-GAI-BP.

Fig. 8. BER comparison results of the proposed BsP detector and RD–GAI-BP in the 8 × 4 MIMO scenario with 16-QAM modulation. For a fair comparison, the PP initialization is also employed for the RD-GAI-BP detector.

\[ \text{if } |\xi_j| > 0.3162 \text{ (error case 03). Therefore, with 16-QAM modulation, } \hat{s}_j \text{ for RD-GAI-BP according to (39) can be listed as} \]

\[ \begin{cases} \hat{s}_j = s_j, & \text{if } |\xi_j| < 0.3162, \\ \hat{s}_j \neq s_j, & \text{if } \begin{cases} s_j = -0.9487, & \text{AND } \xi_j > 0.3162, \\ s_j = (\pm)0.3162, & \text{AND } |\xi_j| > 0.3162, \\ s_j = 0.9487, & \text{AND } \xi_j < 0.3162. \end{cases} \end{cases} \]

(41)

It can be concluded that the error cases are majorly caused by the deviation \( \xi_j \). If \( \xi_j \) could be eliminated, the error performance of the RD-GAI-BP will be greatly improved.

2) Statistical Error Cases for RD-GAI-BP: Fig. 8 demonstrates the BER performance of the RD-GAI-BP and the proposed BsP with \( B(1, 1) \) in 8 × 4 MIMO system with 16-QAM modulation. It is observed that RD-GAI-BP suffers from an aggravated error floor at the high \( E_b/N_0 \) region, whereas the proposed BsP mitigates the error floor. Table II shows the statistical error cases of the RD-GAI-BP and the proposed BsP with \( B(1, 1) \) in the same MIMO scenario as Fig. 8. For a fair comparison, the simulated algorithms are set to receive the same samples (y and H) per channel use. From Table II it is observed that even with the reliable initial information in Case 02, the RD-GAI-BP still produces a relatively large number of error bits. Therefore, it is concluded that the error floor performance of the RD-GAI-BP is caused by the error bits in the case with reliable initial information.

3) Corollary of the Proposed BsP: From the error performance analysis and the statistical error cases above, the reason why the proposed BsP can mitigate the error floor performance could be illustrated as the following corollary.

**Corollary 1:** If the initial priori information is reliable, the proposed BsP would eliminate the deviation \( \xi_j \) in (40), and reduce the error bits, thus mitigating the error floor performance.

**Proof:** Let \( P_k(\eta) \) denote the maximum priori probability for \( s_k \). Since the configuration set \( B(1, 1) \) is the subset of the any \( B(d_m, d_f) \) with \( d_m > 1 \) and \( d_f > 1 \), the proposed BsP with any \( B(d_m, d_f) \) will converge to the same results as that with \( B(1, 1) \) if the initial information is reliable. For the proposed BsP with \( B(1, 1) \) (mapped into the real-domain), (25) can be rewritten as

\[ \begin{align*}
\beta_{ij}(m) &= \frac{-(y_i - \sum_{k=1, k \neq j}^{2N_t} h_{ik} P_k(\eta) \mu_k - h_{ij} \mu_{m})^2}{2\sigma_{ni}^2} \\
& \quad - \frac{-(y_i - \sum_{k=1}^{2N_t} h_{ik} P_k(\eta) \mu_k - h_{ij} \mu_1)^2}{2\sigma_{ni}^2},
\end{align*} \]

(42)

where the maximum priori probability is set as \( P_k(\eta) = 1 \). If the initial priori information for each \( s_k \) is reliable, it is derived that \( \sum_{k=1, k \neq j}^{2N_t} h_{ik} P_k(\eta) \mu_k = z_{ij} \) and \( \sigma_{ij}^2 \) is equal to \( \sigma_{ni}^2 \). Hence the proposed BsP achieves the same results as the RD-GAI-BP with (37) and \( \xi_j = 0 \), greatly enhancing the error performance of the RD-GAI-BP in this case.

**VI. NUMERICAL RESULTS**

In this section, numerical results are provided to demonstrate the error rate performance of the proposed BsP detector. In particular, the proposed BsP detector is first manifested in the medium-scale MIMO scenarios with \( N_r = 8, N_t = 4 \) and high-order modulations. Then, the performance of the proposed BsP detector is investigated in the uplink massive...
MU-MIMO scenarios with 256-QAM modulation, involving a base station (BS) equipped $N_r$ antennas and $N_t$ single-antenna user ends (UEs). Finally, the proposed BsP detector is studied in the massive MU-MIMO orthogonal frequency-division multiplexing (OFDM) system with the least square (LS) channel estimation, to validate its performance in a more realistic system. Without special explanation, the MIMO channels follow the Rayleigh fading channel model, and channel coding is not involved in the MIMO systems. The BsP detector is compared with the EBRDF-BP detector [14], and the RD-GAI-BP detector [20] (the SOA BP detector). The GAI-BP detector [18] and the PP-GF-BP detector [19] are not considered due to their aggravated performance error floor in MIMO systems with high-order modulations. The BP detector [14] is also omitted in the MIMO systems with high-order modulations due to its prohibitive computational complexity. In addition, the near-optimal sphere detection (SD) [5] is employed as the benchmark, to see how far the result between the proposed BsP detector and the near-optimal detector. For fair comparison, RD-GAI-BP detector also employs the initial information provided by LMMSE.

A. Performance in Medium-Scale MIMO System

Fig. 9 shows the BER performance of the proposed BsP detector in $8 \times 4$ MIMO scenarios with different high-order modulations. The number of the iterations for the proposed BsP, the EBRDF-BP [14], the BP [14], and the RD-GAI-BP [20] are consistently set as $Q_L = 10$. Several observations can be drawn from this figure. First, compared with the BP detector in the MIMO scenario with $16$-QAM modulation, the proposed BsP detector with $B(1, 1)$ earns about $0.25$ dB performance gain at BER of $10^{-3}$, and the proposed BsP detector with $B(2, 2)$ enjoys about $1.15$ dB performance gain at the same BER. Second, compared with the RD-GAI-BP detector, the proposed BsP detector suffers from performance loss in low $E_b/N_0$ regions, but enjoys better BER performance in high $E_b/N_0$ regions. In particular, in the MIMO scenario with $64$-QAM modulation, the proposed BsP detector with $B(1, 1)$ and $B(2, 2)$ outperforms the RD-GAI-BP detector at $E_b/N_0 > 19$ dB and $E_b/N_0 > 15$ dB, respectively. Finally, compared to the near-optimal SD, the proposed BsP detector with $B(1, 1)$ in the MIMO scenario with $16$-QAM modulation shows about $1.6$ dB $E_b/N_0$ loss at BER of $10^{-3}$. The proposed BsP detector with $B(2, 2)$ in the same MIMO scenario has about $0.8$ dB $E_b/N_0$ loss at BER of $10^{-3}$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Besides, the proposed BsP detector suffers from larger performance gaps compared with the SD in the MIMO scenarios with higher-order modulations. However, although the BsP detector shows performance degradation with a higher-order modulation, it still holds the waterfall curves in relatively high $E_b/N_0$ regions.

Fig. 10 manifests the BER performance of the proposed BsP detector with different configuration sets in the MIMO scenario with $N_r = 8$, $N_t = 4$ and 16-QAM modulation. Empirically, the proposed BsP detector with the configuration set involving a large $d_m$ and $d_f$ will achieve a better BER performance, due to its large search space of the transmitted symbols. However, Fig. 10 exhibits different experimental results. It is observed that the configuration set $B(2, 2)$ contributes to 1.2 dB performance gain for the proposed BsP detector at the BER of $10^{-4}$, compared with the configuration set $B(1, 1)$. This performance gain is consistent with empirical recognition. However, compared with the configuration set $B(2, 2)$, the configuration sets with large $d_m$ and $d_f$ fail in further bringing performance gains to the proposed BsP detector. The reason is straightforward: since the configuration set $B(d_m, d_f)$ consists of the transmitted symbol vector with the relatively large probability, a larger $d_m$ and $d_f$ will make little contribution to $B(d_m, d_f)$ for involving the optimal transmitted symbol vector, thus resulting in a limited performance gain.

B. Performance in Massive MU-MIMO System

Fig. 11 shows the BER performance of the proposed BsP detector with $B(1, 1)$ in $64 \times 32$ and $128 \times 64$ uplink MIMO scenarios. To meet the requirements of the high data transmission rate in 5G [35], high-order modulation of 256-QAM is employed. The proposed BsP detector with $B(2, 2)$ is omitted here due to its high computational complexity in high-dimensional MIMO systems. In addition to the RD-GAI-BP detector, the proposed BsP detector is compared to $K$-best LSD [6], which involves a very similar procedure of sorting paths and keeping the paths with relatively large reliability. For fairness, the proposed BsP detector with $B(1, 1)$ is compared to the 1-best LSD. In addition, the often-used 8-best LSD is also provided here. The number of iterations for the proposed BsP and RD-GAI-BP is set as $Q_L = 5$. From Fig. 11(a) it can be seen that the proposed BsP detector with $B(1, 1)$ respectively outperforms the RD-GAI-BP detector and the 1-best LSD by about $1.8$ and $1.1$ dB at the BER of $10^{-3}$. Compared to the 8-best LSD and the near-optimal SD, the proposed BsP detector with $B(1, 1)$ suffers from performance loss in relatively low $E_b/N_0$ regions, but achieves the comparable performance in high $E_b/N_0$ regions. Fig. 11(b) shows the same trends as that in Fig. 11(a) but with several differences. In particular, compared to the 8-best LSD and the
near-optimal SD, the proposed BsP detector with $B(1, 1)$ converges to the comparable result more quickly, and the performance gain between the proposed BsP detector with $B(1, 1)$ and the RD-GAI-BP detector degrades. Furthermore, it is worthwhile to notice that the proposed BsP exhibits worse heavy error propagation at low $E_b/N_0$ regions, although it employs the initial information provided by the LMMSE. The reason is that the proposed BsP suffers from heavy error propagation at low $E_b/N_0$ if the LMMSE fails and provides unreliable initial information.

C. Performance in MU-MIMO-OFDM System

Fig. 12 shows the BER performance of the proposed BsP detector in 15-MHz MU-MIMO-OFDM uplink system with $N_r = 32$, $N_t = 16$, and 4096 subcarriers. The modulation scheme is 256-QAM, and the MIMO channels follow an Rx-correlated Rayleigh channel model with a correlation ratio of $\rho_R = 0.5$ here. The LS channel estimation is used to acquire the CSI before the data detection. It can be seen that the proposed BsP detector with $B(1, 1)$ outperforms the LMMSE detector about 2.3 dB at the BER of $10^{-3}$. Compared to the 8-best LSD, the proposed BsP detector with $B(1, 1)$ has about 0.3 dB $E_b/N_0$ loss at the BER of $10^{-3}$. Compared to the near-optimal SD, the proposed BsP detector with $B(1, 1)$ suffers from about 0.8 dB $E_b/N_0$ loss at the same BER as $10^{-3}$.

VII. CONCLUSION

In this paper, the BsP MIMO detector is proposed to balance performance and complexity. By exploiting the symbol probabilities to truncate the search space of the symbol candidates, the BsP detector enjoys a lower complexity than the SOA BP detector. This approach also helps to mitigate the performance error floor of the SOA BP detector. Numerical results have shown that with the proposed simplifying mechanisms, the BsP detector outperforms its counterparts in terms of both performance and complexity. It should be noted that the proposed BsP detector is suitable for message-passing problems with large candidate sets, including but not limited to MIMO detection. Future work will be directed towards joint iterative cooperation with other baseband modules, the efficient hardware implementations of the BsP MIMO detector, and other applications of BsP in a generalized fashion.

ACKNOWLEDGMENT

The authors thank Prof. Xiaosi Tan for useful discussions.

REFERENCES

[1] J. Hoydis, S. T. Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?” IEEE J. Sel. Areas Commun., vol. 31, no. 2, pp. 160–171, Feb. 2013.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” IEEE J. Sel. Topics Signal Process., vol. 8, no. 5, pp. 742–758, Oct. 2014.

[3] F. Rusek et al., “Scaling up MIMO: Opportunities and challenges with very large arrays,” IEEE Signal Process. Mag., vol. 30, no. 1, pp. 40–60, Jan. 2013.

[4] O. Damen, A. Chkeif, and J.-C. Belfiore, “Lattice code decoder for space-time codes,” IEEE Commun. Lett., vol. 4, no. 5, pp. 161–163, May 2000.

[5] B. M. Hochwald and S. T. Brink, “Achieving near-capacity on a multiple-antenna channel,” IEEE Trans. Commun., vol. 51, no. 3, pp. 380–399, Mar. 2003.

[6] Z. Guo and P. Nilsson, “Algorithm and implementation of the K-best sphere decoding for MIMO detection,” IEEE J. Sel. Areas Commun., vol. 24, no. 3, pp. 491–503, Mar. 2006.

[7] B. Hassibi and H. Vikalo, “On the sphere-decoding algorithm I. Expected complexity,” IEEE Trans. Signal Process., vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[8] S. Yang and L. Hanzo, “ Fifty years of MIMO detection: The road to large-scale MIMO,” IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 1941–1988, 4th Quart., 2015.

[9] T. L. Narasimhan and A. Chockalingam, “Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems,” IEEE J. Sel. Topics Signal Process., vol. 8, no. 5, pp. 847–860, Oct. 2014.

[10] Y. Zhang, Z. Wu, C. Li, Z. Zhang, X. You, and C. Zhang, “Expectation propagation detection with Neumann-series approximation for massive MIMO,” in Proc. IEEE Int. Workshop Signal Process. Syst. (SIPS), Oct. 2018, pp. 59–64.

[11] Z. Wu, C. Zhang, Y. Xue, S. Xu, and X. You, “Efficient architecture for soft-output massive MIMO detection with Gauss–Seidel method,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2016, pp. 1886–1889.

[12] Minimum Requirements Related to Technical Performance for IMT2020 Radio Interface(s), document ITU-R M.2410–0, ITU-R M.[IMT2020-TECH PERF REQ], International Telecommunication Union Recommendations, Nov. 2017.

[13] Study on New Radio (NR) Access Technology Physical Layer Aspects, document TR 38.802, 3GPP, Mar. 2017.

[14] 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.

[15] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, San Mateo, CA, USA: Morgan Kaufmann, 1988.

[16] 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.

[17] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” Bell Labs Tech. J., vol. 1, no. 2, pp. 41–59, Aut. 1996.

[18] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, “Low-complexity detection in large-dimension MIMO-CSI channels using graphical models,” IEEE J. Sel. Topics Signal Process., vol. 5, no. 8, pp. 1497–1511, Dec. 2011.

[19] F. Long, T. Lv, R. Cao, and H. Gao, “Single edge based belief propagation algorithms for MIMO detection,” in Proc. 34th IEEE Sarnoff Symp., May 2011, pp. 1–5.

[20] J. Yang, W. Song, S. Zhang, X. You, and C. Zhang, “Low-complexity belief propagation detection for correlated large-scale MIMO systems,” J. Signal Process. Syst., vol. 90, no. 4, pp. 585–599, Apr. 2018.

[21] H. Wyneersch, J. Lien, and M. Z. Win, “Cooperative localization in wireless networks,” Proc. IEEE, vol. 97, no. 2, pp. 427–450, Feb. 2009.

[22] B. Li and N. Wu, “Convergence analysis of Gaussian SPAWN under high-order graphical models,” IEEE Signal Process. Lett., vol. 27, no. 12, pp. 1725–1729, 2020.

[23] J. M. Mooij and H. J. Kappen, “Sufficient conditions for convergence of the sum–product algorithm,” IEEE Trans. Inf. Theory, vol. 53, no. 12, pp. 4422–4437, Dec. 2007.

[24] D. Declercq and M. Fossorier, “Extended minsum algorithm for decoding LDPC codes over GF(q),” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Sep. 2005, pp. 464–468.

[25] D. Declercq and M. Fossorier, “Decoding algorithms for nonbinary LDPC codes over GF(q),” IEEE Trans. Commun., vol. 55, no. 4, pp. 633–643, Apr. 2007.

[26] A. Voicila, D. Declercq, F. Verdier, M. Fossorier, and P. Urard, “Low-complexity decoding for non-binary LDPC codes in high order fields,” IEEE Trans. Commun., vol. 58, no. 5, pp. 1365–1373, May 2010.

[27] M. Nassar, P. Schneiter, and B. L. Evans, “A factor graph approach to joint OFDM channel estimation and decoding in impulsive noise environments,” IEEE Trans. Signal Process., vol. 62, no. 6, pp. 1576–1589, Mar. 2014.

[28] J. Goldberger and A. Leshem, “Pseudo prior belief propagation for LDPC codes over GF(q) in noisy channels,” IEEE Trans. Commun., vol. 57, no. 8, pp. 4973–4982, Aug. 2009.
or coauthored more than 200 peer-reviewed articles and holds more than 80 invention patents. His research interests include intelligent 5G/6G mobile communications and millimeter wave wireless communications. He submitted around 20 technical contributions to IEEE standards, and was awarded the Certificate of Appreciation for outstanding contribution to the development of IEEE standard 802.11aj. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is also an Editor-at-Large of the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY and an Associate Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS.

Chuan Zhang (Senior Member, IEEE) received the B.E. degree in microelectronics and the M.E. degree in very-large scale integration (VLSI) design from Nanjing University, Nanjing, China, in 2006 and 2009, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Minnesota, Twin Cities (UMN), USA, in 2012. He is currently a Young Chair Professor with Southeast University. He is also with the LEADS, National Mobile Communications Research Laboratory, Frontiers Science Center for Mobile Information Communications and Security of MoE, Quantum Information Center, Southeast University, and Purple Mountain Laboratories, Nanjing. His current research interests include algorithms and implementations for signal processing and communication systems.

Dr. Zhang is also a member of the Applied Signal Processing Systems TC of the IEEE Signal Processing Society, the VLSI Systems and Communications TC, the VLSI System and Applications TC, and the Digital Signal Processing TC of the IEEE Circuits and Systems Society. He received the Best Contribution Award of the IEEE Asia Pacific Conference on Circuits and Systems (APCCAS) in 2018, the Best Paper Award in 2016, the Best (Student) Paper Award of the IEEE International Conference on DSP in 2016, three Best (Student) Paper Awards of the IEEE International Conference on ASIC in 2015, 2017, and 2019, the Best Paper Award Nomination of the IEEE Workshop on Signal Processing Systems in 2015, three Excellent Paper Awards and two Excellent Poster Presentation Awards of the International Collaboration Symposium on Information Production and Systems from 2016 to 2018, the Outstanding Achievement Award of the Intel Collaborative Research Institute in 2018, and the Merit (Student) Paper Award of the IEEE APCCAS in 2008. He also received the Three-Year University-Wide Graduate School Fellowship of UMN and the Doctoral Dissertation Fellowship of UMN. He serves as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS and the Corresponding Guest Editor for the IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS twice. He is a Distinguished Lecturer and the Chair Elect of the Circuits and Systems for Communications TC of the IEEE Circuits and Systems Society.

Xiaohu You (Fellow, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from Southeast University, Nanjing, China, in 1985 and 1988, respectively. Since 1990, he has been with the National Mobile Communications Research Laboratory, Southeast University, where he is currently the Director and a Full Professor. He has long been engaged in teaching, scientific research, and organizational management of mobile communication systems, and made important contributions to the development of China’s 3G, 4G, and 5G mobile communications. From 1999 to 2002, he was a Principal Expert of the C3G Project, responsible for organizing China’s 3G Mobile Communications Research and Development Activities. From 2001 to 2006, he was a Principal Expert of the China National 863 Beyond 3G FuTURE Project. Since 2013, he has been a Principal Investigator of the China National 863 5G Project. As the first recipient, he received the 1st Class National Technological Invention Award, the 2nd Class National Science and Technology Progress Award, the Tan Kah Kee Science Award, and other important awards.