[19] D. Sundararajan, *The Discrete Fourier Transform: Theory, Algorithms and Applications*. Singapore: World Scientific, 2001.
[20] M. Buhmann, *Radial Basis Functions*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[21] F. L. Bookstein, "Principle warps: Thin-plate splines and the decomposition and deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
[22] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," 1997.
[23] Specification 3GPP TS 45.005 Radio Transmission and Reception.

Power Allocation for Channel Estimation and Performance of Mismatched Decoding in Wireless Relay Networks

Duy H. N. Nguyen, *Student Member, IEEE*, and
Ha H. Nguyen, *Senior Member, IEEE*

***Abstract*—This paper is concerned with power allocation (PA) among the source and the relays of wireless relay networks to minimize the mean-square error (MSE) of the channel estimation when distributed space–time coding (DSTC) is applied. The optimal PA scheme is numerically obtained by means of geometric programming for the least-squares (LS) estimator, and a closed-form near-optimal PA scheme is also suggested. The impact of imperfect channel estimation on the error performance of DSTC is analyzed for both the LS and linear minimum MSE (LMMSE) channel estimators. It is proved that mismatched decoding of DSTC is able to achieve the same diversity order as coherent decoding of DSTC. Furthermore, when the closed-form PA obtained in the training phase is applied to the transmission phase, mismatched decoding is able to achieve a significant coding gain over the equal-PA scheme (which assigns half of the total power to the source and equally shares the other half to all the relays).**

***Index Terms*—Channel estimation, distributed space–time coding (DSTC), diversity order, geometric programming (GP), power allocation (PA).**

I. Introduction

Distributed space–time coding (DTSC) [1]–[4] has been proposed for wireless relay networks, where the relays cooperate with each other, simulate a virtual array of transmit antennas, and perform space–time coding on the source signal. By exploiting the spatial diversity provided by DSTC, it is well known that the transmission reliability of the source signal over a wireless relay network can be significantly improved. While most of the existing works on DSTC in the literature consider the relay networks with perfect channel state information (CSI) at the destination [1]–[4], only a few of them study the networks with imperfect CSI.

Mismatched decoding with imperfect channel estimation is investigated in [5] for a network with one relay. By combining with the direct source-to-destination $(S \to D)$ transmission, it is shown that the system is able to achieve a diversity order of 2. Channel estimation in a single-relay network has also been reported in [6]. Reference [7] has recently considered channel estimation and optimal training design for multiple-relay networks, where the channel variance of each link can take on any value. The optimal training at the source and the relays are presented in [7] to minimize the channel estimation mean-square error (MSE). However, [7] does not show how to optimally allocate the power among the source and the relays to further minimize the MSE, nor does it provide a diversity analysis of the mismatched decoder. These two important issues shall be examined in this paper.

Manuscript received March 25, 2010; revised July 20, 2010; accepted September 3, 2010. Date of publication September 27, 2010; date of current version November 12, 2010. The review of this paper was coordinated by Dr. C. Ling.

D. H. N. Nguyen is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: huu.n.nguyen@mail.mcgill.ca).

H. H. Nguyen is with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: ha.nguyen@usask.ca).

Digital Object Identifier 10.1109/TVT.2010.2080296

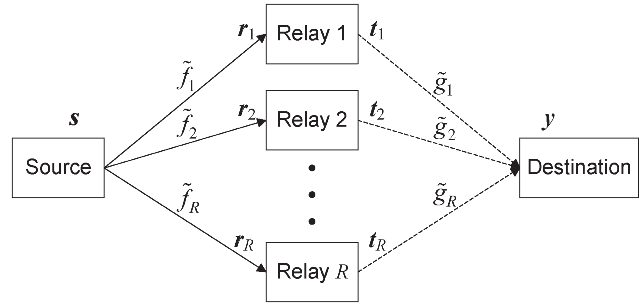

Fig. 1. Block diagram of a wireless relay network with DSTC.

This paper first studies power allocation (PA) schemes to minimize the MSE of the channel estimate, which is obtained with either the least-squares (LS) or the linear minimum MSE (LMMSE) criterion. With the LS criterion, we consider a geometric programming (GP)-based approach to find the optimal PA scheme. We also propose a closed-form PA scheme, whose performance is very close to that of the optimal scheme. Interestingly, this closed-form near-optimal scheme turns out to be same as the optimal PA scheme that has been recently proposed in [8] to maximize the signal-to-noise ratio (SNR) of the coherent DSTC at the destination under the minimum "amount of fading" constraint. With regard to the LMMSE criterion, although no optimal solution is found, the proposed PA schemes obtained under the LS criterion can be readily applied as suboptimal solutions. In Section IV, we prove that the mismatched decoder of DSTC that uses the imperfect channel estimation is able to achieve the *same* diversity order as the coherent decoder (which has the perfect channel estimation).

Notations: Superscripts $(\cdot)^{\mathcal{T}}$ and $(\cdot)^{\dagger}$ stand for transpose and complex conjugate transpose operations, respectively; $\boldsymbol{I}_M$ is an $M \times M$ identity matrix; $\mathrm{tr}(\cdot)$ denotes the trace of a square matrix; $\mathbb{E}_x[\cdot]$ indicates the expectation of random variable x; $\mathcal{CN}(0, \sigma^2)$ denotes a circularly symmetric complex Gaussian random variable with variance σ^2.

II. System Model

Consider a wireless relay network with $R + 2$ nodes, as illustrated in Fig. 1. The system has one source node, one destination node, and R relay nodes. Each node is equipped with only one antenna, which is used for transmission and reception in the half-duplex mode. Assume that there is no direct link from the source to the destination, as all signals from the source are relayed to arrive at the destination. Let $\tilde{f}_i \sim \mathcal{CN}(0, \sigma^2_{F_i})$ and $\tilde{g}_i \sim \mathcal{CN}(0, \sigma^2_{G_i})$ be the channel coefficients from the source to the ith relay and from the ith relay to the destination, respectively, for $i = 1, \ldots, R$. These coefficients are assumed to be independent of each other. It is further assumed that $\tilde{f}_i$ and $\tilde{g}_i$, $i = 1, \ldots, R$ remain constant over the coherence time $T_C = 2T$, which includes both training time T and data-transmission time T. These coefficients then take on independent values over the next coherence time interval. As shall be seen shortly, the assumption of equal training and data intervals is for simplicity, which allows the same processing to be done at the relays in both training and data-transmission phases.

Let $\mathcal{S} = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}$ be the codebook consisting of L distinguished codewords of length T employed by the source, where $\boldsymbol{s}_l^\dagger \boldsymbol{s}_l = 1$, for $l = 1, \ldots, L$. Suppose that $\boldsymbol{s} \in \mathcal{S}$ is the information codeword that the source wants to send to the destination. In the first stage, the source transmits vector $\sqrt{P_0 T}\boldsymbol{s}$ over T symbol intervals such that P_0 is the average power per transmission. The received signal at relay i can be written as

$$\boldsymbol{r}_i = \sqrt{P_0 T}\tilde{f}_i \boldsymbol{s} + \boldsymbol{w}_i \tag{1}$$

where the noise vector $\boldsymbol{w}_i$ contains independent and identically distributed (i.i.d.) $\mathcal{CN}(0, N_0)$ random variables. The amplify-and-forward (AF) protocol [3] with linear signal processing is used at each relay. In particular, similar to [3], a *unitary relaying matrix* $\boldsymbol{A}_i$ of size $T \times T$ is used to linearly process the received signal at the ith relay and form the retransmitted signal as

$$\boldsymbol{t}_i = \sqrt{\frac{P_i}{P_0\sigma_{F_i}^2 + N_0}}\boldsymbol{A}_i \boldsymbol{r}_i = \sqrt{\varepsilon_i}\boldsymbol{A}_i\boldsymbol{r}_i, \qquad i = 1, \ldots, R \tag{2}$$

where the normalization factor $\varepsilon_i = P_i/(P_0\sigma_{F_i}^2 + N_0)$ maintains the average transmitted power of P_i at the ith relay. Let $\boldsymbol{w}$, whose elements are i.i.d. $\sim \mathcal{CN}(0, N_0)$, represent the additive white Gaussian noise vector at the destination. With perfectly synchronized transmissions from the relays, the received signal at the destination can be written as

$$\boldsymbol{y}_S = \sum_{i=1}^{R} \tilde{g}_i \boldsymbol{t}_i + \boldsymbol{w} = \sqrt{P_0 T}\boldsymbol{X}_S \boldsymbol{\Lambda}\boldsymbol{h} + \tilde{\boldsymbol{w}}_S \tag{3}$$

where

$$\begin{aligned}
\boldsymbol{X}_S &= [\boldsymbol{A}_1\boldsymbol{s}, \ldots, \boldsymbol{A}_R\boldsymbol{s}]\\
\boldsymbol{\Lambda} &= \mathrm{diag}\left(\sqrt{\varepsilon_1\sigma_{F_1}^2\sigma_{G_1}^2}, \ldots, \sqrt{\varepsilon_R\sigma_{F_R}^2\sigma_{G_R}^2}\right)\\
\boldsymbol{h} &= [f_1 g_1, \ldots, f_R g_R]^{\mathcal{T}}\\
\tilde{\boldsymbol{w}}_S &= \sum_{i=1}^{R}\sqrt{\varepsilon_i\sigma_{G_i}^2}g_i\boldsymbol{A}_i\boldsymbol{w}_i + \boldsymbol{w}.
\end{aligned} \tag{4}$$

Note that f_i and g_i are the normalized channel coefficients of $\tilde{f}_i$ and $\tilde{g}_i$, $i = 1, \ldots, R$, respectively. They are i.i.d. $\mathcal{CN}(0, 1)$. Since $\boldsymbol{w}_i$ is circularly symmetric and $\boldsymbol{A}_i$ acts as a rotation matrix, the rotated noise vector $\boldsymbol{A}_i\boldsymbol{w}_i$ has the same distribution as that of $\boldsymbol{w}_i$. Therefore, conditioned on $\{g_i\}$, $\tilde{\boldsymbol{w}}_S$ contains i.i.d. Gaussian random variables with variance

$$\gamma = N_0\left(1 + \sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2|g_i|^2\right). \tag{5}$$

In most existing studies on DSTC, it is commonly assumed that the destination has a perfect knowledge of the CSI in $\boldsymbol{h}$. With such an assumption, the coherent maximum-likelihood decoding of $\boldsymbol{X}_S$ is possible [3], [4]. However, in practical wireless relay networks, the channel vector $\boldsymbol{h}$ has to be estimated at the destination, typically via training signals. The next section investigates the channel estimation problem and shows how to optimally allocate the training power among the source and the relays to minimize the MSE of the channel estimate.

III. Optimal Power Allocation for Channel Estimation

In the AF protocol, the effective channel is a product of the source-to-relay $(S \to R)$ and relay-to-destination $(R \to D)$ channels. To estimate the effective channel, it is possible to estimate the $S \to R$ and $R \to D$ channels separately or the overall effective channel directly. As pointed out in [7], the former approach requires more training slots, as well as the transmission of the $S \to R$ channel estimation to the destination, which may be prone to error. Thus, similar to [7], this paper only considers the latter approach in estimating $\boldsymbol{h}$.

As aforementioned, channel estimation is typically performed with the help of training signals, whose location and values are known to the receiver. Assume that the source sends a training sequence $\boldsymbol{z}$ and the codeword $\boldsymbol{s}$ and both of them are affected by the same overall channel vector $\boldsymbol{h}$. After possibly rearranging the order of the transmitted symbols, the destination observes the following vector due to the transmission of the training symbols:

$$\boldsymbol{y}_T = \sqrt{P_0 T}\boldsymbol{X}_T\boldsymbol{\Lambda}\boldsymbol{h} + \tilde{\boldsymbol{w}}_T \tag{6}$$

where $\boldsymbol{X}_T = [\boldsymbol{A}_1\boldsymbol{z}, \ldots, \boldsymbol{A}_R\boldsymbol{z}]$ is the training matrix formed at the relays and known at the destination. The noise vector $\tilde{\boldsymbol{w}}_T$, which is given in a form similar to (4), has the same distribution as that of $\tilde{\boldsymbol{w}}_S$.

For the relay networks considered in this paper, where f_i and g_i, $i = 1, \ldots, R$ are independent of each other, the optimal training matrix $\boldsymbol{X}_T$ was shown to be orthogonal for both the LS and LMMSE estimation criteria [7]. In particular, $\boldsymbol{X}_T^\dagger\boldsymbol{X}_T$ is a multiple of an identity matrix. It should be pointed out that, since at least R independent measurements are needed to estimate the length-R channel vector $\boldsymbol{h}$, the training time T should be no less than the number of relays[1] R, i.e., $T \geq R$. In this paper, to simplify the processing at each relay, it is assumed that the same relaying matrix is applied to both the training sequence and the information codeword. Thus, the data-transmission time is also set at T. The channel $\boldsymbol{h}$ is therefore assumed to remain constant for a block of $T_C = 2T$ channel uses and independently change over the next block. By choosing the relaying matrix $\boldsymbol{A}_i$ as in [4] and [7], and setting $\boldsymbol{z} = 1/\sqrt{T}[1, \ldots, 1]^{\mathcal{T}} \in \mathbb{R}^T$, it is possible to have $\boldsymbol{X}_T^\dagger\boldsymbol{X}_T = \boldsymbol{I}_R$.

A. LS Estimation

Dividing both sides of (6) by $\sqrt{P_0 T}$, one has the following equivalent input–output model for the training phase:

$$\bar{\boldsymbol{y}}_T = \frac{\boldsymbol{y}_T}{\sqrt{P_0 T}} = \boldsymbol{X}_T\boldsymbol{\Lambda}\boldsymbol{h} + \sqrt{\zeta}\bar{\boldsymbol{w}}_T \tag{7}$$

where, conditioned on $\{g_i\}$, $\bar{\boldsymbol{w}}_T$ contains i.i.d. $\mathcal{CN}(0, 1)$ random variables, and

$$\begin{aligned}
\zeta = \frac{\gamma}{P_0 T} &= \frac{N_0\left(1 + \sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2|g_i|^2\right)}{P_0 T}\\
&= \zeta_0\left(1 + \sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2|g_i|^2\right).
\end{aligned} \tag{8}$$

[1]This condition is similar to the requirement of having the training time at least equal to the number of transmit antennas in a multiple-input–multiple-output (MIMO) system [9].

Note that $\zeta_0 = N_0/(P_0T)$ would be the inverse of the SNR at each relay if there was no channel fading. For convenience, ζ_0^{-1} shall be generally referred to as the *channel* SNR (CSNR). The LS estimation of $\boldsymbol{h}$ is [7]

$$\hat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{\Lambda}^{-1}\boldsymbol{X}_T^{\dagger}\bar{\boldsymbol{y}}_T = \boldsymbol{h} + \sqrt{\zeta}\boldsymbol{\Lambda}^{-1}\boldsymbol{X}_T^{\dagger}\bar{\boldsymbol{w}}_T \tag{9}$$

where the property $\boldsymbol{X}_T^{\dagger}\boldsymbol{X}_T = \boldsymbol{I}_R$ has been used. Conditioned on $\{g_i\}$, the covariance of the estimation error $\boldsymbol{\Delta}_h = \hat{\boldsymbol{h}}_{\mathrm{LS}} - \boldsymbol{h}$ can be shown to be $\mathrm{cov}(\boldsymbol{\Delta}_h|\{g_i\}) = \zeta\boldsymbol{\Lambda}^{-2}$ [7]. Then, averaging over $\{g_i\}$, the MSE in estimating $\boldsymbol{h}$ is

$$\mathrm{cov}(\boldsymbol{\Delta}_h) = \bar{\zeta}\boldsymbol{\Lambda}^{-2} \tag{10}$$

where

$$\bar{\zeta} = \mathop{\mathbb{E}}_{\{g_i\}}[\zeta] = \zeta_0\left(1 + \sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2\right). \tag{11}$$

The total MSE with the LS estimation is given by

$$\mathrm{tr}(\bar{\zeta}\boldsymbol{\Lambda}^{-2}) = \bar{\zeta}\sum_{i=1}^{R}\frac{1}{\varepsilon_i\sigma_{F_i}^2\sigma_{G_i}^2}. \tag{12}$$

It is clear that the total MSE depends on how the total power, which is denoted by P, is allocated among the source and the relays. Our objective is to find the optimal PA to minimize the total MSE. The optimization problem is stated as

$$\begin{aligned} &\underset{P_0,P_1,\ldots,P_R}{\text{minimize}} && \frac{1}{P_0}\left(1+\sum_{i=1}^{R}\frac{P_i\sigma_{G_i}^2}{P_0\sigma_{F_i}^2+N_0}\right)\sum_{i=1}^{R}\frac{P_0\sigma_{F_i}^2+N_0}{P_i\sigma_{F_i}^2\sigma_{G_i}^2}\\ &\text{subject to} && \sum_{i=0}^{R}P_i \le P. \end{aligned} \tag{13}$$

It is observed that the optimization problem is not convex due to the nonconvexity of the objective function in $P_0, P_1, \ldots, P_R$. Thus, it might be difficult to find the global optimal solution in $P_0, P_1, \ldots, P_R$ directly. In the following, we investigate a GP approach to transform the problem into a convex GP. We then present a closed-form near-optimal solution, which is more tractable to study the diversity order of the training code in Section IV.

1) PA via GP: Due to the component $P_0\sigma_{F_i}^2 + N_0$ in the denominator of the objective function, the optimization problem (13) is not a GP in $P_0, P_1, \ldots, P_R$.

However, if $P_0, \varepsilon_1, \ldots, \varepsilon_R$ are treated as variables, the problem can be readily transformed into a convex GP, i.e.,

$$\begin{aligned} &\underset{P_0,\varepsilon_1,\ldots,\varepsilon_R}{\text{minimize}} && \frac{1}{P_0}\left(1+\sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2\right)\sum_{i=1}^{R}\frac{1}{\varepsilon_i\sigma_{F_i}^2\sigma_{G_i}^2}\\ &\text{subject to} && P_0+\sum_{i=1}^{R}\varepsilon_i\left(P_0\sigma_{F_i}^2+N_0\right) \le P. \end{aligned} \tag{14}$$

As $1+\sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2$ is linear, and $\sum_{i=1}^{R}(1/\varepsilon_i\sigma_{F_i}^2\sigma_{G_i}^2)$ is a posynomial, the objective function is also a posynomial [10]. In addition, the constraint is a posynomial. Thus, under this formulation, the optimization is a typical GP. The optimal solution to the problem can be easily obtained by any GP solver, such as `cvx` [11].

2) Suboptimal Closed-Form Solution: While the optimal solution to problem (13) can be found by GP, it is also useful to obtain a suboptimal closed-form solution. As shall be seen in Section IV, such a closed-form solution is more convenient in diversity analysis of the mismatched decoder. Introduce the constraint $\varepsilon_1\sigma_{F_1}^2\sigma_{G_1}^2 = \cdots = \varepsilon_R\sigma_{F_R}^2\sigma_{G_R}^2$ into problem (14) and consider the following new optimization problem:

$$\begin{aligned} &\underset{P_0,\varepsilon_1,\ldots,\varepsilon_R}{\text{minimize}} && \frac{1}{P_0}\left(1+\sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2\right)\sum_{i=1}^{R}\frac{1}{\varepsilon_i\sigma_{F_i}^2\sigma_{G_i}^2}\\ &\text{subject to} && P_0+\sum_{i=1}^{R}\varepsilon_i\left(P_0\sigma_{F_i}^2+N_0\right) \le P\\ &&& \varepsilon_1\sigma_{F_1}^2\sigma_{G_1}^2 = \cdots = \varepsilon_R\sigma_{F_R}^2\sigma_{G_R}^2. \end{aligned} \tag{15}$$

Since the feasible set of problem (15) is smaller than that of problem (14), the optimal value of problem (15) is inferior to the optimal value of problem (14). However, numerous simulations show that the gap between the two optimal values is very small. This characteristic can be loosely explained as follows: Extract the objective function of (14) and group it into two summations as

$$\begin{aligned} &\frac{1}{P_0}\left(1+\sum_{i=1}^{R}\varepsilon_i\sigma_{G_i}^2\right)\sum_{i=1}^{R}\frac{1}{\varepsilon_i\sigma_{F_i}^2\sigma_{G_i}^2}\\ &\qquad= \sum_{i=1}^{R}\frac{1}{P_0\varepsilon_i\sigma_{F_i}^2\sigma_{G_i}^2} + \sum_{i,j}\frac{\sigma_{G_j}^2}{P_0\sigma_{F_i}^2\sigma_{G_i}^2}\frac{\varepsilon_j}{\varepsilon_i}. \end{aligned}$$

It can be seen that a small perturbation on any ε_i will have a considerable effect on the first summation. In addition, if the first summation were to be minimized alone, one would get $\varepsilon_1\sigma_{F_1}^2\sigma_{G_1}^2 = \cdots = \varepsilon_R\sigma_{F_R}^2\sigma_{G_R}^2$ as the solution. Now, if this constraint is imposed, one equivalently has

$$\frac{P_1}{\frac{P_0\sigma_{F_1}^2+N_0}{\sigma_{F_1}^2\sigma_{G_1}^2}} = \cdots = \frac{P_R}{\frac{P_0\sigma_{F_R}^2+N_0}{\sigma_{F_R}^2\sigma_{G_R}^2}} = \frac{\sum_{i=1}^{R}P_i}{\sum_{i=1}^{R}\frac{P_0\sigma_{F_i}^2+N_0}{\sigma_{F_i}^2\sigma_{G_i}^2}}. \tag{16}$$

Define $a=\sum_{i=1}^{R}(1/\sigma_{F_i}^2)$, $b=\sum_{i=1}^{R}(1/\sigma_{G_i}^2)$, and $c=\sum_{i=1}^{R}(N_0/\sigma_{F_i}^2\sigma_{G_i}^2)$. It is noted that the power constraint $\sum_{i=0}^{R}P_i \le P$ must be met with equality. Then, the constraint in (16) dictates the following PA among the R relays:

$$\begin{aligned} P_i &= \frac{P-P_0}{P_0b+c}\cdot\frac{P_0\sigma_{F_i}^2+N_0}{\sigma_{F_i}^2\sigma_{G_i}^2}\\ \varepsilon_i &= \frac{P-P_0}{P_0b+c}\cdot\frac{1}{\sigma_{F_i}^2\sigma_{G_i}^2}, \qquad i=1,\ldots,R. \end{aligned} \tag{17}$$

Substituting $\varepsilon_1, \ldots, \varepsilon_R$ from (17) into the objective function of (15), the optimization reduces to finding P_0 that minimizes the following total MSE:

$$\begin{aligned} f(P_0) &= \frac{1}{P_0}\left(1+\sum_{i=1}^{R}\frac{P-P_0}{(P_0b+c)\sigma_{F_i}^2}\right)R\frac{P_0b+c}{P-P_0}\\ &= R\left(\frac{P_0b+c}{P_0(P-P_0)}+\frac{a}{P_0}\right). \end{aligned} \tag{18}$$

Ignoring the constant factor R, the equivalent optimization problem to (15) is

$$\begin{aligned} \underset{P_0}{\text{minimize}} \quad & f(P_0) = \frac{P_0 b + c}{P_0(P - P_0)} + \frac{a}{P_0} \\ \text{subject to} \quad & 0 < P_0 < P. \end{aligned} \tag{19}$$

Since the second derivative of the objective function is always positive in the domain of P_0, the objective function is convex. This problem can be easily solved, and the solution is given as follows:

$$P_0 = \begin{cases} \frac{\sqrt{(Pa+c)(Pb+c)}-(Pa+c)}{b-a}, & \text{if } b \neq a \\ P/2, & \text{if } b = a. \end{cases} \tag{20}$$

The allocated power at each relay is given by (17) accordingly.

At this point, it is interesting to point out that the proposed suboptimal PA scheme in (17) and (20) for MSE minimization with the LS estimation is exactly the same as the PA scheme obtained in [8] for average SNR maximization in the data-transmission phase. The proposed scheme in [8] attempts to maximize the average SNR at the destination while minimizing the so-called "amount of fading" in a wireless relay network. In fact, the condition to achieve the minimum amount of fading turns out to be the same as in (16), whereas the average SNR at the destination is simply proportional to the reciprocal of the total MSE with the LS estimation. Moreover, such a PA scheme has been shown in [8] to enable the maximum diversity order to be achieved with the coherent DTSC in an arbitrary relay network as the SNR goes to infinity.

B. LMMSE Estimation

The LMMSE estimation requires the second-order statistics of the channel to be known at the destination. Letting $\boldsymbol{\Sigma}_h$ be the covariance matrix of $\boldsymbol{h}$, $\boldsymbol{\Sigma}_h = \boldsymbol{I}_R$, as $\{f_i\}$ and $\{g_i\}$ are i.i.d. $\mathcal{CN}(0,1)$ random variables. Based on the training phase model in (7), the LMMSE estimation yields [7]

$$\begin{aligned} \hat{\boldsymbol{h}}_{\text{LMMSE}} &= \boldsymbol{\Sigma}_h \boldsymbol{\Lambda} \boldsymbol{X}_T^{\dagger} \left(\boldsymbol{X}_T \boldsymbol{\Lambda} \boldsymbol{\Sigma}_h \boldsymbol{\Lambda} \boldsymbol{X}_T^{\dagger} + \bar{\zeta} \boldsymbol{I}_T \right)^{-1} \bar{\boldsymbol{y}}_T \\ &= (\boldsymbol{\Lambda}^2 + \bar{\zeta} \boldsymbol{I}_R)^{-1} \boldsymbol{\Lambda} \boldsymbol{X}_T^{\dagger} \bar{\boldsymbol{y}}_T = \boldsymbol{B} \hat{\boldsymbol{h}}_{\text{LS}} \end{aligned} \tag{21}$$

where $\boldsymbol{B} = (\boldsymbol{\Lambda}^2 + \bar{\zeta}\boldsymbol{I}_R)^{-1}\boldsymbol{\Lambda}^2 = (\boldsymbol{I}_R + \bar{\zeta}\boldsymbol{\Lambda}^{-2})^{-1}$ is considered as a biasing matrix to the unbiased estimator $\hat{\boldsymbol{h}}_{\text{LS}}$ of $\boldsymbol{h}$. It can be seen that if the PA scheme meets the constraint in (16), $\boldsymbol{\Lambda}$ and $\boldsymbol{B}$ are multiples of an identity matrix. As a result, $\hat{\boldsymbol{h}}_{\text{LMMSE}}$ is a scaled version of $\hat{\boldsymbol{h}}_{\text{LS}}$.

With an orthogonal training matrix, the covariance of the estimation error $\boldsymbol{\Delta}_h = \hat{\boldsymbol{h}}_{\text{LMMSE}} - \boldsymbol{h}$ can be found as

$$\text{cov}(\boldsymbol{\Delta}_h) = \left(\boldsymbol{\Sigma}_h^{-1} + \frac{1}{\bar{\zeta}} \boldsymbol{\Lambda}^2 \right)^{-1} = \left(\boldsymbol{I}_R + \frac{1}{\bar{\zeta}} \boldsymbol{\Lambda}^2 \right)^{-1}. \tag{22}$$

The total MSE under the LMMSE estimation is then given by

$$\text{tr}\left(\text{cov}(\boldsymbol{\Delta}_h)\right) = \sum_{i=1}^{R} \frac{1}{1 + \varepsilon_i \sigma_{F_i}^2 \sigma_{G_i}^2 / \bar{\zeta}}. \tag{23}$$

Due to the component $1 + \varepsilon_i \sigma_{F_i}^2 \sigma_{G_i}^2/\bar{\zeta}$ in the denominator, it is not possible to transform (23) into a posynomial. Consequently, the problem of minimizing the total MSE with the LMMSE estimation cannot be transformed into a GP, which makes it much more difficult to find the optimal solution. However, it is clear from (12) and (23) that the total MSE in the LMMSE estimation is always less than that in the LS estimation if the same PA scheme is applied. This is an intuitively satisfying result since the LMMSE estimation is optimum with regard to the MSE measure. Nevertheless, at high SNR, the total MSEs with the two estimation schemes become virtually indifferent. Thus, both the optimal and closed-form near-optimal solutions considered for the LS estimation in Section III-A can also be applied as suboptimal schemes for the LMMSE estimation.

IV. Diversity Analysis of Mismatched Decoding

This section analyzes the diversity order of mismatched decoding with imperfect channel estimation. The analysis follows a similar procedure in [9, Sec. 4] for point-to-point MIMO systems, and it is performed in the large CSNR regime, i.e., when CSNR $\to \infty$ or $\zeta_0 \to 0$ (asymptotic analysis). It is important to point out that both ζ in (8) and $\bar{\zeta}$ in (11) are proportional to ζ_0; hence, they are on the same order of ζ_0. The analysis will show that the mismatched decoder of DSTC is able to achieve the same diversity order as that of the coherent decoder if the same PA scheme is applied.

First, for coherent decoding of DSTC, it is assumed that the channel $\boldsymbol{h}$ is perfectly known at the destination. The pairwise error probability (PEP) of mistaking the transmitted codeword $\boldsymbol{s}$ by $\hat{\boldsymbol{s}}$, i.e., mistaking $\boldsymbol{X}_S$ by $\hat{\boldsymbol{X}}_S$, is given by [3]

$$\begin{aligned} \mathbb{P}(\boldsymbol{X}_S \to \hat{\boldsymbol{X}}_S) &= \underset{\{f_i\},\{g_i\}}{\mathbb{E}} \mathbb{P}\left(\boldsymbol{X}_S \to \hat{\boldsymbol{X}}_S | \{f_i\}, \{g_i\} \right) \\ &= \underset{\{f_i\},\{g_i\}}{\mathbb{E}} \left[Q\left(\sqrt{\frac{\|\boldsymbol{\Delta}_S \boldsymbol{\Lambda} \boldsymbol{h}\|^2}{2\zeta}} \right) \right] \end{aligned} \tag{24}$$

where $\boldsymbol{\Delta}_S = \boldsymbol{X}_S - \hat{\boldsymbol{X}}_S$. Now, for mismatched decoding, the destination uses the estimated CSI $\hat{\boldsymbol{h}}$ in the same way as with the perfect CSI $\boldsymbol{h}$. The decoder performs

$$\hat{\boldsymbol{X}}_S = \arg\min_{\boldsymbol{X}_S} \|\bar{\boldsymbol{y}}_S - \boldsymbol{X}_S \boldsymbol{\Lambda} \hat{\boldsymbol{h}}\|^2. \tag{25}$$

It is noted that under the LMMSE channel estimation in (21), the Taylor series expansion of the biasing matrix is $\boldsymbol{B} = (\boldsymbol{I}_R + \bar{\zeta}\boldsymbol{\Lambda}^{-2})^{-1} = \boldsymbol{I}_R - O(\zeta_0 \boldsymbol{\Lambda}^{-2})$. This also implies that $\boldsymbol{B} \to \boldsymbol{I}_R$ when $\zeta_0 \to 0$.

Thus, under either the LS or LMMSE estimation, the channel estimate $\hat{\boldsymbol{h}}$ can be expressed as

$$\hat{\boldsymbol{h}} = \boldsymbol{h} + \sqrt{\zeta} \boldsymbol{\Lambda}^{-1} \boldsymbol{X}_T^{\dagger} \bar{\boldsymbol{w}}_T - O(\zeta_0 \boldsymbol{\Lambda}^{-2}). \tag{26}$$

The mismatched metric for codeword $\hat{\boldsymbol{X}}_S$ is

$$\begin{aligned} &\|\bar{\boldsymbol{y}}_S - \hat{\boldsymbol{X}}_S \boldsymbol{\Lambda} \hat{\boldsymbol{h}}\|^2 \\ &\quad = \Big\| \boldsymbol{X}_S \boldsymbol{\Lambda} \boldsymbol{h} + \sqrt{\zeta} \bar{\boldsymbol{w}}_S \\ &\qquad - \hat{\boldsymbol{X}}_S \boldsymbol{\Lambda} \left[\boldsymbol{h} + \sqrt{\zeta} \boldsymbol{\Lambda}^{-1} \boldsymbol{X}_T^{\dagger} \bar{\boldsymbol{w}}_T - O(\zeta_0 \boldsymbol{\Lambda}^{-2}) \right] \Big\|^2 \\ &\quad = \left\| \boldsymbol{\Delta}_S \boldsymbol{\Lambda} \boldsymbol{h} + \sqrt{\zeta} \left(\bar{\boldsymbol{w}}_S - \hat{\boldsymbol{X}}_S \boldsymbol{X}_T^{\dagger} \bar{\boldsymbol{w}}_T \right) - O(\zeta_0 \boldsymbol{\Lambda}^{-1}) \right\|^2 \end{aligned}$$

$$= \|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2 + 2\sqrt{\zeta}\mathrm{Re}\left\{\bar{\boldsymbol{w}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\} - 2\sqrt{\zeta}\mathrm{Re}\left\{\bar{\boldsymbol{w}}_T^\dagger\boldsymbol{X}_T\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\} + O(\zeta_0). \quad (27)$$

Note that, since $\bar{\boldsymbol{w}}_S$ and $\bar{\boldsymbol{w}}_T$ contain i.i.d. $\mathcal{CN}(0,1)$, the term $\|\sqrt{\zeta}(\bar{\boldsymbol{w}}_S - \hat{\boldsymbol{X}}_S\boldsymbol{X}_T^\dagger\bar{\boldsymbol{w}}_T)\|^2 = \zeta\|\bar{\boldsymbol{w}}_S - \hat{\boldsymbol{X}}_S\boldsymbol{X}_T^\dagger\bar{\boldsymbol{w}}_T\|^2 \to 0$ almost surely as $\zeta_0 \to 0$. Therefore, this term can be included in $O(\zeta_0)$ in (27). The asymptotic PEP of the mismatched decoder can be calculated as

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{X}_S \to \hat{\boldsymbol{X}}_S)\\
&= \mathbb{P}\left(\left\|\bar{\boldsymbol{y}}_S - \hat{\boldsymbol{X}}_S\boldsymbol{\Lambda}\hat{\boldsymbol{h}}\right\|^2 < \left\|\bar{\boldsymbol{y}}_S - \boldsymbol{X}_S\boldsymbol{\Lambda}\hat{\boldsymbol{h}}\right\|^2\right)\\
&= \mathbb{P}\Bigg(\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2 + 2\sqrt{\zeta}\mathrm{Re}\left\{\bar{\boldsymbol{w}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\}\\
&\qquad - 2\sqrt{\zeta}\mathrm{Re}\left\{\bar{\boldsymbol{w}}_T^\dagger\boldsymbol{X}_T\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\} + O(\zeta_0) < 0\Bigg)\\
&= \mathop{\mathbb{E}}_{\{f_i\},\{g_i\}}\Bigg[\mathbb{P}\Bigg(\frac{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}{\sqrt{\zeta}} + O(\sqrt{\zeta_0})\\
&\qquad < 2\mathrm{Re}\left\{\bar{\boldsymbol{w}}_T^\dagger\boldsymbol{X}_T\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\}\\
&\qquad - 2\mathrm{Re}\left\{\bar{\boldsymbol{w}}_S^\dagger\boldsymbol{\Delta}_s\boldsymbol{\Lambda}\boldsymbol{h}\right\} |\{f_i\},\{g_i\}\Bigg)\Bigg]. \qquad (28)
\end{aligned}$$

Note that, conditioned on $\{g_i\}$ and $\{f_i\}$, the elements of $\bar{\boldsymbol{w}}_T$ and $\bar{\boldsymbol{w}}_S$ are i.i.d. $\mathcal{CN}(0,1)$ random variables. Thus, $2\mathrm{Re}\{\bar{\boldsymbol{w}}_T^\dagger\boldsymbol{X}_T\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\} - 2\mathrm{Re}\{\bar{\boldsymbol{w}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\}$ is Gaussian distributed with zero mean and variance $2\|\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2 + 2\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2$, since $\boldsymbol{X}_T^\dagger\boldsymbol{X}_T = \boldsymbol{I}_R$. The probability term in (28) is just the probability that a zero-mean Gaussian random variable is bigger than some constant. Therefore, (28) can be written as

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{X}_S \to \hat{\boldsymbol{X}}_S)\\
&= \mathop{\mathbb{E}}_{\{f_i\},\{g_i\}}\left[Q\left(\sqrt{\frac{\left[\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2/\sqrt{\zeta} + O(\sqrt{\zeta_0})\right]^2}{2\left\|\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\|^2 + 2\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}}\right)\right]\\
&= \mathop{\mathbb{E}}_{\{f_i\},\{g_i\}}\left[Q\left(\sqrt{\frac{\frac{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^4}{\zeta} + 2\frac{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}{\sqrt{\zeta}}O(\sqrt{\zeta_0})}{2\left\|\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\|^2 + 2\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}}\right)\right]\\
&= \mathop{\mathbb{E}}_{\{f_i\},\{g_i\}}\left[Q\left(\sqrt{\frac{1}{2\zeta}\frac{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}{1 + \frac{\left\|\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\|^2}{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}} + O(1)}\right)\right]. \qquad (29)
\end{aligned}$$

Similar to [9], by applying the Cauchy–Schwarz inequality to the Frobenius norm [12], one has

$$1 \le 1 + \frac{\left\|\hat{\boldsymbol{X}}_S^\dagger\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\right\|^2}{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2} \le 1 + \|\hat{\boldsymbol{X}}_S\|^2. \qquad (30)$$

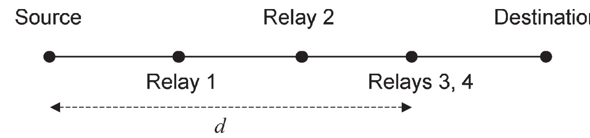

Fig. 2. Locations of relays relatively to the source and the destination.

Thus, due to the monotonic decreasing of the Q function, the asymptotic PEP is bounded as

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\{f_i\},\{g_i\}}\left[Q\left(\sqrt{\frac{1}{2\zeta}\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2 + O(1)}\right)\right] \le \mathbb{P}(\boldsymbol{X}_S \to \hat{\boldsymbol{X}}_S)\\
&\qquad \le \mathop{\mathbb{E}}_{\{f_i\},\{g_i\}}\left[Q\left(\sqrt{\frac{1}{2\zeta}\frac{\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2}{1 + \|\hat{\boldsymbol{X}}_S\|^2} + O(1)}\right)\right]. \qquad (31)
\end{aligned}$$

Note that, at high CSNR, i.e., small ζ, the constant term $O(1)$ becomes negligible compared with the term $\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2/(2\zeta)$ or the term $\|\boldsymbol{\Delta}_S\boldsymbol{\Lambda}\boldsymbol{h}\|^2/(2\zeta(1 + \|\hat{\boldsymbol{X}}_S\|^2))$. Therefore, it can be neglected as far as the diversity order analysis is concerned. Comparing the PEP expression in (24) and the bounds in (31) clearly shows that the diversity order of the mismatched decoder is the same as that of the coherent decoder if the same PA scheme is applied.

In [3], it was shown that, if $\boldsymbol{\Delta}_S = \boldsymbol{X}_S - \hat{\boldsymbol{X}}_S$ is full rank, the PEP in (24) decays on the order of $R(1 - \log\log\eta/\log\eta)$, where η is the average SNR computed at the destination. However, the derivation of the diversity order in [3] is only for the special case of a network with $\sigma_{F_i}^2 = \sigma_{G_i}^2 = 1$, $i = 1, \ldots, R$. The optimal PA scheme for such a network is the equal PA [3], which assigns half of the total power to the source and equally shares the other half with all the relays, i.e., $P_0 = P/2$, $P_1 = \cdots = P_R = P/(2R)$.

In an arbitrary network topology, where $\sigma_{F_i}^2$ and $\sigma_{G_i}^2$ can take on any values, the aforementioned equal-PA scheme is clearly suboptimal with respect to the achievable coding gain. It has been recently shown in [8] that the closed-form PA scheme in (17) and (20) can extract the maximum diversity order from the underlying coherent DSTC when SNR $\to \infty$ and can considerably outperform the equal PA in terms of coding gain. As a consequence of the PEP analysis in (31), under the PA scheme in (20), the mismatched decoder is also able to realize the maximum diversity order of the underlying DSTC and improve the coding gain over the equal-PA scheme. In addition, as the closed-form PA scheme is suboptimal to the GP-based PA scheme with the LS estimation, it is expected that the mismatched decoder in conjunction with the GP-based PA scheme is able to further enhance the coding gain.

V. Simulation Results

This section presents numerical results on the MSE of both the LS and LMMSE estimators obtained by the GP-based and closed-form near-optimal PA schemes, as well as validating our theoretical findings on the performance of mismatched decoding. We consider a relay network with four relays, as illustrated in Fig. 2. Assume that the source and the destination are located at (0, 0) and (1, 0). The first relay is located near the source at (0.25, 0), and the second relay is at midway between the source and the destination at (0.5, 0). The third and fourth relays are located at the same distance d from the source, i.e., $(d, 0)$. The fading variance is proportionally assigned to the distance between the transmit and receive terminals, taking into account the path-loss exponent, which is set at 4. Thus, if $\sigma_{G_1}^2$ is normalized to 1, then

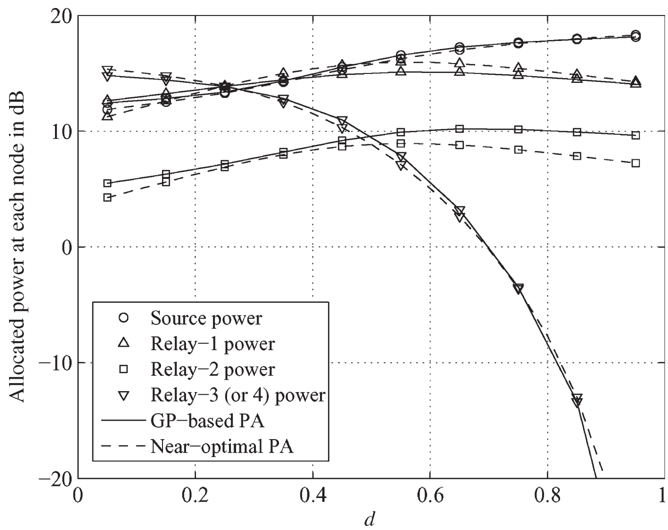

Fig. 3. Allocated power at each node (source and relays) obtained by the GP-based and closed-form near-optimal PA schemes when the location of relays 3 and 4 is varied and $P = 20$ dB.

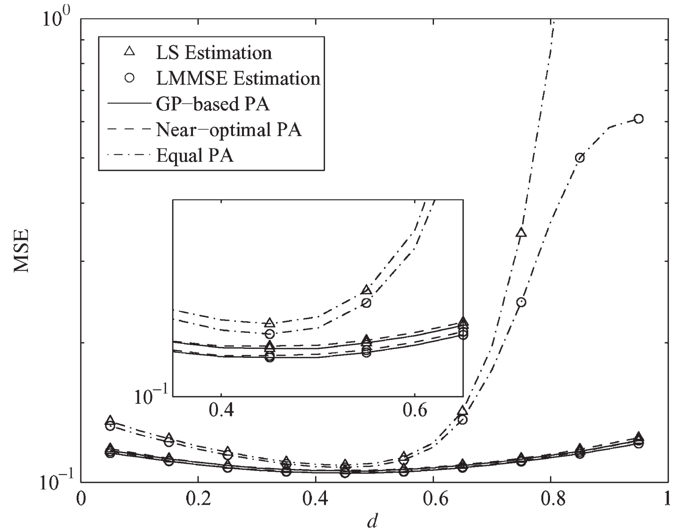

Fig. 4. Total MSE achieved with the LS and LMMSE estimators when the location of relays 3 and 4 is varied and $P = 20$ dB: GP-based, closed-form near-optimal, and equal-PA schemes.

$\sigma^2_{F_1} = 3^4$. Similarly, $\sigma^2_{F_2} = \sigma^2_{G_2} = (3/2)^4$, $\sigma^2_{F_3} = \sigma^2_{F_4} = (4d/3)^{-4}$, and $\sigma^2_{G_1} = \sigma^2_{G_2} = (4(1-d)/3)^{-4}$. The parameter N_0 is normalized to 1. The training (and data transmission) interval T is set to be equal to the number of relays R.

Fig. 3 illustrates how the total power P (set at 20 dB) is split between the source and the relays in the GP-based and closed-form PA schemes when the network topology is changed by varying the location of relays 3 and 4 (i.e., the distance d). When d is varied, the parameters a, b, and c are effectively changed, which then triggers the change to power allocated to each node by the closed-form scheme. As can be seen from the figure, the proposed closed-form solution is very close to the optimal GP-based solution for all the values of d shown in the figure. Fig. 4 then compares the total MSE achieved by the GP-based, closed-form near-optimal, and equal-PA schemes for both the LS and LMMSE estimations. While the GP-based scheme slightly outperforms the near-optimal scheme as expected, the MSEs achieved by the GP-based and near-optimal schemes are superior to the MSE achieved by the equal-PA scheme, particularly when relays 3 and 4 are placed closer to the destination.

It is worth noting that the proposed closed-form scheme is near optimal at all SNR regimes. To illustrate this, with a fixed location of relays 3 and 4 at (0.75, 0), the MSE is plotted over a wide range of the total power P in Fig. 5. It is observed again that the closed-form scheme performs practically the same as the GP-based scheme, and both schemes significantly outperform the equal-PA scheme. A closer look at Fig. 5 also confirms the fact that the MSE of the LMMSE estimation is smaller than that of the LS estimation with all the three PA schemes considered. At high CSNR, the difference is negligible, which validates the common representation of the two estimators in (26).

The symbol error rate (SER) performance of the mismatched decoder is compared with that of the coherent decoder under the three PA schemes in Fig. 6. The distributed quasi-orthogonal space–time block code is applied in the four-relay network [4]. As can be seen from the figure, the diversity order achieved with mismatched decoding is the same as that of coherent decoding, as long as the same PA scheme is applied. This agrees with our analysis in Section IV. Moreover, under either the GP-based or the near-optimal PA schemes, the mismatched decoder significantly outperforms the decoder under the equal-PA scheme by about 4.5 dB at the SER level of 10^{-4}. It is worth

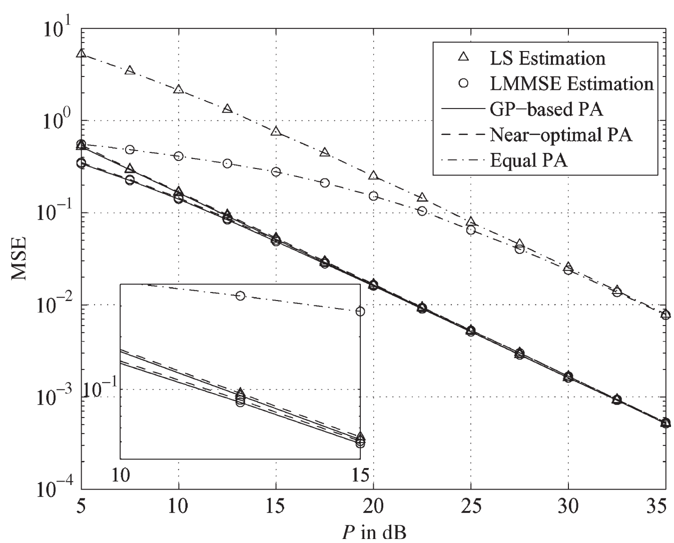

Fig. 5. Total MSE achieved with the LS and LMMSE estimators in a four-relay network with $d = 0.75$: GP-based, closed-form near-optimal, and equal-PA schemes.

noting that the proposed closed-form PA scheme performs basically the same as the GP-based PA scheme. Finally, the figure shows that the mismatched decoders perform almost the same with both the LS and LMMSE channel estimations and under all three PA schemes, particularly at high CSNR.

VI. Conclusion

This paper has considered PA schemes to minimize the total MSE of both the LS and LMMSE channel estimations for DSTC in wireless relay networks. The diversity order of the error performance of the mismatched decoder that works with the estimated channel information was also analyzed. It was shown that, with a given PA scheme in the data transmission phase, the mismatched decoder is able to achieve the *same* diversity order as the coherent decoder. In particular, if the GP-based or the proposed closed-form PA schemes obtained in the training phase are also applied to the transmission phase, then the

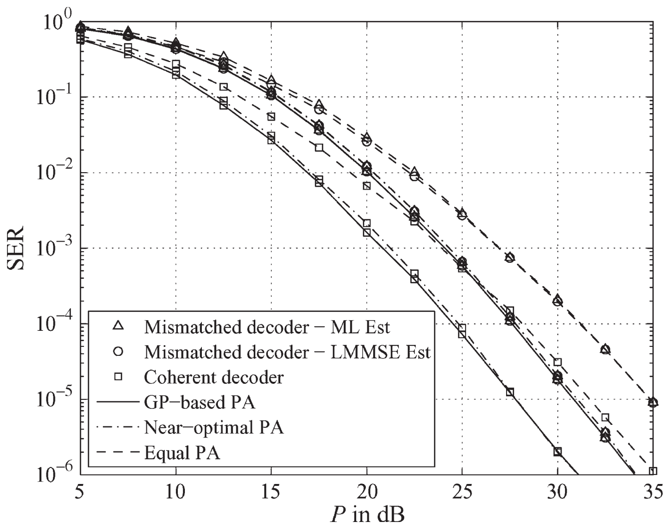

Fig. 6. Error performance of DSTC with the mismatched and coherent decoders in a four-relay network, in which relays 3 and 4 are at (0.75, 0).

mismatched decoder achieves the same *maximum* diversity order as that of the coherent decoder and significantly outperforms the equal-PA scheme in terms of coding gain.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose comments helped improve the presentation of this paper.

REFERENCES

[1] J. N. Laneman and G. W. Wornell, "Distributed space–time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[2] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, "Fading relay channels: Performance limits and space–time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[3] Y. Jing and B. Hassibi, "Distributed space–time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[4] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4106–4118, Nov. 2007.

[5] H. Mheidat and M. Uysal, "Non-coherent and mismatched-coherent receivers for distributed STBCs with amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4060–4070, Nov. 2007.

[6] C. S. Patel and G. L. Stüber, "Channel estimation for amplify and forward relay based cooperation diversity systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2348–2356, Jun. 2007.

[7] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[8] D. H. N. Nguyen, H. H. Nguyen, and H. D. Tuan, "A novel power allocation scheme for distributed space–time coding," in *Proc. IEEE ICC*, Dresden, Germany, Jun. 2009, pp. 1–6.

[9] G. Taricco and E. Biglieri, "Space–time decoding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 7, pp. 1874–1888, Jul. 2005.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[11] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, Jun. 2009. [Online]. Available: http://stanford.edu/~boyd/cvx

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

Convergence and Performance of Distributed Power Control Algorithms for Cooperative Relaying in Cellular Uplink Networks

Young-Keum Song, *Student Member, IEEE*, and Dongwoo Kim, *Senior Member, IEEE*

***Abstract*—We investigate an application of distributed power control (DPC) and opportunistic power control (OPC), respectively, to cooperative relaying systems that consist of multiple sources, multiple relays, and a single destination. Two types of relaying gains are considered for half-duplex amplify-and-forward (AF) relays: fixed and variable gains, respectively. Using half-duplex AF relays, remote users benefit by making the path loss smaller by two shorter paths. However, they lose bandwidth and suffer from additional interference due to amplification of the noise at the relays and the unnecessary signal from the relays, where power control plays an important role. We begin with redefining and extending the existing power control algorithms to be fit for the relaying environments and then, using standard techniques, construct the convergence, except when the number of variable-gain relays is even. With numerical investigation, we show the advantages and disadvantages of the power-controlled systems with relays in terms of the outage probability, the average transmit power consumption, and the transmission capacity. Simulation results show that DPC with relays achieves significant improvement in the outage performance and power consumption. On the other hand, in OPC, some negative effects arise by using the relays in the capacity since the relays increase the interference that severely affects the opportunistic capacity.**

***Index Terms*—Amplify and forward (AF), convergence, cooperative relaying, distributed power control (DPC), opportunistic power control (OPC).**

I. INTRODUCTION

Power control in wireless relaying systems has been considered for maximizing throughput with power constraints [1], [2] or minimizing power consumption under quality-of-service requirements [3], [4]. The previous works assume that the signals transmitted from multiple relays can be separated at the destination without any interference from other relays and that each relay serves only one source. The former assumption would normally result in increasing the inefficiency in using bandwidth since it requires as many orthogonal channels as the relays. The latter assumption would require a complicated procedure to allocate the relays to sources and for orthogonal transmission between multiple sources. When relays can support multiple sources and do not use orthogonal channels, relays can generate a great deal of interference, and power control becomes more important in providing adequate signal quality (mostly predefined). The power control issue is a focus of this paper.

Power control in environments with overwhelming interference but without relays has extensively been studied for the last two decades. Among the effective methods, distributed power control (DPC) [5]–[8] and opportunistic power control (OPC) [9], [10] are two

Manuscript received January 6, 2010; revised April 16, 2010 and June 9, 2010; accepted July 30, 2010. Date of publication August 16, 2010; date of current version November 12, 2010. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Science, and Technology under Grant KRF-2008-314-D00274. The review of this paper was coordinated by Prof. N. Arumugam.

The authors are with the Department of Electrical Engineering and Computer Science, Hanyang University, Ansan 425-791, Korea (e-mail: dkim@hanyang.ac.kr).

Digital Object Identifier 10.1109/TVT.2010.2066590