On the Improved Path Metric for Soft-input Soft-output Tree Detection

Jun Won Choi*, Byonghyo Shim**, and Andrew C. Singer*
*University of Illinois at Urbana-Champaign
** Korea University

***Abstract*—In this paper, we propose a new path metric, which improves the performance of soft-input soft-output (SISO) tree detection for iterative detection and decoding (IDD) systems. While the conventional path metric accounts for the contribution of symbols on a visited path due to the causal nature of tree search, the new path metric, called *improved path metric*, reflect the contribution of unvisited paths using an unconstrained minimum mean squared error (MMSE) estimate of undecided symbols. The improved path metric is applied to SISO M-algorithm, which finds a list of symbol candidates based on breadth-first search strategy and computes *a posteriori* probability of each entry of the symbol vector. We study the probability of correct path loss (CPL) for the improved path metric and confirm the performance improvement over the conventional path metric.**

I. Introduction

The complex-domain relationship between the transmitted symbol and received signal vector in many communication systems can be expressed as

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

where $\mathbf{x}$ is the transmitted vector whose entries are chosen from a finite symbol alphabet, $\mathbf{y}$ and $\mathbf{n}$ are the received signal and noise vectors, respectively, and $\mathbf{H}$ is a complex channel matrix. Basically, *a posteriori probability* (APP) detection is formulated as obtaining *a posteriori* probabilities on the bits comprising $\mathbf{x}$ based on *a priori* probabilities on them and the observed vector $\mathbf{y}$. The APP detector is a key component of *iterative detection and decoding* (IDD) system, where a form of soft information is exchanged between a soft-input soft-output (SISO) symbol detector and channel decoder to achieve bit error rate (BER) improvement [1].

Direct computation of APP involves marginalization over all possible candidates of $\mathbf{x}$, incurring significant amount of complexity for large systems. In order to reduce the complexity for the APP detection, tree detection techniques [1]–[4] have been popularly used, which, in essence, limit the set of symbol candidates over which the APP is estimated. In [1], a list of multiple symbol candidates with high likelihood functions are found by a fixed-radius sphere search and then the APP is estimated by counting the symbol candidates in the list. In [3], a hard sphere decoder is employed to find a *maximum a posteriori* (MAP) symbol estimate, and a candidate list is generated by bit-flipping of the MAP estimate. An approach for computing the APP of all bits comprising the symbol vector through a single search is proposed in [4], and more sophisticated extension of this idea is presented in [5]. While the complexity of these APP detectors depends on noise and channel realizations, there have been other algorithms that offer fixed complexity. In [6], the M-algorithm was used to find a fixed size candidate list and in [7], the stack algorithm was combined with soft augmentation of tail bits for list generation. Other fixed complexity SISO detection methods include [8] and [9].

A search space spanned by the symbol vector $\mathbf{x}$ can be represented by a tree, where each complete path corresponds to a realization of $\mathbf{x}$. In a nutshell, a goal of tree detection algorithms is to find the complete paths with the minimum cost metric. Towards this end, path metric is assigned to each node, which contains partial information on the cost metric. Owing to the causal nature of the search, the conventional path metric [1] is determined by contributions of the visited path on the cost metric only. In this case, the path metric in the upper level of tree considers only a few symbol spans so that it does not reflect the validity of a path, causing erroneous decisions or backtracking operations often. This property of the conventional path metric degrades the efficiency of tree search.

There were several different path metrics used for tree search. A *Fano metric*, proposed for sequential decoding, introduces a bias term proportional to tree depth to penalize short paths [10]. Similar bias term is derived for equalization of intersymbol interference (ISI) channel in [11] and for multi-input multi-output (MIMO) detection in [12]. In [7], a probability density of received vector was used as a bias term to improve the efficiency of stack algorithm.

In this paper, we propose a new path metric for improving a sorting (or pruning) process of the tree search. This path metric, referred to as *improved path metric* looks ahead the contribution of unvisited paths using soft unconstrained linear symbol estimates to provides stronger information on the visited path. By applying the improved path metric to the SISO M-algorithm, more promising symbol candidates can be selected, which eventually brings the improvement in detection performance. In this paper, we particularly focus on the analysis of correct path loss (CPL) probability to demonstrate the performance gain achieved by the improved path metric.

II. Problem Description

In this section, we first describe the IDD system. Then, we will provide a brief introduction of SISO tree detection algorithm.

A. Description of IDD System

The rate R_c recursive systematic convolutional (RSC) encoder is used to convert a sequence of *i.i.d.* binary bits $\{b_i\}$ to a sequence of coded bits $\{c_i\}$. The bit sequence $\{c_i\}$ is permuted using a random interleaver and Q interleaved bits are mapped to a finite alphabet symbol. These symbols comprise the symbol vector $\mathbf{x} = [x_1, \cdots, x_N]^T$. We label the interleaved bits associated with x_k as $\bar{c}_{k,1}, \cdots, \bar{c}_{k,Q}$. Due to the interleaving, we can assume that the entries of symbol vector are uncorrelated with each other. The vector of received signal can be expressed as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{y}$ and $\mathbf{n}$ are the $L \times 1$ received signal and noise vectors, respectively, and $\mathbf{H}$ is a $L \times N$ complex channel matrix.

Given the observation $\mathbf{y}$ and *a priori* probabilistic knowledge on $\mathbf{x}$, the SISO symbol detector computes the *a posteriori* log-likelihood ratio (LLR) of $\bar{c}_{k,i}$, i.e., $L_{\text{post}}(\bar{c}_{k,i}) = \ln Pr(\bar{c}_{k,i} = +1|\mathbf{y}) - \ln Pr(\bar{c}_{k,i} = -1|\mathbf{y})$. Based on the Gaussian noise model $\mathbf{n} \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$, we have

$$L_{\text{post}}(\bar{c}_{k,i}) = \ln \frac{\sum_{\mathbf{x}\in X_{k,i}^{+1}} \exp(\lambda(\mathbf{x}))}{\sum_{\mathbf{x}\in X_{k,i}^{-1}} \exp(\lambda(\mathbf{x}))} \tag{2}$$

where

$$\begin{aligned} \lambda(\mathbf{x}) &= -\frac{1}{\sigma_n^2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \sum_{i=1}^{N}\sum_{j=1}^{Q} \ln Pr(\bar{c}_{i,j}), \\ Pr(\bar{c}_{i,j}) &= \frac{1}{2}\left(1 + \bar{c}_{i,j}\tanh\left(\frac{L_{\text{pri}}(\bar{c}_{i,j})}{2}\right)\right), \end{aligned} \tag{3}$$

where the set $X_{k,i}^{\pm 1}$ is the set of all symbol combinations satisfying $\bar{c}_{k,i} = \pm 1$, and $L_{\text{pri}}(\bar{c}_{k,i})$ is *a priori* LLR defined as $L_{\text{pri}}(\bar{c}_{k,i}) = \ln Pr(\bar{c}_{k,i} = +1) - \ln Pr(\bar{c}_{k,i} = -1)$. Once $L_{\text{post}}(\bar{c}_{k,i})$ is found, the extrinsic LLR is computed using $L_{\text{ext}}(\bar{c}_{k,i}) = L_{\text{post}}(\bar{c}_{k,i}) - L_{\text{pri}}(\bar{c}_{k,i})$. Then, the extrinsic LLRs of all bits contained in a processing block are delivered to the channel decoder as *a priori* LLRs. Under the IDD framework, the exchange of these extrinsic LLRs is carried out repeatedly until the performance ceases to improve.

B. SISO Tree Detection Algorithm

Since a tree search algorithm searches for the symbol vectors minimizing $-\lambda(\mathbf{x})$, we will call the function $d_{\text{APP}}(\mathbf{x}) = -\sigma_n^2 \lambda(\mathbf{x})$ a *cost metric*. Tree search is operated on a tree representing the search space of $\mathbf{x}$. Consider the N entries of $\mathbf{x}$, i.e., $x_1, \cdots, x_N$. Staring from the root node, we can expand a tree depending on the realizations of x_N to x_1 [2]. For each node, we extend 2^Q branches according to each symbol realization. These branch extensions repeat until all branches for $x_1, \cdots, x_N$ are generated. This yields a tree of the depth N, where each "complete" path traversing from the root to the bottom level is associated with a realization of $\mathbf{x}$. In the sequel, we will denote a path associated with a set of symbols $x_i, \cdots, x_j, (i < j)$ by a column vector $\mathbf{x}_i^j = [x_i, \cdots, x_j]^T$. For convenience, we call a level of tree associated with the symbol x_i "the ith level", e.g., the bottom level is called the "first" level.

For systematic search of the candidates minimizing the cost metric, a path metric is assigned to each path. To determine the path metric, we perform QR decomposition of $\mathbf{H} = \mathbf{Q}\mathbf{R}$ where $\mathbf{R}$ has an $N \times N$ upper-triangular matrix whose diagonals are non-negative and $\mathbf{Q}$ is an $L \times N$ unitary matrix[1]. Then, we can express $d_{\text{APP}}(\mathbf{x})$ as

$$\begin{aligned} d_{\text{APP}}(\mathbf{x}) &= \|\mathbf{y}' - \mathbf{R}\mathbf{x}\|^2 - \sigma_n^2 \sum_{k=1}^{N}\sum_{i=1}^{Q} \ln Pr(\bar{c}_{k,i}) \qquad (4) \\ &= \sum_{i=1}^{N} b(\mathbf{x}_i^N), \qquad (5) \end{aligned}$$

where $b(\mathbf{x}_i^N) = \left|y_i' - \sum_{j=i}^{N} r_{i,j}x_j\right|^2 - \sigma_n^2 \sum_{i=1}^{Q} \ln Pr(\bar{c}_{k,i})$ and $\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$. The path metric associated with the path $\mathbf{x}_k^N$ can be defined as a partial sum in the cost metric, i.e., $p^{(c)}(\mathbf{x}_k^N) = \sum_{i=k}^{N} b(\mathbf{x}_i^N)$. Since the path metric depends on the terms $b(\mathbf{x}_i^N)(k \leq i \leq N)$ associated with the symbols on the visited path, we call it *causal path metric*. According to predefined node extension order, the tree detection algorithm produces the most likely symbol candidates minimizing the cost metric. If we denote this candidate list as $\mathcal{L}$, the approximated APP is obtained as

$$L_{\text{post}}(\bar{c}_{k,i}) \approx \ln \frac{\sum_{\mathbf{x}\in\mathcal{L}\cap X_{k,i}^{+1}} \exp(\lambda(\mathbf{x}))}{\sum_{\mathbf{x}\in\mathcal{L}\cap X_{k,i}^{-1}} \exp(\lambda(\mathbf{x}))}. \tag{6}$$

III. Improved Path Metric

In this section, we derive a new path metric that improves the search efficiency of tree detection algorithms.

A. Motivation

We consider the following path metric;

Definition 3.1: A *genie-aided path metric* is defined as

$$p^{(G)}(\mathbf{x}_k^N) = p^{(T)}(\mathbf{x}_k^N) + \underbrace{\min_{\mathbf{x}_1^{k-1}}\left(\sum_{i=1}^{k-1} b(\mathbf{x}_i^N)\right)}_{\text{bias term}}. \tag{7}$$

The genie-aided path metric is obtained by minimizing the sum of $b(\mathbf{x}_i^N)(1 \leq i \leq k-1)$ over all combinations of undecided symbols $\mathbf{x}_1^{k-1}$. This minimal term is added to the causal path metric. This bias term accounts for the contributions of unvisited paths, which were dropped in the causal path metric.

Lemma 3.2: The modified M-algorithm, which uses the genie-aided path metric, can find a closest path with minimal number of node visitations, i.e., with $M = 1$.

This lemma can be easily proved using the fact that from any node, the genie-aided path metric provides the smallest cost metric among all tail paths.

[1]For the sake of simplicity, we consider a square matrix for $\mathbf{H}$ here, i.e., $L = N$. It is straightforward to extend our derivation to general rectangular matrices.

Theorem 3.3: Given the actual transmitted symbol vector $\widetilde{\mathbf{x}}_k^N$ (i.e., $\mathbf{x}_k^N = \widetilde{\mathbf{x}}_k^N$), the bias metric of the genie-aided path metric is given by

$$\min_{\mathbf{x}_1^{k-1}} \left(\sum_{i=1}^{k-1} b\left(\mathbf{x}_i^N\right) \right) = \sum_{i=1}^{k-1} b\left(\mathbf{x}_i^N\right) \Bigg|_{\mathbf{x}_1^{k-1} = \check{\mathbf{x}}_1^{k-1}}, \qquad (8)$$

where the minimizer $\check{\mathbf{x}}_1^{k-1}$ is the MAP estimate of $\mathbf{x}_1^{k-1}$, i.e.,

$$\check{\mathbf{x}}_1^{k-1} = \arg\max_{\mathbf{x}_1^{k-1}} \ln Pr\left(\mathbf{x}_1^{k-1} \middle| \mathbf{y}', \mathbf{x}_k^N = \widetilde{\mathbf{x}}_k^N\right). \qquad (9)$$

Theorem 3.3 implies that the bias term of the genie-aided path metric is obtained by evaluating $\sum_{i=1}^{k-1} b\left(\mathbf{x}_i^N\right)$ at the MAP estimate of $\mathbf{x}_1^{k-1}$. This MAP estimate is derived under the condition that a current visited path $\mathbf{x}_k^N$ is associated with the actual transmitted symbols. Though it sounds promising, computation of the bias term incurs high complexity for large systems. Hence, in what follows, we will present a practical way to improve the quality of path metric.

B. Derivation of New Path Metric

Consider a particular path $\mathbf{x}_k^N$. We define a new path metric as follows;

Definition 3.4: The *improved path metric*, denoted by $p^{(i)}\left(\mathbf{x}_k^N\right)$, is defined as

$$p^{(i)}\left(\mathbf{x}_k^N\right) \triangleq p^{(c)}\left(\mathbf{x}_k^N\right) + \underbrace{\sum_{i=1}^{k-1} b\left(\mathbf{x}_i^N\right) \Bigg|_{\mathbf{x}_1^{k-1} = \hat{\mathbf{x}}_1^{k-1}}}_{\text{bias term } p^{(b)}\left(\mathbf{x}_k^N\right)} \qquad (10)$$

where $\hat{\mathbf{x}}_1^{k-1}$ is the linear MMSE estimate of $\mathbf{x}_1^{k-1}$ derived assuming that $\mathbf{x}_k^N = \widetilde{\mathbf{x}}_k^N$.

In order to reduce the complexity for finding the MAP estimate in the genie-aided path metric, we relax finite alphabet constraint on the undecided symbols $\mathbf{x}_1^{k-1}$. Then, we use the linear MMSE estimate $\hat{\mathbf{x}}_1^{k-1}$ instead of the MAP estimate $\check{\mathbf{x}}_1^{k-1}$. Since the bias term depends on the visited path $\mathbf{x}_k^N$, we denote it as $p^{(b)}\left(\mathbf{x}_k^N\right)$.

In order to derive the linear MMSE estimate, we partition the vector $\mathbf{y}'$ and $\mathbf{n}'$ to $(k-1)\times 1$ and $(N-k+1)\times 1$ vectors, i.e.,

$$\mathbf{y}' = \begin{bmatrix} \mathbf{y}_1^{k-1} \\ \mathbf{y}_k^N \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11,k} & \mathbf{R}_{12,k} \\ \mathbf{0} & \mathbf{R}_{22,k} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1^{k-1} \\ \mathbf{x}_k^N \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1^{k-1} \\ \mathbf{n}_k^N \end{bmatrix}, \qquad (11)$$

where $\mathbf{R}_{11,k}$, $\mathbf{R}_{12,k}$, and $\mathbf{R}_{22,k}$ are the partitioned submatrices of $\mathbf{R}$. Using (11), we can express the improved path metric as $p^{(i)}\left(\mathbf{x}_k^N\right) = p^{(c)}\left(\mathbf{x}_k^N\right) + p^{(b)}\left(\mathbf{x}_k^N\right)$, where

$$p^{(c)}\left(\mathbf{x}_k^N\right) = \left\|\mathbf{y}_k^N - \mathbf{R}_{22,k}\mathbf{x}_k^N\right\|^2 + \xi\left(\mathbf{x}_k^N\right) \qquad (12)$$

$$p^{(b)}\left(\mathbf{x}_k^N\right) = \left\|\mathbf{y}_1^{k-1} - \mathbf{R}_{11,k}\hat{\mathbf{x}}_1^{k-1} - \mathbf{R}_{12,k}\mathbf{x}_k^N\right\|^2 \qquad (13)$$

and $\xi\left(\mathbf{x}_k^N\right) = -\sigma_n^2 \sum_{i=k}^N \sum_{j=1}^Q \ln Pr\left(\bar{c}_{i,j}\right)$. Given the condition $\mathbf{x}_k^N = \widetilde{\mathbf{x}}_k^N$, the linear MMSE estimate of $\mathbf{x}_1^{k-1}$ can be derived as

$$\hat{\mathbf{x}}_1^{k-1} = = \mathbf{F}_k\left(\mathbf{y}_1^{k-1} - \mathbf{R}_{11,k}\overline{\mathbf{x}}_1^{k-1} - \mathbf{R}_{12,k}\mathbf{x}_k^N\right) + \overline{\mathbf{x}}_1^{k-1} \qquad (14)$$

where $\overline{\mathbf{x}}_1^{k-1} = E[\mathbf{x}_1^{k-1}]$ and $\mathbf{F}_k = \mathrm{Cov}(\mathbf{x}_1^{k-1}, \mathbf{y}_1^{k-1} | \mathbf{x}_k^N = \widetilde{\mathbf{x}}_k^N)\ \mathrm{Cov}^{-1}\left(\mathbf{y}_1^{k-1} | \mathbf{x}_k^N = \widetilde{\mathbf{x}}_k^N\right)$. We can obtain $\overline{\mathbf{x}}_1^{k-1}$ from *a priori* LLRs [1] and

$$\mathbf{F}_k = \mathbf{\Lambda}_k(\mathbf{R}_{11,k})^H\left((\mathbf{R}_{11,k})\,\mathbf{\Lambda}_k(\mathbf{R}_{11,k})^H + \sigma_n^2\mathbf{I}\right)^{-1}, \qquad (15)$$

where $\mathbf{\Lambda}_k = \mathrm{diag}\left(\lambda_1, \cdots, \lambda_{k-1}\right)$ and $\lambda_i = \sum_{\theta\in\Theta} |\theta - \overline{x}_i|^2 \prod_{q=1}^Q \frac{1}{2}\left(1 + \bar{c}_{i,q}\tanh\left(\frac{L_{\mathrm{pri}}(\bar{c}_{i,q})}{2}\right)\right)$. The set Θ includes all constellation points that a symbol can take on. Notice that in the absence of *a priori* LLRs (e.g. in the first iteration), $\mathbf{\Lambda}_k = \mathbf{I}$ and $\overline{\mathbf{x}}_1^{k-1} = \mathbf{0}$. Denoting $\mathbf{q}_k = \mathbf{Z}_k\left(\mathbf{y}_1^{k-1} - \mathbf{R}_{11,k}\overline{\mathbf{x}}_1^{k-1}\right)$ and $\mathbf{P}_k = \mathbf{Z}_k\mathbf{R}_{12,k}$, we can express the improved path metric as

$$p^{(i)}\left(\mathbf{x}_k^N\right) = p^{(c)}\left(\mathbf{x}_k^N\right) + \left\|\mathbf{q}_k - \mathbf{P}_k\mathbf{x}_k^N\right\|^2, \qquad (16)$$

where $\mathbf{Z}_k = \sigma_n^2\left(\mathbf{R}_{11,k}\mathbf{\Lambda}_k(\mathbf{R}_{11,k})^H + \sigma_n^2\mathbf{I}\right)^{-1}$.

C. New SISO M-algorithm

We apply the improved path metric to the SISO M-algorithm. At every tree level, the best M paths are selected based on $p^{(i)}\left(\mathbf{x}_k^N\right)$ of $2^Q M$ survival paths. Beginning from the root node, this candidate selection proceeds for from x_N to x_1. At the bottom level, $2^Q M$ complete paths are selected as a candidate list $\mathcal{L}$. Over $\mathcal{L}$, the *a posteriori* LLR for each bit is calculated from (6).

IV. Performance Analysis

In the previous section, we showed that if the genie-aided path metric is used, the transmitted symbols are always found with $M = 1$. In this section, we investigate the performance of the improved path metric for the case of $M = 1$. We derive a probability of correct path loss (CPL), that is, the probability that the tree search rejects a path associated with transmitted symbols. Though the study is limited to the case of $M = 1$, better CPL probability for $M = 1$ implies better quality of path metric which would hold for other values of M.

Given the channel matrix $\mathbf{R}$ and the *a priori* LLRs, the probability of CPL can be expressed as

$$P_{\mathrm{CPL}} = 1 - \Pr\left(\widetilde{\mathbf{x}} \in \mathcal{L} | \widetilde{\mathbf{x}} \text{ is sent}\right) \qquad (17)$$

$$= 1 - \prod_{k=1}^N \left(1 - \underline{\Pr}\left(\widetilde{\mathbf{x}}_k^N \notin \mathcal{L}_k \middle| \widetilde{\mathbf{x}}_{k+1}^N \in \mathcal{L}_{k+1}\right)\right) \qquad (18)$$

where $\mathcal{L}_k$ denotes the set of paths selected at the kth level and $\underline{\Pr}(\cdot)$ is the probability given that $\widetilde{\mathbf{x}}$ is sent. Since we consider the case of $M = 1$, $\widetilde{\mathbf{x}}_{k+1}^N \in \mathcal{L}_{k+1}$ implies that a correct path have been selected up to the $k+1$th level. Given this condition and from (11), (12), and (13), we can express the path metric, $p^{(i)}\left(\mathbf{x}_k^N\right)$ as

$$p^{(i)}\left(\mathbf{x}_k^N\right) = \left\| \sqrt{\mathbf{r}_k^H\mathbf{Z}_k\mathbf{r}_k + |r_{k,k}|^2}\,(\widetilde{x}_k - x_k) + \frac{\begin{bmatrix}\mathbf{Z}_k\mathbf{r}_k & r_{k,k}\end{bmatrix}}{\sqrt{\mathbf{r}_k^H\mathbf{Z}_k\mathbf{r}_k + |r_{k,k}|^2}} \begin{bmatrix}\mathbf{Z}_k\mathbf{b}_k \\ n_k\end{bmatrix} \right\|^2 + \xi(x_k) + C \qquad (19)$$

where $\mathbf{b}_k = \mathbf{R}_{11,k}\left(\widetilde{\mathbf{x}}_1^{k-1} - \overline{\mathbf{x}}_1^{k-1}\right) + \mathbf{n}_1^{k-1}$, and $\mathbf{r}_k = \mathbf{R}_{12,k}\mathbf{e}_1 = [r_{1,k}, \cdots, r_{k-1,k}]^T$. Note that C is the term independent of the selection of x_k. The first term in (19) is seen as a distance metric between the output of scalar channel $\sqrt{\mathbf{r}_k^H\mathbf{Z}_k\mathbf{r}_k + |r_{k,k}|^2}\widetilde{x}_k + \frac{[\mathbf{Z}_k\mathbf{r}_k \;\; r_{k,k}]}{\sqrt{\mathbf{r}_k^H\mathbf{Z}_k\mathbf{r}_k+|r_{k,k}|^2}}\left[(\mathbf{Z}_k\mathbf{b}_k)^T \;\; n_k\right]^T$ and symbol candidate $\sqrt{\mathbf{r}_k^H\mathbf{Z}_k\mathbf{r}_k + |r_{k,k}|^2}x_k$. The proposed SISO M-algorithm chooses x_k that minimizes the value of $p^{(i)}\left(\mathbf{x}_k^N\right)$. Since the *a priori* term $\xi(x_k)$ in (19) would always lead to better detection, we neglect the impact of it for now. If we let $E\left[\mathbf{b}_k\mathbf{b}_k^H\right] = \mathbf{\Sigma}_k = \left(\mathbf{R}_{11,k}\mathbf{\Lambda}_k\mathbf{R}_{11,k}^H + \sigma_n^2\mathbf{I}\right)$ and $\mathbf{Z}_k = \sigma_n^2\mathbf{\Sigma}_k^{-1}$, the signal to interference plus noise ratio (SINR) of the scalar channel is given by

$$\text{SINR} = \frac{1}{\sigma_n^2}\frac{\left(\mathbf{r}_k^H\left(\sigma_n^4\mathbf{\Sigma}_k^{-2}\right)\mathbf{r}_k + |r_{k,k}|^2\right)^2}{\mathbf{r}_k^H\left(\sigma_n^6\mathbf{\Sigma}_k^{-3}\right)\mathbf{r}_k + |r_{k,k}|^2}. \tag{20}$$

Lemma 4.1: The SINR in (20) is bounded by

$$\sigma_n^2\mathbf{r}_k^H\mathbf{\Sigma}_k^{-2}\mathbf{r}_k + \frac{|r_{k,k}|^2}{\sigma_n^2} \leq \text{SINR} \leq \mathbf{r}_k^H\mathbf{\Sigma}_k^{-1}\mathbf{r}_k + \frac{|r_{k,k}|^2}{\sigma_n^2}. \tag{21}$$

In a similar manner, we can show that the SINR for the causal path metric is $\frac{|r_{k,k}|^2}{\sigma_n^2}$. Hence, the terms B_k^{upper} and B_k^{lower} represent the upper-bound and lower-bound of the SINR gain achieved by the improved path metric, respectively.

It would be interesting to see how the upper-bound and lower-bound of SINR gain converge as the size of $\mathbf{H}$ increases using *random matrix theory*. Suppose that $N, L \to \infty$ with a fixed aspect ratio $\beta = N/L$ $(0 < \beta \leq 1)$.

Theorem 4.2: Let $\lambda_{\min}$ and $\lambda_{\max}$ be smallest and largest diagonals of $\mathbf{\Lambda}_k$. For $L \times N$ matrix $\mathbf{H}$ whose elements are *i.i.d.* random variable with zero mean and variance of $\frac{1}{L}$, the upper-bound B_k^{upper} and lower-bound B_k^{lower} of the SINR gain for $k = \gamma N + 1$ $(0 < \gamma < 1)$ converge as $B_k^{\text{upper}} \longrightarrow B_k^{\text{upper},\infty}$ and $B_k^{\text{lower}} \longrightarrow B_k^{\text{lower},\infty}$, where

$$B_k^{\text{upper},\infty} = \frac{1}{2\lambda_{\min}}\left(-1 - (1-\gamma\beta)\frac{\lambda_{\min}}{\sigma_n^2} + G\left(\frac{\lambda_{\min}}{\sigma_n^2}, \gamma\beta\right)\right) \tag{22}$$

$$B_k^{\text{lower},\infty} = \frac{1}{2\sigma_n^2}\left(-(1-\gamma\beta) + \frac{1 + \gamma\beta + (1-\gamma\beta)^2\frac{\lambda_{\max}}{\sigma_n^2}}{G\left(\frac{\lambda_{\max}}{\sigma_n^2}, \gamma\beta\right)}\right) \tag{23}$$

as $N, L \to \infty$ with $\beta = N/L$, where $G(x, b) = \sqrt{1 + 2(1+b)x + (1-b)^2x^2}$.

Corollary 4.3: As $\sigma_n^2 \to 0$, we have

$$B_k^{\text{upper},\infty} \longrightarrow \lambda_{\min}\frac{\gamma\beta}{(1-\gamma\beta)} \tag{24}$$

$$B_k^{\text{lower},\infty} \longrightarrow 0. \tag{25}$$

In addition, $B_k^{\text{upper},\infty}$ is strictly bounded by $B_k^{\text{upper},\infty} < \lambda_{\min}\frac{\gamma\beta}{(1-\gamma\beta)}$.

According to the above results, the upper-bound and lower-bound of the SINR gain approach deterministic values which lie between $[0, \lambda_{\min}\frac{\gamma\beta}{(1-\gamma\beta)}]$. We can show that $B_k^{\text{upper},\infty}$ in (22) is an increasing function of $\gamma\beta \in (0, 1)$. Since γ indicates an index for tree depth, the upper-bound of the SINR gain is maximum at the top tree level.

Now, we derive the probability of CPL using the SINR we obtained. In order to make derivation tractable, we make Gaussian approximation for the MMSE estimation error $\mathbf{e}_1^{k-1} = \mathbf{x}_1^{k-1} - \hat{\mathbf{x}}_1^{k-1}$. We can show that this approximation leads to the assumption that the interference plus noise in the scaler channel is also Gaussian. The validity of the Gaussian approximation has been supported in [13]. Assume a random channel $\mathbf{H}$ whose elements are independent complex Gaussian with $\mathcal{CN}(0, 1)$. We can express the average probability of CPL, denoted as $\overline{P}_{\text{CPL}}$, as

$$\overline{P}_{\text{CPL}} = 1 - E_{\mathbf{H}}\left[\prod_{k=1}^{N}\left(1 - \underline{\Pr}\left(\widetilde{\mathbf{x}}_k^N \notin \mathcal{L}_k \middle| \widetilde{\mathbf{x}}_{k+1}^N \in \mathcal{L}_{k+1}\right)\right)\right] \tag{26}$$

$$= \sum_{k=1}^{N} E_{\mathbf{H}}\left[\underline{\Pr}\left(\widetilde{\mathbf{x}}_k^N \notin \mathcal{L}_k \middle| \widetilde{\mathbf{x}}_{k+1}^N \in \mathcal{L}_{k+1}\right)\right] + \text{h.o.t.}, \tag{27}$$

where $E_{\mathbf{H}}[\cdot]$ denotes the expectation operation in terms of $\mathbf{H}$. The higher order terms (h.o.t.) are neglected since they are smaller than the leading terms in high SNR regime. From (21), the upper-bound of the CPL probability can be given

$$\underline{\Pr}\left(\widetilde{\mathbf{x}}_k^N \notin \mathcal{L}_k \middle| \widetilde{\mathbf{x}}_{k+1}^N \in \mathcal{L}_{k+1}\right) \leq 4\left(1 - \frac{1}{\sqrt{2^Q}}\right)Q\left(\sqrt{K\left(\sigma_n^2\mathbf{r}_k^H\mathbf{\Sigma}_k^{-2}\mathbf{r}_k + \frac{|r_{k,k}|^2}{\sigma_n^2}\right)}\right). \tag{28}$$

Using the relationship $Q(\sqrt{x+y}) \leq Q(\sqrt{x})\exp\left(-\frac{y}{2}\right)$ for $x, y > 0$ and from (28), we have

$$E_{\mathbf{H}}\left[\underline{\Pr}\left(\widetilde{\mathbf{x}}_k^N \notin \mathcal{L}_k \middle| \widetilde{\mathbf{x}}_{k+1}^N \in \mathcal{L}_{k+1}\right)\right] \leq 4\left(1 - \frac{1}{\sqrt{2^Q}}\right) \cdot E_{\mathbf{H}}\left[Q\left(\sqrt{K\frac{|r_{k,k}|^2}{\sigma_n^2}}\right)\right]E_{\mathbf{H}}\left[\exp\left(-\frac{K\sigma_n^2\mathbf{r}_k^H\mathbf{\Sigma}_k^{-2}\mathbf{r}_k}{2}\right)\right], \tag{29}$$

where (29) follows from that $r_{k,k}$ and $\mathbf{r}_k$ are independent. We can show that with the causal path metric, the upper bound of the average CPL probability becomes $4\left(1 - \frac{1}{\sqrt{2^Q}}\right)E_{\mathbf{H}}\left[Q\left(\sqrt{K\frac{|r_{k,k}|^2}{\sigma_n^2}}\right)\right]$. Note that the original CPL probability is scaled by the term $\xi_k = E_{\mathbf{H}}\left[\exp\left(-\frac{K}{2}\sigma_n^2\mathbf{r}_k^H\mathbf{\Sigma}_k^{-2}\mathbf{r}_k\right)\right]$, which is strictly less than unity. Hence, we refer to the term as a *scaling gain*. Using the fact that $r_{k,k}$ has a Chi-square distribution with $2(L-k+1)$ degree of freedom, the closed-form expression of the first term in (29) can be obtained.

Lemma 4.4: The upper-bound of the scaling gain in (29) is given by

$$\xi_k \leq \int_0^\infty \cdots \int_0^\infty \left(\prod_{i=1}^{k-1} \frac{1}{1+\frac{K}{2}\frac{\sigma_n^2}{(\lambda_{\max} x_i + \sigma_n^2)^2}} \right) \times f_{\eta_1,\cdots,\eta_{k-1}}(x_1,\cdots,x_{k-1})\, dx_1 \cdots dx_{k-1}, \qquad (30)$$

where $f_{\eta_1,\cdots,\eta_{k-1}}(x_1,\cdots,x_{k-1})$ is the joint distribution of the unordered eigenvalues of $\mathbf{R}_{11,k}\mathbf{R}_{11,k}^H$.

While the term $\exp\left(-\frac{K\sigma_n^2 \mathbf{r}_k^H \mathbf{\Sigma}_k^{-2}\mathbf{r}_k}{2}\right)$ goes to one as $\sigma_n^2 \rightarrow 0$, the first term in (29) decreases to zero with a slope $\lim_{\sigma_n^2\rightarrow 0} \ln(P_e)/\ln(\sigma_n^2) = L-k+1$. Therefore, in high SNR regime, the average CPL probability would be dominated by that for $k=N$. Fig. 1 (a) provides the plot of the average CPL probability versus SNR for the 15×15 uncoded QPSK system. The average CPL rate are obtained from semi- Monte-Carlo simulation method, i.e., the exact CPL rate is averaged over random $\mathbf{H}$. The average CPL rate for the causal path metric is included for comparison. For all cases considered, this average CPL curve is quite close to that obtained from the simulations, which demonstrates the validity of Gaussian approximation. The upper bound is tight for high SNR but gets slightly loose as the SNR decreases. Fig. 1 (b) shows the plot of the scaling gain in (30) in terms of SNR for different system sizes. Larger performance gain is observed for larger system size and low to moderate SNR range. This behavior is beneficial for SISO detection, where the performance in low-to-mid SNR range is critical in triggering performance improvement through the IDD.

Due to the lack of space, we could not provide bit error rate (BER) performance of the proposed SISO detector here. We could observe better BER performance at the cost of slight complexity increase for large size of systems. We will include it in our subsequent works.

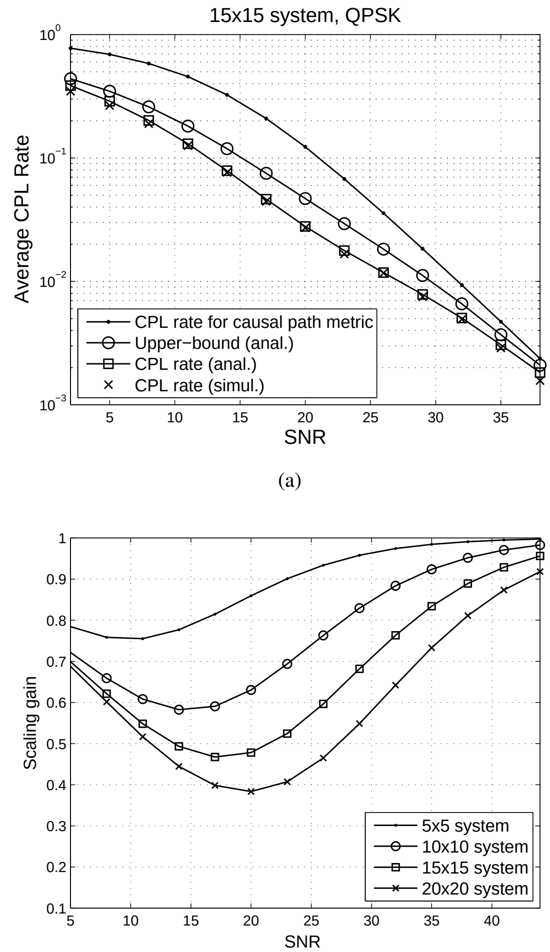

Fig. 1. (a) Average CPL rate versus SNR for 15×15 systems. (b) Scaling gain versus SNR for different system sizes $N=5,10,15,20$. QPSK uncoded transmission is considered.

V. Conclusions

In this paper, we proposed a new path metric, which improves the search efficiency of SISO tree detection algorithm. By studying the probability of correct path loss, we proved the superiority of the improved path metric over the existing causal path metric. Though we have focused on the application to the breadth-first search algorithm, the improved path metric can be applied to other tree search algorithms based on a depth-first or best-first search as well.

References

[1] B. Hochwald and S. T. Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389-399, Mar. 2003.

[2] A. D. Murugan, H. E. Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: rediscovering the sequential decoder," *IEEE Trans. Information Theory*, vol. 52, pp. 933-953, March 2006.

[3] R. Wang and G. B. Giannakis, "Approaching MIMO channel capacity with soft detection based on hard sphere decoding," *IEEE Trans. Commun.*, vol. 54, pp. 587-590, April 2006.

[4] J. Jalden and B. Ottersten, "Parallel implementation of a soft output sphere decoder," *Proc. IEEE Ashilomar Conference on Signals, Systems, and Computers*, Nov. 2005, pp. 581-585.

[5] C. Studer, A. Burg, and H. Bolcskei, "Soft-output sphere decoding: algorithms and VLSI implementation," *IEEE Journal on Selected Areas in Commun.*, vol. 26, pp. 290-300, Feb. 2008.

[6] Y. L. C. de Jong and T. J. Wilink, "Iterative tree search detection for MIMO wireless systems," *IEEE Trans. Commun.*, vol. 53, pp. 930-935, June 2005.

[7] J. Hagenauer and C. Kuhn, "The list-sequential (LISS) algorithm and its application," *IEEE Trans. Commun.*, vol. 55, pp. 918-928, May 2007.

[8] L. G. Overlinebero, T. Ratnarajah, and C. Cowan, "A low-complexity soft-MIMO detector based on the fixed-complexity sphere decoder," *IEEE Acoustic, Speech and Signal Processing (ICASSP)*, April 2008, pp. 2669-2672.

[9] E. G. Larsson and J. Jalden, "Fixed-complexity soft MIMO detection via partial marginalization," *IEEE Trans. Signal Processing*, vol. 56, pp. 3397-3407, Aug. 2008.

[10] R. Fano, "A heuristic discussion of probabilistic decoding," *IEEE Trans. Information Theory,* vol. 55, pp. 918-928, May 2007.

[11] F. Xiong, A. Zerik, and E. Shwedyk, "Sequential sequence estimation for channels with intersymbol interference of finite or infinite length," *IEEE Trans. Commun.*, vol. 38, pp. 795-804, June 1990.

[12] R. Gowaikar and B. Hassibi, "Statistical Pruning for near-Maximum Likelihood Decoding," *IEEE Trans. Signal Processing*, vol. 55, pp. 2661-2675, June 2007.

[13] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection," *IEEE Trans. Information Theory*, vol. 43, pp. 858-871, May 1997.