Hard/Soft Detection with Limited CSI for Multi-Hop Systems

Mamoun Malkawi and Il-Min Kim, Senior Member, IEEE

Abstract—We consider a single-branch, Mobile Multihop Relaying (MMR) network with limited Channel State Information (CSI) at the destination. The considered CSI signalling strategy reduces the signalling overhead at each relay by at least 50%. We show that this significant overhead reduction comes at the expense of no performance loss at all when hard Maximum Likelihood detection is carried out at the destination. Furthermore, we consider the use of the MMR system with concatenated channel codes to carry out soft Maximum a Posteriori (MAP) detection, and demonstrate that with channel codes employed the optimum detection rule becomes prohibitively complex to implement. We propose two approximate soft MAP detection schemes to make the detection feasible for the system, and demonstrate that the performance is either almost identical to or slightly degraded from the ideal case with full CSI at the destination.

Index Terms—Amplify-and-forward protocol, MMR networks, limited CSI.

I. INTRODUCTION

WITH the recent upsurge in demand for wireless communication services, next-generation wireless systems are expected to provide very high data rates and improved quality of service while staying within commercial feasibility by being affordable to deploy. One very promising tool to aid in meeting market expectations is the concept of Mobile Multihop Relaying (MMR) [1], which describes systems having an intermediary station(s) between a source and destination to act as a relay point by forwarding data symbols along a specified source-destination path. MMR has proven to be very effective with various wireless systems such as cellular networks and ad hoc sensor networks among others [2], [3]. Some advantages of deploying relay terminals include extending the coverage of existing base station terminals, eliminating coverage problems caused by dead spots and deep-shadowed regions, and reducing power consumption at base stations. In fact, realizing the gains achieved by utilizing MMR in wireless systems, an official IEEE task group was established to introduce IEEE 802.16j to standardize the use of MMR in the existing IEEE 802.16e standard [4].

Considering signal detection, the destination in almost all works has been assumed to have perfect instantaneous Channel State Information (CSI). Although this is a widely accepted assumption, its implications can be very costly for relay terminals because it means an increased signalling overhead burden. Few works studied non-coherent hard detection schemes in the context of cooperative diversity networks [5], [6], but they do suffer some drawbacks. The optimum Maximum Likelihood (ML) detection schemes proposed in those works involve numerical integral evaluations that are very costly to implement at the destination, while the other low-complexity hard detection schemes proposed are sub-optimum. Furthermore, all the schemes proposed in [5] and [6] are valid only for Binary Frequency Shift Keying and On-Off Keying schemes. Other works like [7]–[11] studied blind hard detection schemes through the use of differential modulation. Other than these works, almost all works assume perfect instantaneous CSI at the destination [12]–[14].

In order to address the problems described above, we consider an MMR system with limited CSI and a reduced relay signalling overhead. The system is detected coherently, but does not require the knowledge of instantaneous CSI of each individual hop. For this system, we study the ML detector and show that it can still be implemented as a Minimum Euclidian Distance detector. We also propose two novel low complexity approximate Maximum a posteriori Probability (MAP) detection schemes to be used with channel codes and show that they achieve comparable performance to the case where the detector obtains full CSI. Furthermore, our proposed system is applicable to any type of channel fading environment.

Notation: Bold lower letters denote vectors. \( x \sim \mathcal{CN}(\mu, \Omega) \) means that \( x \) is a circularly symmetric complex Gaussian random variable with mean \( \mu \) and variance \( \Omega \). \( E[\cdot] \) and \( \text{Var}[\cdot] \) are the expectation and variance operators, and \( x|y \) means that \( x \) is conditioned on the knowledge of \( y \).

II. SYSTEM MODEL

We consider a single-branch MMR system consisting of a source, a destination, and \((K-1)\) intermediary relays. Consequently, this systems is called a \(K\)-hop system. Each node possesses a single receive and transmit antenna working in a half-duplex mode. At the first time slot, the source transmits a symbol \( x \in \mathcal{S} \), where \( \mathcal{S} \) is any arbitrary two-dimensional \( \mathbb{M} \)-Quadrature Amplitude Modulation (QAM) or \( \mathbb{M} \)-Phase Shift Keying constellation, and \( E[|x|^2] = E_s \). We consider the amplify-and-forward protocol at each relay and assume that each relay has no instantaneous CSI, but possesses knowledge of channel statistics. In this sense the relays can be considered as blind relays since the \( i \)-th relay amplifies the received symbol using a fixed amplification coefficient \( \alpha_i \). The received signal \( y_K \) at the destination can therefore be written...
as

\[ y_K = \left( \prod_{i=1}^{K} h_i \right) \left( \prod_{j=1}^{K-1} \alpha_j \right) x + \tilde{n} \quad (1) \]

where the noise component \( \tilde{n} \) is given by

\[ \tilde{n} = \sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} h_j \right) \left( \prod_{q=i}^{K-1} \alpha_q \right) n_i + n_K. \quad (2) \]

Each \( h_i, i \in \{1, \cdots, K\} \) represents the channel gain for the \( i \)-th hop, and is randomly distributed according to any arbitrary channel model with \( E[|h_i|^2] = \Omega_i \). Each value \( n_i \sim \mathcal{CN}(0, \sigma_i^2), i \in \{1, \cdots, K\} \) represents the additive white Gaussian noise (AWGN) at the \( i \)-th hop. Each \( \alpha_i \) is set just as in [14]–[18] and is given by

\[ \alpha_i = \sqrt{\frac{E_{r,i}}{E_s \Omega_i + \sigma_i^2}}. \quad (3) \]

This setting for \( \alpha_i \) ensures that each relay satisfies a long term power constraint given by \( E_{r,i} \), where \( E_{r,i} \) is the average transmission power at the \( i \)-th relay. The average Signal to Noise Ratio at the \( i \)-th hop is therefore given by \( SNR_i = \Omega_i E_s / \sigma_i^2 \) for \( i = 1 \), and \( SNR_i = \alpha_{i-1} \Omega_i E_{r,i-1} / \sigma_i^2 \) for \( i \in \{2, \cdots, K\} \).

A. CSI Signalling

1) Full CSI Signalling Strategy: In full CSI signalling strategies, the destination is assumed to have full instantaneous CSI for every hop. This may be achieved by using either of two different pilot signalling strategies: the first being a strategy where the source and each relay transmit a pilot signal to the next terminal. Using the pilot signal, the next terminal (the next relay or the destination) estimates the channel of its own individual hop. The estimated channel coefficient at each relay is then quantized and transmitted down the path all the way to the destination. In this scenario, each relay must have the ability to estimate its own channel coefficient. This signalling strategy suffers from two major drawbacks. The first drawback is the need of intelligent relays with the ability of estimating the channel. In systems with blind relays, such as the one adopted in this paper, estimation and quantization of the channel is not possible. Even if this drawback is avoided by using more intelligent relays, this pilot signalling strategy still suffers from another disadvantage that the channel coefficients at each relay will have to be quantized and retransmitted. This digital retransmission of the channel coefficients will incur an increased overhead since the channel coefficient, being a complex number, will need a high bit resolution to quantize accurately. Even if the system uses a quantized codebook for the channel distribution where the codebook index of the estimated channel is transmitted instead of the actual channel coefficient, this codebook index will still need several bits of transmission.

In the other pilot signalling strategy, each terminal, including the source and each relay, generates its own pilot signal, which in turn is relayed all the way to the destination. That is, each relay is responsible for forwarding the pilot signals for all previous terminals, as well as starting the transmission of

- a pilot signal. The destination eventually will have received \( K \) different pilot signals, and will be able to estimate the CSI for each individual hop. Fig. 1 provides an illustration of this CSI signalling strategy for a three-hop network. This strategy, unlike the previous one, does not suffer from the need of intelligent relays since no channel estimation is carried out at the relays. It also avoids the need for relays to quantize and transmit the channel coefficients of the individual hops. However, the signalling overhead can still be an issue since each relay is responsible for forwarding all the pilot signals generated by the previous terminals. That means the \( i \)-th relay needs to transmit \( (i+1) \) pilot signals (one that it generates, and \( i \) that it relays from previous terminals). That can be a significant burden especially as \( K \) increases.

Since we consider blind relays in the MMR system, the second CSI signalling strategy would be the most suitable one out of the two to obtain full CSI at the destination. For the rest of this paper, we will refer to the second CSI signalling strategy as the full CSI signalling strategy.

2) Proposed CSI Signalling Strategy: We consider a new CSI signalling strategy that eliminates much of the overhead of the full CSI signalling strategy. We will assume that the relays avoid generating their own pilot signals, and that only the source generates one pilot signal that is amplified and forwarded by every relay. Fig. 2 illustrates this CSI signalling strategy for a three-hop network. The implications of this assumption are as follows: 1) The signalling overhead at the \( i \)-th hop is reduced to \((i+1)\) of the overhead compared to the full signalling strategy; 2) The destination has no knowledge of the instantaneous CSI of each hop. Instead, the destination acquires the overall product of channel gain coefficients only, which is given by \( \left( \prod_{i=1}^{K} h_i \right) \).

This reduction in pilot signalling is very significant in terms of transmission power. The ratio between the total number of pilot signals required by the proposed strategy over that required by the full strategy is \( 2/(K+1) \). An attractive feature of the proposed strategy is that its savings improve as \( K \) increases. In fact, as \( K \) tends to infinity, the ratio tends to zero. In what follows, we will see that this overhead reduction comes at absolutely no loss in performance in terms of hard ML detection, and at a small loss in terms of soft MAP detection.
B. Likelihood Function and Extrinsic Information

To derive the likelihood function of \(y_K\), we first look at the distribution of \(\tilde{n}\). Note that in the full CSI signalling strategy, \(\tilde{n}\) has a Gaussian Probability Density Function (PDF) with zero mean and \(\text{Var}[\tilde{n}|\{h_i: i=1, \ldots, K\}] = \sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |h_j|^2 \right) \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2\), and therefore the likelihood function \(p(y_K|x, \{h_i : i = 1, \ldots, K\})\) is given in closed form as a Gaussian function. With that property, ML detection at the receiver is simply carried out using the classical minimum Euclidian distance rule. Soft MAP detection can also be carried out using any of the existing schemes for the AWGN channel with known CSI.

However, the distribution of the noise component \(\tilde{n}\) in the proposed CSI signalling strategy is no longer Gaussian because the individual \(h_i, i \in \{1, \ldots, K\}\) are not known and are still random entities. Expanding (2) will give a summation of \(K\) terms, each of which is a multiplication of multiple random variables. This makes deriving the PDF of \(\tilde{n}\) quite involved. Even if a closed form PDF is derived for \(\tilde{n}\), then it would only be valid for a specific PDF model for each \(h_i, i \in \{1, \ldots, K\}\).

To approach the problem of finding a general expression for the PDF of \(\tilde{n}\), we condition the distribution of \(\tilde{n}\) on the knowledge of the channels at each hop, then we take the expectation over all the channel PDFs. That would give us an integral form for the likelihood function as follows

\[
p(y_K|x, \prod_{i=1}^{K} h_i) = \frac{1}{\pi} \int_{\Psi_K} \cdots \int_{\Psi_2} \exp\left( -\beta_K(x|\prod_{i=1}^{K} h_i) \sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |z_j|^2 \right) \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2 \right) \times f_{|h_2|^2, \ldots, |h_K|^2}|\prod_{i=1}^{K} h_i|^{2z_2, \ldots, z_K} \cdot d_{z_2} \cdots d_{z_K} \tag{4}
\]

where \(f_{|h_2|^2, \ldots, |h_K|^2}|\prod_{i=1}^{K} h_i|^{2z_2, \ldots, z_K}\) represents the joint PDF of \(|h_i|^2, i \in \{2, \ldots, K\}\) conditioned on the knowledge of \(\prod_{i=1}^{K} h_i\), and

\[
\beta_K(x|\prod_{i=1}^{K} h_i) = \left( \prod_{i=1}^{K} h_i \right) \left( \prod_{j=1}^{K-1} \alpha_j \right) x^2. \tag{5}
\]

Each range \(\Psi_i, i \in \{2, \ldots, K\}\) represents the positive range of each \(|h_i|^2\) given that \(\prod_{i=1}^{K} h_i\) is known. We have found that it was extremely difficult to solve the integration in (4) in closed form, even for the simplest case for \(K = 2\) and a Rayleigh fading channel. Interestingly, however, we will see that achieving ML detection is still possible without the need for the integration in (4).

Note that with our proposed CSI signalling strategy, our system is still a coherent system. Since the effective channel coefficient, \(\prod_{i=1}^{K} h_i \left( \prod_{j=1}^{K-1} \alpha_j \right)\), is known at the receiver, the detector is coherently detected. The only difference between our proposed strategy and the full CSI strategy is that the noise component \(\tilde{n}\) in the full CSI strategy is Gaussian, while in our proposed strategy the noise is non-Gaussian.

When the use of capacity approaching channel codes, e.g. Turbo codes or Low Density Parity Check (LDPC) codes, is desired, the detector will have to produce soft extrinsic likelihood information which is used by the soft-in soft-out channel decoder to decode the received frame. In that case, the extrinsic information is passed on in the form of a log-likelihood ratio for each modulated bit that is given by

\[
L_E(y_K|h_m \prod_{i=1}^{K} h_i) = \ln \sum_{x \in S_{m,+1}} p(y_K|x, \prod_{i=1}^{K} h_i) \exp(\lambda)
- \ln \sum_{x \in S_{m,-1}} p(y_K|x, \prod_{i=1}^{K} h_i) \exp(\lambda) \tag{6}
\]

where \(\lambda\) represents the \(a\ priori\) information and \(S_{m,+1}\) and \(S_{m,-1}\) are a partition of \(S\). For more details on soft decoding of concatenated systems, the reader is referred to [21] where the issue is addressed in the context of multi-antenna systems. From (6), we see that evaluating the extrinsic information requires evaluating \((K - 1) \times M\) numerical integrations per modulated bit, which is intolerable even for the simple dual-hop scenario.

III. HARD ML DETECTION

Based on the Likelihood function in (4), one may obtain the ML solution by numerically integrating (4) for all the symbols in the constellation, and then decide in favor of the symbol giving the integration the maximum value. Obviously, the complexity of this detection is prohibitively high. It is easily shown however, that computing the integral of (4) is actually unnecessary and that the ML symbol estimate can be simplified to the classical minimum Euclidian distance rule as follows

\[
\hat{x}_{ML} = \arg \max_{\hat{x} \in \mathcal{S}} p(y_K|x, \prod_{i=1}^{K} h_i)
= \arg \max_{\hat{x} \in \mathcal{S}} \frac{1}{\pi} \exp\left( -\beta_K(x|\prod_{i=1}^{K} h_i) \sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |z_j|^2 \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2 \right) \right)
\times \int_{\Psi_K} \cdots \int_{\Psi_2} \sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |z_j|^2 \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2 \right) \times f_{|h_2|^2, \ldots, |h_K|^2}|\prod_{i=1}^{K} h_i|^{2z_2, \ldots, z_K} \cdot d_{z_2} \cdots d_{z_K} \right) \tag{7}
\]

This in turn becomes

\[
\hat{x}_{ML} = \arg \max_{\hat{x} \in \mathcal{S}} \exp \left( -\beta_K(x|\prod_{i=1}^{K} h_i) \sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |z_j|^2 \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2 \right) \right) \tag{8}
\]

\[
\sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |z_j|^2 \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2 \right)
\sum_{i=1}^{K-1} \left( \prod_{j=i+1}^{K} |z_j|^2 \left( \prod_{q=i+1}^{K-1} \alpha_q^2 \right) \sigma_i^2 + \sigma_K^2 \right)
\]
where each $\tilde{z}_i, i \in \{2, \cdots, K\}$ is any arbitrary positive constant in $\Psi_i$, (8) is true since the likelihood function is positive and strictly decreasing with respect to $\beta_K(x) \prod_{i=1}^{K} h_i$, and therefore maximizing the integral will be equivalent to maximizing the integrand. We are therefore able to reduce the $i$-th relay's signalling overhead by a factor of $i/(i+1)$ compared to the full CSI signalling strategy at no performance loss at all when hard ML detection is carried out.

IV. SOFT MAP DETECTION

In order to make our analysis more tractable, throughout Section IV we introduce an approximation to $p(y_k|x, \prod_{i=1}^{K} h_i)$ that ignores the knowledge of $\prod_{i=1}^{K} h_i$ in the joint PDF, $f_{[h_2,\cdots,|h_K|]}(\prod_{i=1}^{K} h_i, z_2, \cdots, z_K)$. Then $p(y_k|x, \prod_{i=1}^{K} h_i)$ is approximated by $\tilde{p}(y_k|x, \prod_{i=1}^{K} h_i)$ where

$$\tilde{p}(y_k|x, \prod_{i=1}^{K} h_i) = \frac{1}{\pi} \int_0^\infty \cdots \int_0^\infty \frac{\exp\left(\frac{-\beta_K(x) \prod_{i=1}^{K} h_i}{\sum_{i=1}^{K-1} \left(\prod_{j=i+1}^{K} z_j \right) \left(\prod_{q=i+1}^{K-1} \alpha_2 q^2 + \alpha_2^2 \right) \sigma_1^2 + \sigma_2^2 \right)} \times \prod_{i=1}^{K} f_{[h_i]}(z_i) \, dz_2 \cdots dz_K. \tag{9}$$

Numerical results in Section V will demonstrate that this approximation is very accurate.

A. Gaussian Approximation

Our first sub-optimum MAP detection scheme makes the assumption that the noise component $\tilde{n}$ of (2) is of a Gaussian distribution. Specifically, we define a Gaussian random variable $\tilde{n}_{\mathcal{G}} \sim \mathcal{CN}(0, \sum_{i=1}^{K-1} \left(\prod_{j=i+1}^{K} \alpha_2 q^2 \right) \sigma_1^2 + \sigma_2^2)$ such that $E[\tilde{n}_{\mathcal{G}}] = E[\tilde{n}]$ and $\text{Var}[\tilde{n}_{\mathcal{G}}] = \text{Var}[\tilde{n}]$. Then we can rewrite (6) as (12) on top of the next page.

We can justify this Gaussian assumption particularly for the cases where the variance of the noise component at the last hop dominates the variance at the first hop (and earlier hops), or equivalently when $SNR_R, i \in \{1, \cdots, K-1\}$ are much higher than $SNR_K$. This is a typical scenario in the downlink of a relay network where the channel between the base-station and the relays (which can be analog repeaters) is a Ricean channel with a strong Line of Sight component. However, when $SNR_K \gg SNR_1$ the Gaussian approximation falls short. In order to address this case, we introduce the second sub-optimum MAP detection scheme in what follows.

B. Mean Value Theorem Based Approximation for Dual-hop

In this subsection, we consider the special class of dual-hop networks, which have been shown to be very practical [19]. We propose an approximate MAP detection scheme suitable for dual-hop systems to avoid the need of evaluating the integrals given in (6), and that performs better than the Gaussian approximation when $SNR_R > SNR_1$. Based on the First Mean Value Theorem for Integration (MVT) [20, §3.01-Theorem 3.012], we can write $\tilde{p}(y_{2k}|x, \prod_{i=1}^{2} h_i)$ as follows:

$$\int_0^\infty \exp\left(\frac{-\beta_2(x) \prod_{i=1}^{2} h_i}{\alpha_2^2 \gamma \sigma_1^2 + \sigma_2^2}\right) f_{[h_2]}(z) \, dz = \frac{1}{\alpha_2^2 \gamma \sigma_1^2 + \sigma_2^2} \exp\left(\frac{-\beta_2(x) \prod_{i=1}^{2} h_i}{\alpha_2^2 \gamma \sigma_1^2 + \sigma_2^2}\right) \tag{10}$$

where $\gamma \in (0, \infty)$. We dropped the constant $1/\pi$ since it cancels out in the calculation of (6). Worthy of note here is our justification for choosing the appropriate functions to separate the integrand in (10) since the MVT allows some flexibility in choosing the functions that form the integrand. Our solution is to always separate the integrand as in (10) into a Gaussian function and the PDF of $[h_2]^2$. First of all, this solution is valid for all channels regardless of their distributions. Our integrand will always be equivalent to the Gaussian function on the right hand side of (10) multiplied by the integration of the distribution of $[h_2]^2$, which is equal to 1. The second reason for using the Gaussian function is that it provides us with the convenience of using the very standard Maxlog approximation [21], and that can provide a significant complexity reduction in the soft detection process.

Therefore, our problem is now directed towards finding the right $\gamma$ value to evaluate $\tilde{p}(y_{2k}|x, \prod_{i=1}^{2} h_i)$ in (10). Note that in (10), $\gamma$ is a function of $\{\beta_2(x) \prod_{i=1}^{2} h_i, \mu, \sigma_2^2\}$. If we were able to know $\gamma$ for a fixed $\{\beta_2(x) \prod_{i=1}^{2} h_i, \mu, \sigma_2^2\}$ then we could exactly evaluate $\tilde{p}(y_{2k}|x, \prod_{i=1}^{2} h_i)$ by the right hand side of (10). One may solve for $\gamma$ in (10) numerically for different possible values of $\{\beta_2(x) \prod_{i=1}^{2} h_i, \mu, \sigma_2^2\}$ and use it to exactly evaluate $\tilde{p}(y_{2k}|x, \prod_{i=1}^{2} h_i)$. That solution however may not be feasible because there are three variables and the range of each variable is also too large to handle. To solve this dependency on many variables, we make the following manipulations to (10)

$$\int_0^\infty \frac{1}{\epsilon + 1} \exp\left(\frac{-\beta_2(x) \prod_{i=1}^{2} h_i}{\epsilon + 1}\right) f_{[h_2]}(z) \, dz = \frac{1}{\epsilon + 1} \exp\left(\frac{-\beta_2(x) \prod_{i=1}^{2} h_i}{\epsilon + 1}\right) \tag{11}$$

where $\beta_2(x) \prod_{i=1}^{2} h_i = \beta_2(x) \prod_{i=1}^{2} h_i / \sigma_2^2$ and $\epsilon = \mu / \sigma_2^2$. Dividing by $\sigma_2^2$ will allow us to model $\gamma$ as a function of $\{\beta_2(x) \prod_{i=1}^{2} h_i, \epsilon\}$, reducing the dependency of $\gamma$ to two variables, $\beta_2(x) \prod_{i=1}^{2} h_i$ and $\epsilon$. When studying the behavior of $\gamma$, we noticed that it could be modeled, with a great deal of accuracy, by a simple function with respect to $\beta_2(x) \prod_{i=1}^{2} h_i$ given that $\epsilon$ is fixed. We model $\gamma|\epsilon$ as

$$\gamma|\epsilon \approx b \left(\beta_2(x) \prod_{i=1}^{2} h_i\right)^a \tag{13}$$

where $b > 0$ and $a \in [0, 1]$ are found numerically by minimizing the expected Mean Squared Error (MSE), $E[|\gamma| - \gamma|\epsilon|^2$.
the tabulated function to obtain the range to be categorized for the dual-hop network can be theoretically extended to the multi-hop case. Unlike the dual-hop case, modeling the relationship between \( \gamma_{1} \) and the variables that influence them is much more difficult because they are interrelated. This makes modeling each \( \gamma_{j} \) value a multi-variable problem. Overall, when more than two hops exist, the Gaussian approximation seems to be more suitable than the MVT-based approximation due to its simplicity.

\[
L_{E}(y_{K}|b_{m}, \prod_{i=1}^{K} h_{i}) = \\
\ln \sum_{x \in S_{m+1}} \exp \left( \frac{-\beta_{K}(x) \prod_{i=1}^{K} h_{i}}{\sum_{i=1}^{K-1} \left( \prod_{q=i}^{K-2} \alpha_{q} \right) \left( \prod_{j=i+1}^{K} \Omega_{j} \right) \sigma_{i}^{2} + \sigma_{K}^{2}} + \lambda \right) \\
- \ln \sum_{x \in S_{m-1}} \exp \left( \frac{-\beta_{K}(x) \prod_{i=1}^{K} h_{i}}{\sum_{i=1}^{K-1} \left( \prod_{q=i}^{K-2} \alpha_{q} \right) \left( \prod_{j=i+1}^{K} \Omega_{j} \right) \sigma_{i}^{2} + \sigma_{K}^{2}} + \lambda \right)
\]

(12)

The function model is calculated once for each \( \epsilon \) value, and this one-dimensional tabulation can be stored in the function model in Fig. 3 loses some accuracy when \( \epsilon \) is large. However, that is still acceptable since the PDF of \( \beta_{2}(x|\prod_{i=1}^{2} h_{i}) \) is obtained, it is divided by the current \( \sigma_{2}^{2} \) to obtain \( \beta_{2}(x|\prod_{i=1}^{2} h_{i}) \). The resulting value is then used in the tabulated function to obtain \( \gamma_{1} \). Note that the function model in Fig. 3 loses some accuracy when \( \beta_{2}(x|\prod_{i=1}^{2} h_{i}) \) is large. However, that is still acceptable since the PDF of \( \beta_{2}(x|\prod_{i=1}^{2} h_{i}) \) closely resembles the exponential PDF, and the large \( \beta_{2}(x|\prod_{i=1}^{2} h_{i}) \) region happens very rarely. Table I gives an example of the tabulated \( \gamma_{1} \) functions represented by the coefficients \( a \) and \( b \) for different values of \( \epsilon \in \{0, \cdots , 30\} \) dB. Again this tabulation is done assuming a Rayleigh fading channel. The MSE value, \( d_{\text{min}} \), between the function model of (13) and the actual \( \gamma_{1} \) is also given as a normalized percentage with respect to \( E[|\gamma_{1}|^{2}] \). It was observed numerically that the error, \( d_{\text{min}} \), values are extremely small, with a value always less than 0.2%. Note that while this table is generated assuming a Rayleigh fading channel, the methodology of our MVT-based approximation is valid for any channel distribution.

C. On the Mean Value Theorem Based Approximation for Multi-hop

By [22, §12.11-Theorem 3] the MVT is applicable to (4), which means that the MVT-based MAP approximating scheme for the dual-hop network can be theoretically extended to the multi-hop case. Unlike the dual-hop case, modeling the relationship between \( \gamma_{2}, \cdots , \gamma_{K} \) and the variables that influence them is much more difficult because they are interrelated. This makes modeling each \( \gamma_{j} \) value a multi-variable problem. Overall, when more than two hops exist, the Gaussian approximation seems to be more suitable than the MVT-based approximation due to its simplicity.

V. SIMULATION RESULTS

For the performance of hard ML detection, It is obvious from Section III that the performance of the MMR system with our signalling strategy and the full CSI strategy is identical. Therefore, we do not provide any simulation results for hard ML detection and only focus on the soft MAP detection case. We have conducted numerical simulations for both the dual-hop and three-hop relay systems \((K = 2 \text{ and } K = 3)\) employing soft MAP detection at the destination. For all our simulations we have used a rate 1/2 LDPC code of length 204. We have assumed a block Rayleigh fading channel with each \( h_{i} \sim CN(0, 1), \) \( i \in \{1, \cdots , K\} \), and changes every transmitted frame of 204 bits. We have used a 4-QAM constellation in all of our simulations. We have also used the standard Max-log approximation to reduce the complexity of decoding.

Fig. 4 shows the Frame Error Rate (FER) performance as a function of \( SNR_{1} \) for a dual-hop network. We have assumed that the transmitted energy at the relay is normalized such that \( E_{r_{1}} = E_{s} \). To compare the performance of our proposed schemes as opposed to the case with full CSI signalling, we varied the ratio \( \sigma_{1}^{2}/\sigma_{2}^{2} \) for three different cases such that we have \( SNR_{1} \gg SNR_{2}, SNR_{1} \approx SNR_{2}, \) and \( SNR_{1} \ll SNR_{2} \). For the case of \( SNR_{1} \gg SNR_{2} \) \((\sigma_{1}^{2}/\sigma_{2}^{2} = 30 \text{ dB})\), as expected in Section IV-A, we see that the Gaussian approximation performs almost identically to the full CSI signalling strategy, and therefore there’s no real incentive to use the MVT-based approximation. For the case of \( SNR_{1} \approx SNR_{2} \) \((\sigma_{1}^{2}/\sigma_{2}^{2} = 0 \text{ dB})\), we see that
TABLE I

<table>
<thead>
<tr>
<th>SNR_1 (dB)</th>
<th>a</th>
<th>b</th>
<th>d_{min}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.7</td>
<td>0.5</td>
<td>0.1960 %</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>0.6</td>
<td>0.0486 %</td>
</tr>
<tr>
<td>2</td>
<td>0.6</td>
<td>4.7</td>
<td>0.1406 %</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
<td>0.4</td>
<td>0.1754 %</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
<td>0.6</td>
<td>0.1187 %</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>0.9</td>
<td>0.0547 %</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
<td>5.1</td>
<td>0.1334 %</td>
</tr>
<tr>
<td>7</td>
<td>0.8</td>
<td>0.5</td>
<td>0.1494 %</td>
</tr>
<tr>
<td>8</td>
<td>0.7</td>
<td>0.7</td>
<td>0.0628 %</td>
</tr>
<tr>
<td>9</td>
<td>0.7</td>
<td>1.1</td>
<td>0.0341 %</td>
</tr>
<tr>
<td>10</td>
<td>0.5</td>
<td>4.1</td>
<td>0.0973 %</td>
</tr>
<tr>
<td>11</td>
<td>0.7</td>
<td>0.6</td>
<td>0.0587 %</td>
</tr>
<tr>
<td>12</td>
<td>0.7</td>
<td>0.7</td>
<td>0.0725 %</td>
</tr>
<tr>
<td>13</td>
<td>0.7</td>
<td>1.1</td>
<td>0.0289 %</td>
</tr>
<tr>
<td>14</td>
<td>0.5</td>
<td>4.6</td>
<td>0.1168 %</td>
</tr>
<tr>
<td>15</td>
<td>0.7</td>
<td>0.6</td>
<td>0.0638 %</td>
</tr>
<tr>
<td>16</td>
<td>0.6</td>
<td>0.8</td>
<td>0.0537 %</td>
</tr>
<tr>
<td>17</td>
<td>0.7</td>
<td>1.1</td>
<td>0.0325 %</td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>4.7</td>
<td>0.1013 %</td>
</tr>
<tr>
<td>19</td>
<td>0.7</td>
<td>0.6</td>
<td>0.0817 %</td>
</tr>
<tr>
<td>20</td>
<td>0.6</td>
<td>0.8</td>
<td>0.0475 %</td>
</tr>
<tr>
<td>21</td>
<td>0.6</td>
<td>1.1</td>
<td>0.0279 %</td>
</tr>
<tr>
<td>22</td>
<td>0.9</td>
<td>3.2</td>
<td>0.0030 %</td>
</tr>
<tr>
<td>23</td>
<td>0.6</td>
<td>0.7</td>
<td>0.0855 %</td>
</tr>
<tr>
<td>24</td>
<td>0.6</td>
<td>0.8</td>
<td>0.0484 %</td>
</tr>
<tr>
<td>25</td>
<td>0.6</td>
<td>1.1</td>
<td>0.0305 %</td>
</tr>
<tr>
<td>26</td>
<td>0.9</td>
<td>3.2</td>
<td>0.0033 %</td>
</tr>
<tr>
<td>27</td>
<td>0.6</td>
<td>0.7</td>
<td>0.0845 %</td>
</tr>
<tr>
<td>28</td>
<td>0.6</td>
<td>0.8</td>
<td>0.0449 %</td>
</tr>
<tr>
<td>29</td>
<td>0.6</td>
<td>1.1</td>
<td>0.0257 %</td>
</tr>
<tr>
<td>30</td>
<td>0.5</td>
<td>4.7</td>
<td>0.0715 %</td>
</tr>
</tbody>
</table>

Note in this study that we have assumed perfect channel estimation. In reality, the quality of channel estimation can have a negative effect on signal detection, whether full CSI signalling or our proposed signalling is used. The effect of imperfect channel estimation on our proposed signalling strategy is certainly interesting, and thus, it can be studied as a further work.

VI. Conclusion

This letter looked at the detection of MMR networks with limited CSI. The proposed system enjoys a significant reduction in the signalling overhead at each relay (at least 50%). We have further proposed the frameworks for hard ML as well as soft MAP detection at the receiver when our signalling strategy is utilized. We have shown that this significant reduction in signalling overhead does not entail any performance penalty in terms of hard ML detection. For the case of soft MAP detection we have proposed two approximate MAP schemes that avoid the need of necessary integrations at the destination.

For dual-hop networks, both our schemes perform almost identical or very close to the full CSI signalling case. For the multi-hop case (i.e., more than one relay), we have shown that our Gaussian approximation scheme extended nicely to the multi-hop case (i.e., more than one relay). We have shown that this significant reduction in signalling overhead does not entail any performance penalty in terms of hard ML detection.

Fig. 5 shows the FER performance of a three-hop network as a function of \( SNR_1 \). When \( \sigma^2_1/\sigma^2_2 = -30 \text{ dB} \) and \( \sigma^2_1/\sigma^2_2 = 0 \text{ dB} \), we see that the performance of both the Gaussian approximation and the full CSI signalling case is almost identical, which is expected when \( SNR_1 \) and \( SNR_2 \) dominate \( SNR_3 \). At \( \sigma^2_1/\sigma^2_2 = \sigma^2_1/\sigma^2_3 = 0 \text{ dB} \), it is observed that the Gaussian approximation performs less than 1dB away from the full CSI signalling case. In the case of \( \sigma^2_1/\sigma^2_2 = \sigma^2_1/\sigma^2_3 = 30 \text{ dB} \) we see an extreme case of performance degradation for the Gaussian approximation. That is because when \( SNR_1 \) is much worse than the other hops, the PDF of \( \tilde{n} \) deviates away from a Gaussian PDF, and that is when the Gaussian approximation loses some of its accuracy. Note here that the performance loss for the \( \sigma^2_1/\sigma^2_2 = \sigma^2_1/\sigma^2_3 = 0 \text{ dB} \) case and \( \sigma^2_1/\sigma^2_2 = \sigma^2_1/\sigma^2_3 = 30 \text{ dB} \) case comes at a significant 60% reduction in relay pilot signalling overhead.

For dual-hop networks, both our schemes perform almost identical or very close to the full CSI signalling case. For the multi-hop case (i.e., more than one relay), we have shown that our Gaussian approximation scheme extended nicely to the multi-hop case and demonstrated that the performance is comparable to the full CSI signalling case especially when \( SNR_1 \ll SNR_K \).

The Gaussian approximation suffers a small loss of about 0.5 dB, whereas the MVT-based approximation performs almost identically to the full CSI signalling strategy. In the last case of \( SNR_1 \ll SNR_2 \) (\( \sigma^2_1/\sigma^2_2 = 30 \text{ dB} \)), the Gaussian approximation loses about 1.7 dB in performance while the MVT-based approximation loses about 0.5 dB. Note that this small performance loss in the \( \sigma^2_1/\sigma^2_2 = 0 \text{ dB} \) and \( \sigma^2_1/\sigma^2_2 = 30 \text{ dB} \) cases comes at a significant 50% reduction in pilot signalling overhead at the relay. Remember that this reduction is particularly effective when channels vary quickly, making CSI signalling an increased burden.
REFERENCES


