

# Optimal Flashy Transmission in Training-Based MISO TDD Systems

Xiangyun Zhou<sup>†</sup>, Parastoo Sadeghi<sup>‡</sup>, Tharaka A. Lamahewa<sup>‡</sup>, and Are Hjørungnes<sup>†</sup>

<sup>†</sup>UNIK - University Graduate Center, University of Oslo, Norway

<sup>‡</sup>School of Engineering, Australian National University, Australia

Email: xiangyun@unik.no, {parastoo.sadeghi, tharaka.lamahewa}@anu.edu.au, arehj@unik.no

**Abstract**—We consider a flashy transmission scheme in multiple-input single-output (MISO) systems. In order to efficiently utilize the limited energy budget, only a portion of the available time is used for the actual transmission. For block-wise transmissions, we define the flashiness as the percentage of blocks used for the actual transmission. Each transmission block consists of a reverse training phase and a data transmission phase. Using a capacity lower bound as the figure of merit, we show that the optimal training and data transmit power does not change with the energy budget as long as the flashiness is strictly less than one. We provide a simple design guideline to obtain closed-form solutions for the optimal flashiness, training power, as well as data transmit power which maximize the capacity lower bound. Our numerical results demonstrate a significant rate improvement by using the optimal flashy transmission in the low signal-to-noise ratio (SNR) regime, compared to the non-flashy scheme which uses all blocks for transmission.

## I. INTRODUCTION

Information-theoretic studies on multi-antenna transmissions have shown that the information capacity of the wireless channel can be significantly improved if the transmitter is able to utilize the channel state information (CSI) [1, 2]. To this end, various techniques were proposed to acquire CSI at the transmitter in the past few years. For non-reciprocal channels, such as in frequency-division duplex (FDD) systems, the use of limited feedback is a popular technique that allows the transmitter to adaptively design the precoder according to the instantaneous CSI [3, 4]. The overhead in obtaining the CSI at the transmitter is characterized by the number of bits used to quantize the feedback information. When the channels are reciprocal, such as in time-division duplex (TDD) systems, the outgoing and incoming channels are symmetric. In this case, training-based schemes can be used to allow the transmitter to estimate the CSI [5–8]. The overhead in obtaining the CSI can be characterized by the amount of energy used in the pilot transmission. Usually, both forward training and reverse training are needed to acquire the CSI at both the receiver and the transmitter, respectively. However, the use of reverse

training alone, which allows only the transmitter to estimate the channel using pilots sent from the receiver, was studied in [5, 7] and shown to be an energy-efficient way to achieve good symbol error rate performance at low signal-to-noise ratio (SNR) in [8].

In this paper, we consider TDD systems with reverse training only and focus on the low SNR regime. The low SNR regime is very important for wideband communications in which the overall data rate is improved by increasing the bandwidth while reducing the SNR [9]. For systems with reverse training, however, the poor quality of channel estimation at the transmitter due to low SNR can significantly reduce the data rate. To overcome this problem, we consider the flashy transmission scheme, which was proposed and studied in [10, 11] for systems with forward training, and extend it to reverse-training-based systems. The flashy transmission scheme uses a certain proportion of (but not necessarily all) the available blocks for transmission. With a long-term average energy constraint, the operating SNR during both training and data transmission increases, if a smaller fraction of blocks is used for actual transmission. Hence, the fraction of blocks for transmission, which is referred to as the *flashiness*, becomes an important design parameter to achieve good rate performance. In this work, we apply the flashy transmission scheme to systems with reverse training and obtain simple design guidelines for optimizing the flashiness in two different scenarios. In the first scenario, the designer has the most degrees of freedom and jointly optimizes the transmit power for pilot and data, as well as the flashiness to maximize a lower bound on the information capacity. In the second scenario, the pilot transmit power (at the receiver) is fixed and the designer optimizes the data transmit power (at the transmitter) as well as the flashiness.

The rest of the paper is organized as follows. In Section II, we introduce the system model and the optimization problem. A closed-form lower bound on the ergodic capacity is derived in Section III. Using the ergodic capacity lower bound as the objective function, the optimal flashiness and power allocation is studied in Section IV. In Section V, we present the numerical results. Finally, we draw the conclusion in Section VI.

Notations:  $[\cdot]^\dagger$  denotes the complex conjugate transpose operation.  $|\cdot|$  denotes the absolute value of a complex scalar.  $\|\cdot\|$  denotes the Euclidean norm of a vector.  $\mathbb{E}\{\cdot\}$  denotes the mathematical expectation.

This work was supported by the Research Council of Norway through the project 197565/V30 entitled “Theoretical Foundations of Mobile Flexible Networks - THEFONE”, and under Australian Research Council’s Discovery Projects funding scheme (project no. DP110102548). T. A. Lamahewa is also with the National ICT Australia (NICTA), Canberra Research Laboratory, 7 London Circuit, Canberra, ACT 2600, Australia. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the digital Economy and the Australian Research Council through the ICT Center of Excellence program.

## II. SYSTEM MODEL

We consider a point-to-point communication system where the transmitter is equipped with  $N_t$  antennas ( $N_t \geq 1$ ) and the receiver has a single antenna. This scenario is representative, for example, of downlink transmission in wireless local area networks. For simplicity, we refer to the transmitter as the base station (BS) and the receiver as the user terminal (UT). We assume the communication channel is flat fading and hence, the input-output relationship is given by

$$y = \mathbf{h}\mathbf{w}x + n, \quad (1)$$

where  $x$  and  $y$  are the transmit symbol from the BS and received symbol at the UT, respectively,  $\mathbf{h}$  denotes the channel gain vector of size  $1 \times N_t$  with independent and identically distributed (i.i.d.) zero-mean circular-symmetric complex Gaussian entries,  $\mathbf{w}$  represents the transmitter beamforming vector of size  $N_t \times 1$ , and  $n$  is the additive white Gaussian noise (AWGN). We normalize the variance of each element of  $\mathbf{h}$  and  $n$  to be one and hence, the power of transmitted signal also represents the receive SNR. Furthermore, we consider TDD scheme in which the channel reciprocity holds.

The time available for transmission is divided into blocks of size  $L$  symbols. We assume that the channel gains remain constant during one block and change to independent values in the next block. Although every block is available for transmission, we only use a portion of them. We refer to the percentage of blocks used for actual transmission as *flashiness*, denoted by  $\delta \in [0, 1]$ . For example, one in five blocks is used for transmission if  $\delta = 0.2$ .

Each transmission block consists of two phases, namely the reverse training phase and the data transmission phase. During reverse training, the UT sends  $L_p$  pilot symbols to the BS with pilot transmit power  $\mathcal{P}_p$ . The BS uses the linear minimum mean square error (LMMSE) estimator to obtain an estimate of the incoming channel. Using the channel reciprocity, an estimate of the outgoing channel is obtained at the BS. We denote the channel estimate as  $\hat{\mathbf{h}}$  and denote the estimation error as  $\tilde{\mathbf{h}} = \mathbf{h} - \hat{\mathbf{h}}$ . For simplicity, we normalize the receiver noise power at the BS to one. The variance of each element of  $\tilde{\mathbf{h}}$  is hence given by  $\sigma_{\tilde{h}}^2 = 1/(1 + \mathcal{P}_p L_p)$  [12]. Treating the channel estimate as the true channel, the BS designs the beamforming vector which maximizes the receive SNR at the UT, given by  $\mathbf{w} = \hat{\mathbf{h}}^\dagger / \|\hat{\mathbf{h}}\|$ . During the data transmission,  $L_d$  information symbols are transmitted by the BS through the beamforming vector and each received symbol  $y$  is given in (1). The data transmit power is given by  $\mathbb{E}\{|x|^2\} = \mathcal{P}_d$ . The total duration of one transmission block<sup>1</sup> is given by  $L_p + L_d = L$ .

We consider an energy constrained communication system, in which the flashiness  $\delta$  and the transmit power  $\mathcal{P}_p$  and  $\mathcal{P}_d$  need to be optimally designed to meet a certain objective. The

<sup>1</sup>In practice, there is a time interval between the training phase and the data transmission phase. Hence, the total duration of one block is larger than  $L_p + L_d$ . For simplicity, we ignore this time interval in this work.

average energy per block available for transmission is given by

$$\mathcal{E} = \delta(\mathcal{P}_p L_p + \mathcal{P}_d L_d). \quad (2)$$

We see that  $\mathcal{E}/\delta$  is the actual energy consumption in any block used for transmission. We also define the average power constraint as  $\mathcal{P} = \mathcal{E}/L$ .

We consider the following two optimization problems:

- Q1. For a given average energy constraint  $\mathcal{E}$ , what are the optimal flashiness and transmit power that maximize the data rate?
- Q2. For a target data rate, what are the optimal flashiness and transmit power that minimize the average energy consumption  $\mathcal{E}$ ?

We will primarily focus on the solution to Q1 in this work. As we will see later, Q2 can be readily solved with the solution to Q1.

## III. CAPACITY LOWER BOUND

In this section, we find a closed-form expression for the achievable data rate, which will be used as the objective function for the proposed optimization problem. The received signal during the data transmission phase can be rewritten as

$$\begin{aligned} y &= \hat{\mathbf{h}}\mathbf{w}x + \tilde{\mathbf{h}}\mathbf{w}x + n \\ &= \|\hat{\mathbf{h}}\|x + \tilde{\mathbf{h}}\mathbf{w}x + n. \end{aligned} \quad (3)$$

Note that the UT does not know the channel  $\mathbf{h}$  or the estimate  $\hat{\mathbf{h}}$ . However, the UT can accurately obtain the statistics of the channel estimate, e.g.,  $\mathbb{E}\{\|\hat{\mathbf{h}}\|\}$ , since the statistics change much more slowly than  $\tilde{\mathbf{h}}$  [5, 7]. Therefore, we rewrite (3) as

$$y = \mathbb{E}\{\|\hat{\mathbf{h}}\|\}x + \underbrace{(\|\hat{\mathbf{h}}\| - \mathbb{E}\{\|\hat{\mathbf{h}}\|\})x + \tilde{\mathbf{h}}\mathbf{w}x + n}_e. \quad (4)$$

where  $e$  denotes the effective noise at the UT.

Ideally one would find the expression of the information capacity, which characterizes the maximum achievable data rate. However, the exact capacity expression is generally unknown with imperfect channel estimation [13]. Therefore, we adopt the method in [13, 14] to find a lower bound on the ergodic capacity, which gives an achievable data rate for systems with reverse training only. Following [13, 14], a capacity lower bound can be obtained by approximating the effective noise to be complex Gaussian, as

$$\begin{aligned} C_{\text{LB}} &= \delta \frac{L_d}{L} \log_2 \left( 1 + \frac{\mathbb{E}\{\|\hat{\mathbf{h}}\|\}^2 \mathcal{P}_d}{\mathbb{E}\{|e|^2\}} \right) \\ &= \delta \frac{L_d}{L} \log_2 \left( 1 + \frac{\mathbb{E}\{\|\hat{\mathbf{h}}\|\}^2 \mathcal{P}_d}{1 + \mathcal{P}_d(\mathbb{E}\{\|\hat{\mathbf{h}}\|^2\} - [\mathbb{E}\{\|\hat{\mathbf{h}}\|\}]^2 + \sigma_{\tilde{h}}^2)} \right), \end{aligned} \quad (5)$$

where  $\delta$  accounts for the percentage of time used for actual transmission and  $L_d/L$  accounts for the training overhead. To obtain a closed-form expression of the capacity lower bound, we proceed as follows: Firstly, the fact that  $\mathbf{h}$  has a Gaussian distribution implies that its LMMSE estimate  $\hat{\mathbf{h}}$  also has a

$$C_{\text{LB}} = \frac{\mathcal{E}L_d}{L}(\mathcal{E}/\delta)^{-1} \log_2 \left( 1 + \frac{v\mathcal{P}_pL_p(\mathcal{E}/\delta - \mathcal{P}_pL_p)}{k\mathcal{P}_pL_p(\mathcal{E}/\delta - \mathcal{P}_pL_p) + \mathcal{E}/\delta - \mathcal{P}_pL_p + \mathcal{P}_pL_pL_d + L_d} \right). \quad (7)$$

Gaussian distribution. Then, it is easy to show that  $\|\hat{\mathbf{h}}\|^2$  has a Gamma distribution with parameter  $(N_t, 1 - \sigma_h^2)$ . Hence,

$$\mathbb{E}\{\|\hat{\mathbf{h}}\|^2\} = N_t(1 - \sigma_h^2) = \frac{N_t\mathcal{P}_pL_p}{1 + \mathcal{P}_pL_p}.$$

Secondly, we compute  $\mathbb{E}\{\|\hat{\mathbf{h}}\|\}$  using the distribution of  $g = \|\hat{\mathbf{h}}\|^2$  as

$$\begin{aligned} \mathbb{E}\{\|\hat{\mathbf{h}}\|\} &= \int_0^\infty g^{1/2} g^{N_t-1} \frac{\exp\{-\frac{g}{1-\sigma_h^2}\}}{(1-\sigma_h^2)^{N_t} \Gamma(N_t)} dg \\ &= \frac{1}{\sqrt{(1-\sigma_h^2)\Gamma(N_t)}} \int_0^\infty \left(\frac{g}{1-\sigma_h^2}\right)^{N_t-1/2} \exp\left\{-\frac{g}{1-\sigma_h^2}\right\} dg \\ &= \frac{\sqrt{(1-\sigma_h^2)}}{\Gamma(N_t)} \int_0^\infty t^{N_t-1/2} \exp\{-t\} dt \\ &= \sqrt{\frac{\mathcal{P}_pL_p}{1 + \mathcal{P}_pL_p} \frac{\Gamma(N_t + 1/2)}{\Gamma(N_t)}}, \end{aligned}$$

where  $t = \frac{g}{1-\sigma_h^2}$  and  $\Gamma(\cdot)$  denotes the Gamma function.

Therefore, the capacity lower bound in (5) can be simplified as

$$C_{\text{LB}} = \delta \frac{L_d}{L} \log_2 \left( 1 + \frac{v\mathcal{P}_d\mathcal{P}_pL_p}{k\mathcal{P}_d\mathcal{P}_pL_p + \mathcal{P}_d + \mathcal{P}_pL_p + 1} \right), \quad (6)$$

where  $v = \left[ \frac{\Gamma(N_t+1/2)}{\Gamma(N_t)} \right]^2$  and  $k = N_t - v$ . The capacity lower bound in (6) will be used as the figure of merit in solving the optimization problem in the next section.

#### IV. OPTIMAL FLASHINESS AND POWER ALLOCATION

In this section, we aim to solve the two optimization problems, Q1 and Q2, stated at the end of Section II. The degrees of freedom in a practical system determine the parameters that can be controlled by the designer. In the following, we consider two scenarios:

- 1) The transmit power at both the BS ( $\mathcal{P}_d$ ) and UT ( $\mathcal{P}_p$ ) can be jointly optimized.
- 2) The transmit power at the UT ( $\mathcal{P}_p$ ) is fixed and the transmit power at the BS ( $\mathcal{P}_d$ ) can be optimized.

We denote the optimum values of the pilot transmit power, data transmit power, and the flashiness as  $\mathcal{P}_p^*$ ,  $\mathcal{P}_d^*$ , and  $\delta^*$ , respectively. In the following, we focus on the solution to Q1. Using (2), we can rewrite the capacity lower bound in (6) as in (7) on the top of this page.

##### A. Joint Transmitter-Receiver Optimization

In the first scenario, the designer has the most degrees of freedom and tries to jointly optimize the training power  $\mathcal{P}_p$

and the data transmit power  $\mathcal{P}_d$ , as well as the flashiness  $\delta$ . Here, we write the optimization problem as

$$\begin{aligned} &\max_{\delta, \mathcal{P}_p, \mathcal{P}_d} C_{\text{LB}}, \\ &\text{subject to} \quad \mathcal{E} = \delta(\mathcal{P}_pL_p + \mathcal{P}_dL_d), \end{aligned}$$

where the constraint is on the average energy consumption given in (2). The maximization over  $\delta$  can be viewed as the outer problem and the maximization over  $\mathcal{P}_p$  and  $\mathcal{P}_d$  can be viewed as the inner problem.

For any given  $\delta$ , the optimal  $\mathcal{P}_p$  that maximizes (7) can be directly solved as

$$\mathcal{P}_p^* = \frac{\sqrt{(\mathcal{E}/\delta)^2 L_d + (\mathcal{E}/\delta)L_d^2 + (\mathcal{E}/\delta)L_d + L_d^2} - \mathcal{E}/\delta - L_d}{L_p(L_d - 1)}. \quad (8)$$

It is easy to check that  $\mathcal{P}_p^* > 0$  as long as  $L_d > 1$  and hence, the solution in (8) is always valid for systems with  $L_d > 1$ . The optimal  $\mathcal{P}_d$  for a given  $\delta$  can be obtained using (2) and (8). Furthermore, we see that  $\mathcal{P}_p^*$  in (8) is independent of the constants  $v$  and  $k$ . This implies that the optimal training and data transmission power is independent of the number of transmit antennas  $N_t$  for any fixed  $\delta$ .

To solve the outer optimization problem (i.e., optimization over  $\delta$ ), we substitute  $\mathcal{P}_p^*$  from (8) into  $C_{\text{LB}}$  in (7). Then, we obtain the capacity lower bound in the form of

$$C_{\text{LB}} = \frac{\mathcal{E}L_d}{L} f(\mathcal{E}/\delta, L_p, L_d), \quad (9)$$

where  $f(\mathcal{E}/\delta, L_p, L_d)$  denotes a function of  $\mathcal{E}/\delta$ ,  $L_p$  and  $L_d$ . For any given average energy constraint  $\mathcal{E}$ , the optimal value of  $\delta$  that maximizes  $C_{\text{LB}}$  is obtained from the optimal value of  $\varepsilon$  that maximizes  $f(\varepsilon, L_p, L_d)$ , where we have defined  $\varepsilon = \mathcal{E}/\delta$ .

Since  $\delta \in [0, 1]$ , the feasible range of  $\varepsilon$  is  $[\mathcal{E}, \infty)$ . Denote the optimal value of  $\varepsilon$  that maximizes  $f(\varepsilon, L_p, L_d)$  as  $\varepsilon^*$ . One simple but important observation can be made here: For fixed values of  $L_p$  and  $L_d$ , the value of  $\varepsilon^*$  does not change with  $\mathcal{E}$ , as long as  $\mathcal{E} < \varepsilon^*$ . This is because  $\mathcal{E}$  only affects the feasible range of  $\varepsilon$ . If  $\mathcal{E} < \varepsilon^*$  for some value of  $\mathcal{E}$ , decreasing  $\mathcal{E}$  never changes the value of  $\varepsilon^*$ , whilst increasing  $\mathcal{E}$  does not change the value of  $\varepsilon^*$  unless  $\mathcal{E}$  reaches  $\varepsilon^*$ . Note that  $\mathcal{E} < \varepsilon$  when  $\delta < 1$ . Hence, we have the following result.

*Lemma 1: When the optimal flashiness satisfies  $\delta^* < 1$ , the value of  $\mathcal{E}/\delta^*$  does not change with  $\mathcal{E}$ .*

From the derived closed-form solution of  $\mathcal{P}_p^*$  (and hence  $\mathcal{P}_d^*$ ) in (8), an immediate result following *Lemma 1* is obtained:

*Corollary 1: When the optimal flashiness satisfies  $\delta^* < 1$ , the optimal training and data transmission power does not change with  $\mathcal{E}$ .*

With the results in *Lemma 1* and *Corollary 1*, a simple design guideline for finding the optimal parameters  $\delta^*$ ,  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$  is derived:

- 
- Step 1. Choose a sufficiently low probing value of  $\mathcal{E}$ , denoted as  $\mathcal{E}_0$ . Obtain  $\mathcal{P}_p$  and  $\mathcal{P}_d$  using (8) and (2), and then numerically find  $\delta^*$  that maximizes  $C_{\text{LB}}$  in (7) for this probing energy constraint  $\mathcal{E}_0$ . If  $\delta^* < 1$ , denote  $\delta^*$  as  $\delta_0$ . Otherwise, choose a lower  $\mathcal{E}_0$  and repeat this step.
  - Step 2. Obtain  $\mathcal{P}_p$  and  $\mathcal{P}_d$  using (8) and (2) with  $\mathcal{E}_0$  and  $\delta_0$ . Denote these power values as  $\mathcal{P}_{p0}$  and  $\mathcal{P}_{d0}$ . Obtain the critical value of  $\mathcal{E}$  above which  $\delta^*$  reaches 1 as  $\varepsilon_0 = \mathcal{E}_0/\delta_0$ .
  - Step 3. Now, the closed-form solutions of  $\delta^*$ ,  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$  are available for any energy constraint  $\mathcal{E}$  as follows: If  $\mathcal{E} < \varepsilon_0$ , the optimal transmit power  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$  are given by  $\mathcal{P}_{p0}$  and  $\mathcal{P}_{d0}$ , and the optimal flashiness  $\delta^*$  is given by  $\frac{\mathcal{E}}{\mathcal{P}_{p0}L_p + \mathcal{P}_{d0}L_d}$ . Otherwise,  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$  are computed from (8) and (2) with  $\delta^* = 1$ .
- 

Note that Steps 1 and 2 in the above design guideline are the initialization stage, as they are performed without knowing the energy budget for transmission. After the initialization, the optimal values of  $\delta^*$ ,  $\mathcal{P}_p^*$ , and  $\mathcal{P}_d^*$  can be easily found for any given energy budget without the need of numerical optimization. Therefore, the optimization problem Q1 can be efficiently solved with minimal computational complexity.

To solve Q2, the designer can express the capacity lower bound (which gives an achievable data rate) as a function of  $\mathcal{E}$  by using the closed-form solution to Q1, and then numerically search for the value of  $\mathcal{E}$  that gives the target achievable data rate.

*Comments on the Optimal  $L_p$  and  $L_d$ :* If the block transmission (or frame) structure can be changed by the designer, one may also be able to choose the values of  $L_p$  and  $L_d$  that gives the best rate performance. Indeed, the optimal training and data lengths are given by  $L_p = 1$  and  $L_d = L - 1$ , respectively, provided that the optimal  $\mathcal{P}_p$  and  $\mathcal{P}_d$  are used. The proof of this result follows the proof of [15, Corollary 1]. Intuitively, it is desirable to minimize the proportion of time used for training as long as any quality of channel estimation can be achieved by changing the training power. In practice, there is a peak power constraint at the UT. If the required training power is higher than the peak power allowed, one needs to increase the training length to reduce the required training power.

## B. Transmitter-Side Optimization

In the second scenario, the designer only has control over the data transmit power at the BS  $\mathcal{P}_d$  and tries to optimize  $\mathcal{P}_d$  as well as the flashiness  $\delta$  for a given training power at the UT  $\mathcal{P}_p$ .

For a selfish designer, who only cares about the energy consumption at the BS, he/she will consider an average energy

constraint per block for the BS only, i.e.,  $\mathcal{E}_d = \delta\mathcal{P}_dL_d$ . For a given  $\mathcal{E}_d$ , it is easy to show that  $\frac{dC_{\text{LB}}}{d\delta} > 0$  by directly computing the derivative of  $C_{\text{LB}}$  in (7). Therefore, the optimal strategy for a selfish design is to use all available time for transmission, i.e.,  $\delta^* = 1$ .

When the energy consumption in both the training and data transmission is considered, the average energy constraint per block is given in (2). Hence, the optimization problem can be written as

$$\begin{aligned} \max_{\delta, \mathcal{P}_d} \quad & C_{\text{LB}}, \\ \text{subject to} \quad & \mathcal{E} = \delta(\mathcal{P}_pL_p + \mathcal{P}_dL_d), \end{aligned}$$

Note that the (optimal) values of  $\delta$  and  $\mathcal{P}_d$  are directly related by (2). Similar to the previous scenario, one can develop a simple design guideline to optimize  $\mathcal{P}_d$  and  $\delta$  with minimal computational complexity as follow.

It is clear that the optimization problem is trivial when  $\delta^* = 1$ . Similar to *Lemma 1* in the previous scenario, we have the following fact: When  $\delta^* < 1$ , the value of  $\mathcal{E}/\delta^*$  which maximizes the capacity lower bound in (7) does not change with  $\mathcal{E}$  (due to the variation in  $\delta^*$ ), for any fixed values of  $\mathcal{P}_p$ ,  $L_p$ , and  $L_d$ . When  $\delta^*$  is found, the optimal data transmit power is readily obtained using (2) as

$$\mathcal{P}_d^* = \frac{\mathcal{E}/\delta^* - \mathcal{P}_pL_p}{L_d}, \quad (10)$$

which does not change with  $\mathcal{E}$  if  $\delta^* < 1$ . With this property, a simple design guideline is derived:

- 
- Step 1. Choose a sufficiently low probing value of  $\mathcal{E}$ , denoted as  $\mathcal{E}_0$ . Numerically find  $\delta^*$  that maximizes  $C_{\text{LB}}$  in (7) for this probing energy constraint  $\mathcal{E}_0$ . If  $\delta^* < 1$ , denote  $\delta^*$  as  $\delta_0$ . Otherwise, choose a lower  $\mathcal{E}_0$  and repeat this step.
  - Step 2. Obtain  $\mathcal{P}_d$  using (10) with  $\mathcal{E}_0$  and  $\delta_0$ , and denote this power value as  $\mathcal{P}_{d0}$ . Obtain the critical value of  $\mathcal{E}$  above which  $\delta^*$  reaches 1 as  $\varepsilon_0 = \mathcal{E}_0/\delta_0$ .
  - Step 3. Now, the closed-form solutions of  $\delta^*$  and  $\mathcal{P}_d^*$  are available for any given energy constraint  $\mathcal{E}$  as follows: If  $\mathcal{E} < \varepsilon_0$ , the optimal data transmit power  $\mathcal{P}_d^*$  is given by  $\mathcal{P}_{d0}$ , and the optimal flashiness  $\delta^*$  is given by  $\frac{\mathcal{E}}{\mathcal{P}_pL_p + \mathcal{P}_{d0}L_d}$ . Otherwise,  $\mathcal{P}_d^*$  is computed from (10) with  $\delta^* = 1$ .
- 

Again we see that Steps 1 and 2 in the above design guideline are the initialization stage, as they are performed without knowing the energy budget for transmission. After the initialization, the optimal values of  $\delta^*$  and  $\mathcal{P}_d^*$  can be easily found for any given energy budget without the need of numerical optimization. Therefore, Q1 can be efficiently solved with minimal computational complexity.

To solve Q2, the designer can express the capacity lower bound (which gives an achievable data rate) as a function of  $\mathcal{E}$  by using the closed-form solution to Q1, and then numerically search for the value of  $\mathcal{E}$  that gives the target achievable data rate.

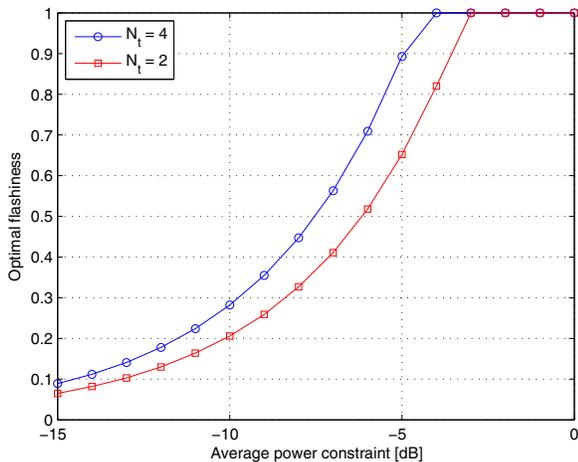


Fig. 1. The optimal flashiness  $\delta^*$  versus average power constraint  $\mathcal{P}$  for different numbers of antennas. The block length is  $L = 20$  and the training length is  $L_p = 1$ .

## V. NUMERICAL RESULTS

In this section, we present numerical results to illustrate the rate improvement from using the optimal flashy transmission in the low SNR regime. The optimal flashiness and transmit power is found by following the design guidelines described in the previous section. One can also use numerical search to obtain the same results. We first show the results for the case of joint transmitter-receiver optimization and then look at the case of transmitter-side optimization.

Fig. 1 shows the optimal flashiness  $\delta^*$  versus average power constraint  $\mathcal{P} = \mathcal{E}/L$ . The value of  $\delta^*$  stays at 1 when  $\mathcal{P}$  is above some critical value. Once  $\mathcal{P}$  drops below this critical value,  $\delta^*$  decreases at the rate such that  $\mathcal{P}/\delta^*$  is a constant (in linear scale), and this constant value equals the critical value of  $\mathcal{P}$ . We see that the optimal flashiness can reach as low as 0.2 at  $\mathcal{P} = -10$  dB for systems with  $N_t = 2$ . Comparing the two curves with different  $N_t$  in Fig. 1, we observe that  $\delta^*$  is higher in a system with a larger number of transmit antennas when  $\delta^* < 1$ . This trend is also observed for other values of  $N_t$  and the plots are omitted for brevity.

Fig. 2 shows the optimal training and data transmit power,  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$ , versus average power constraint  $\mathcal{P}$ . We see that the optimal training and data power is independent of the number of transmit antennas for the range of  $\mathcal{P} > -3$  dB, since the optimal flashiness is the same ( $\delta^* = 1$ ) for both  $N_t = 2$  and  $N_t = 4$ . Furthermore, the values of  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$  stay constant when  $\mathcal{P}$  is below its critical value, i.e., when  $\delta^* < 1$ .

Fig. 3 shows the capacity lower bound  $C_{LB}$  versus average power constraint  $\mathcal{P}$ . Here, we compare  $C_{LB}$  achieved by optimal flashy transmission with  $C_{LB}$  achieved by non-flashy transmission, i.e.,  $\delta = 1$ . The training and data transmit power is optimized in both transmission schemes. It is clear from Fig. 3 that a significant rate improvement can be achieved by using flashy transmission at low SNR. For example, this

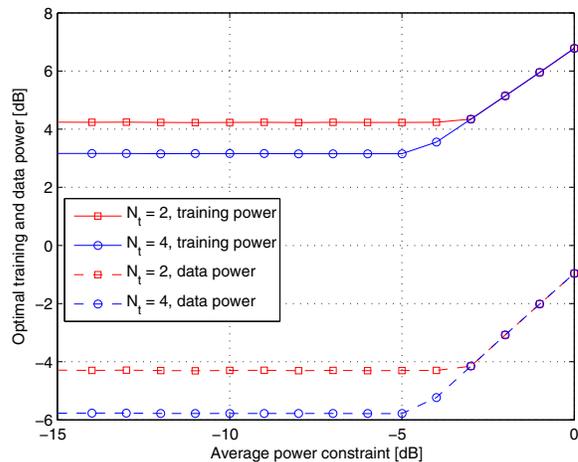


Fig. 2. The optimal training and data transmit power,  $\mathcal{P}_p^*$  and  $\mathcal{P}_d^*$ , versus average power constraint  $\mathcal{P}$  for different numbers of antennas. The block length is  $L = 20$  and the training length is  $L_p = 1$ .

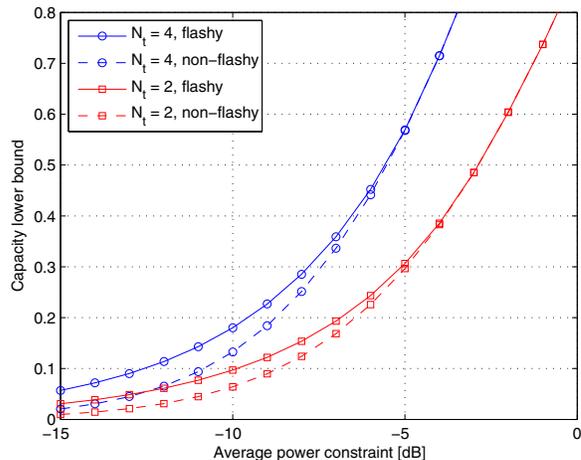


Fig. 3. The capacity lower bound  $C_{LB}$  versus average power constraint  $\mathcal{P}$  for different numbers of antennas. The solid lines indicate the capacity lower bound achieved using the optimal flashy transmission, while the dashed lines indicate the capacity lower bound achieved using non-flashy transmission, i.e.,  $\delta = 1$ . The block length is  $L = 20$  and the training length is  $L_p = 1$ .

rate improvement is 36% and 51% for  $N_t = 4$  and  $N_t = 2$ , respectively, at  $\mathcal{P} = -10$  dB.

Next we present numerical results for the case of transmitter-side optimization. Fig. 4 shows the optimal flashiness  $\delta^*$  versus average power constraint  $\mathcal{P}$  for different training power  $\mathcal{P}_p$ . Similar to the previous case, we see that  $\delta^*$  equals 1 when  $\mathcal{P}$  is above some critical value and it reduces as  $\mathcal{P}$  drops below the critical value. The rate at which  $\delta^*$  decreases with  $\mathcal{P}$  is constant (in linear scale) such that  $\mathcal{P}/\delta^*$  always equals the critical value of  $\mathcal{P}$ .

Fig. 5 shows the capacity lower bound  $C_{LB}$  versus average power constraint  $\mathcal{P}$ . Note that  $C_{LB}$  for non-flashy transmission only makes sense when  $\mathcal{E} = \mathcal{P}L > \mathcal{P}_p L_p$ . Again, we see a

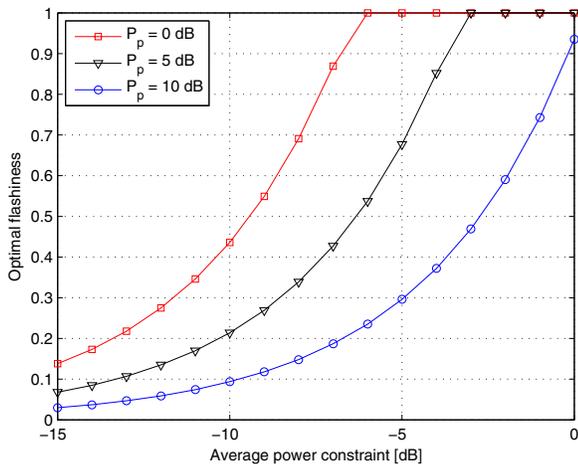


Fig. 4. The optimal flashiness  $\delta^*$  versus average power constraint  $\mathcal{P}$  for different training power. The block length is  $L = 20$ , the training length is  $L_p = 1$  and the number of transmit antennas is  $N_t = 4$ .

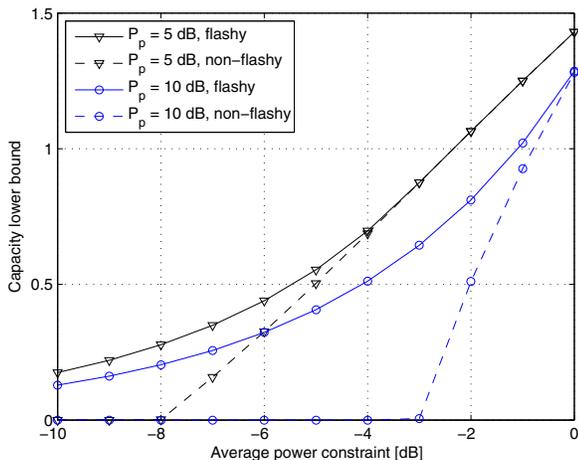


Fig. 5. The capacity lower bound  $C_{LB}$  versus average power constraint  $\mathcal{P}$  for different training power. The solid lines indicate the capacity lower bound achieved using the optimal flashiness transmission, while the dashed lines indicate the capacity lower bound achieved using non-flashiness transmission, i.e.,  $\delta = 1$ . The block length is  $L = 20$ , the training length is  $L_p = 1$  and the number of transmit antennas is  $N_t = 4$ .

significant rate improvement from flashiness transmission in the low SNR regime.

## VI. CONCLUSION

In this paper, we considered TDD systems with reverse training only and studied the flashiness transmission in the low SNR regime. We derived a closed-form lower bound on the ergodic capacity, which gives an achievable data rate. We provided simple design guidelines for flashiness transmission and demonstrated a significant data rate improvement from using the optimized parameters. The results obtained in this work provide useful insights into the design of wideband communication systems in the low SNR regime.

## REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New Jersey: Wiley, 2006.
- [3] D. J. Love, R. W. Heath Jr., W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Commun. Mag.*, pp. 54–59, Oct. 2004.
- [4] D. J. Love, "Duplex distortion models for limited feedback MIMO communications," *IEEE Trans. Signal Processing*, vol. 54, no. 2, pp. 766–774, Feb. 2006.
- [5] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC)*, Pacific Grove, CA, Oct. 2006, pp. 359–363.
- [6] C. Steger and A. Sabharwal, "Single-input two-way SIMO channel: Diversity-multiplexing tradeoff with two-way training," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4877–4885, Dec. 2008.
- [7] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Scheduling and pre-conditioning in multi-user MIMO TDD systems," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Beijing, China, May 2008, pp. 4100–4105.
- [8] X. Zhou, T. A. Lamacchia, P. Sadeghi, and S. Durrani, "Two-way training: Optimal power allocation for pilot and data transmission," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 564–569, Feb. 2010.
- [9] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
- [10] L. Zheng, D. N. C. Tse, and M. Médard, "Channel coherence in the low-SNR regime," *IEEE Trans. Inform. Theory*, vol. 53, no. 3, pp. 976–997, Mar. 2007.
- [11] X. Zhou, P. Sadeghi, and T. A. Lamacchia, "Optimizing training-based MIMO systems: How much time is needed for actual transmission?" in *Proc. IEEE Veh. Tech. Conf. (VTC-Spring)*, Taipei, Taiwan, May 2010, pp. 1–5.
- [12] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- [13] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.
- [14] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [15] X. Zhou, P. Sadeghi, T. A. Lamacchia, and S. Durrani, "Design guidelines for training-based MIMO systems with feedback," *IEEE Trans. Signal Processing*, vol. 57, no. 10, pp. 4014–4026, Oct. 2009.