This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2010 proceedings

A Novel Energy-Efficient Training Method for Receive Antenna Selection

Vinod Kristem
Neelesh B. Mehta, *Senior Member, IEEE*
Andreas F. Molisch, *Fellow, IEEE*

***Abstract*—Receive antenna selection (AS), in which only a subset of antennas receive simultaneously at any time, requires the transmitter to send pilots multiple times so that the receiver can acquire channel state of all antennas and select the best subset. In conventional AS training, sensitivity of coherent *reception* to channel estimation errors forces the transmitter to boost the energy allocated to all pilots to ensure accurate channel estimates. Energy for pilots received by unselected antennas is mostly wasted, especially since the *selection* process is robust to estimation errors. In this paper, we propose a novel training method uniquely tailored for AS. Using one extra pilot symbol, accurate channel estimates get generated for the antenna subset that actually receives data. Consequently, the transmitter can selectively boost the energy allocated to the extra pilot. Other pilots, which now primarily help in subset selection, get allocated less energy. We derive closed-form expressions for the proposed scheme's symbol error probability, and optimize the energy allocated to pilot and data symbols. We show that the optimal solution achieves full diversity; it is provably unique and strictly better than the conventional model.**

I. Introduction

Receive antenna selection (AS) provides a low hardware complexity solution for exploiting the spatial diversity benefits of multiple receive antennas [1], [2]. In AS, the receiver dynamically selects a subset of the antennas with the 'best' instantaneous channel conditions to the transmitter, and processes signals through them. This enables the receiver to employ fewer of the expensive radio frequency (RF) chains. Consequently, AS has been adopted in next generation wireless systems such as IEEE 802.11n [3].

In practice, the channel state information (CSI) is acquired using a pilot-based training scheme [4]. The low hardware complexity of AS imposes unique constraints on how its training gets done. *Specifically, only L antennas can be estimated at any time with L RF chains*. Therefore, in receive AS, the transmitter needs to transmit the pilots multiple times so that the receiver can estimate the channels of all the available antennas and choose the antennas with the best channels. For example, with 4 receive antennas and 2 RF chains, at least 2 pilot symbols are needed – the first one to estimate antennas 1 and 2, and the second one to estimate antennas 3 and 4. In general, with N antennas and L RF chains, the minimum number of training symbols required is $\lceil N/L \rceil$, where $\lceil . \rceil$ denotes the ceil function.

V. Kristem is with Beceem Communications, Bangalore India. He was with the Dept. of Electrical Communication Eng. (ECE) at the Indian Institute of Science (IISc), Bangalore, India during the course of this work. N. B. Mehta is with the ECE Dept. at IISc. A. F. Molisch is with the Dept. of Electrical Eng. at the University of Southern California, Los Angeles, CA, USA.
Emails: vinod.kristem@gmail.com, nbmehta@ece.iisc.ernet.in, andreas.molisch@ieee.org

In AS, estimation errors may cause a suboptimal antenna subset to get selected and will also degrade the performance of coherent demodulation. Since the transmitter does not know a priori which antennas will be selected, it needs to uniformly boost the energy of all the $\lceil \frac{N}{L} \rceil$ pilots to ensure good coherent demodulation performance. This process is energy-inefficient because it obtains highly accurate estimates of unselected antennas. Under a total energy constraint, it even draws energy away from data symbols. Altogether, despite AS being a widely studied technique, the energy-efficiency of its training method has received relatively limited attention. While AS with channel estimation errors was considered in [5]–[7], the optimal pilot and data energy allocation was not considered.

In this paper, we propose a novel training method for AS that significantly improves the energy-efficiency of AS. In the proposed method, the transmitter sends an extra pilot symbol after the first $\lceil N/L \rceil$ pilots, *i.e.*, after the subset that will do the reception has been selected. This one pilot enables the receiver to refine its channel estimates of the selected antenna subset. *Consequently, this last pilot enables the transmitter to exploit the robustness of AS to selection errors and reduce the total energy allocated to the first $\lceil N/L \rceil$ pilots that are used for selection. Only the energy of this last pilot needs to be boosted to get accurate channel estimates for antennas actually used to receive data. The energy saved can be transferred to the data symbols to reduce their symbol error probability.* The method, thus, enables the transmitter to optimally balance the selection errors and inaccuracy of channel estimates used for coherent reception, and best utilize the total available energy. Such a power allocation across both pilots and data was studied in [8], [9] but not for AS.

Under a total energy constraint, we derive closed-form expressions for the fading-averaged SEP of MPSK constellation of the proposed method as a function of the fractions of energy allocated to the various pilot and data symbols. The analysis accounts for imperfect estimation in both the selection and data transmission phases, and enables the fraction of energy allocated to pilots and data to be optimized. To provide a fair comparison, we also optimize the energies allocated to pilots and data in the conventional AS training method, and develop new closed-form solutions for its optimal energy allocation. The analysis, which is verified using extensive Monte Carlo simulations, shows that a coding gain as large as 5 dB over conventional AS training methods can be achieved,

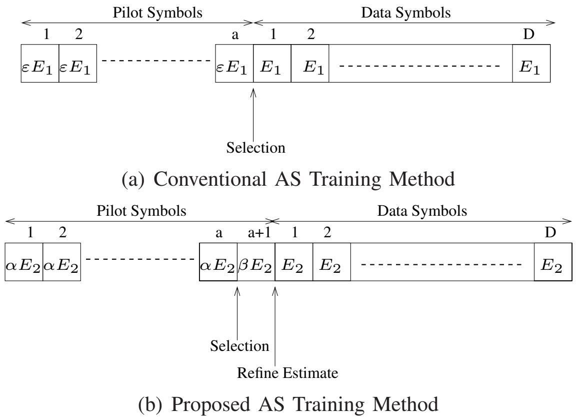

Fig. 1. Conventional and proposed training methods for receive AS.

depending on the constellation size. Our asymptotic analysis provides more insights into the proposed method, and shows that it is strictly better than the conventional method and retains full diversity order. It also leads to an approximate, but very accurate, closed-form expression for the optimal energy allocation.

The outline of this paper is as follows. The AS training models are described in Sec. II. They are analyzed in Sec. III, and are followed by results in Sec. IV. We conclude in Sec. V.

II. MODEL

Consider a system with one transmit, N receive antennas, and L receive RF chains. Let h_k denote the channel gain of the frequency-flat, time-invariant block-fading channel between the transmitter and the k^{th} receive antenna. It is modeled as a circularly symmetric complex Gaussian random variable (RV) with unit variance. Therefore, the amplitude of the channel gain, $|h_k|$, is a Rayleigh RV. The channel gains for different receive antennas are assumed to be independent and identically distributed (i. i. d.), which is the case when the receive antennas are spaced sufficiently apart [10].

We first describe the conventional AS training method, and then the proposed method.

A. Conventional AS Training Method

To enable the receiver to estimate the channel gains of all the N links, the transmitter sequentially transmits $a = \lceil \frac{N}{L} \rceil$ pilot symbols [5] each with energy εE_1, sequentially, as shown in Figure 1(a), where E_1 is the energy allocated to a data symbol and $\varepsilon \geq 0$ is the energy scaling factor for pilots. Each pilot symbol is received by at most L receive antennas.

The signal received by the k^{th} receive antenna is given by

$$r_k = \sqrt{\varepsilon E_1} p h_k + n_k, \tag{1}$$

where p is the (complex) pilot symbol with $|p| = 1$, n_k is a circular symmetric complex Gaussian RV with variance N_0 that is independent across k. The minimum mean square error (MMSE) channel estimate for the k^{th} antenna is [11]

$$\hat{h}_k = \frac{\sqrt{\varepsilon E_1} p^*}{\varepsilon E_1 + N_0} r_k. \tag{2}$$

Since the antenna elements are independent of each other, the receiver selects the L antennas with the highest estimated channel gains. Let $\hat{\Omega}_L$ denote the selected subset of antennas. Note that $\hat{\Omega}_L$ depends on $\left\{\hat{h}_k\right\}_{k=1}^{N}$.

The pilot symbols are followed by D data symbols, each transmitted with energy E_1. All the D data symbols are received by the same antenna subset $\hat{\Omega}_L$. The constraint on total energy, E_T, takes the form

$$(a\varepsilon + D)E_1 = E_T. \tag{3}$$

Let $\gamma \triangleq \frac{E_T}{N_0}$ denote the total energy normalized with respect to noise variance. Therefore, $\frac{E_1}{N_0} = \frac{\gamma}{a\varepsilon + D}$. Notice that the choice of ε affects the energy allocated, E_1, to each data symbol.

B. Proposed AS Training Method

We now describe the proposed AS training method, which is shown in Figure 1(b). As before, the transmitter now first sequentially transmits $a = \lceil \frac{N}{L} \rceil$ pilot symbols so that all the N channels can be estimated. But, each pilot symbol is now transmitted with energy αE_2, where E_2 is the energy per data symbol and α is now the energy scaling factor for these pilots.

The pilot symbol received by the k^{th} receive antenna is $r_k = \sqrt{\alpha E_2} p h_k + n_k$, where n_k, as before, is a circular symmetric complex Gaussian RV with variance N_0. As in (2), the MMSE channel estimate for the k^{th} antenna is given by

$$\hat{h}_k = \frac{\sqrt{\alpha E_2} p^*}{\alpha E_2 + N_0} r_k = a_1 h_k + e_k, \tag{4}$$

where $a_1 \triangleq \frac{\alpha E_2}{\alpha E_2 + N_0}$, and the zero-mean Gaussian noise term $e_k \triangleq \frac{n_k p^* \sqrt{\alpha E_2}}{\alpha E_2 + N_0}$ with variance $\sigma_e^2 = \frac{\alpha E_2 N_0}{(\alpha E_2 + N_0)^2}$.

The L antennas with the largest channel estimate magnitudes are selected. As before, $\hat{\Omega}_L$ denote the selected antenna subset, which depends on $\left\{\hat{h}_k\right\}_{k=1}^{N}$.

Extra Pilot and Refined Estimates: The key difference in the proposed method is that an extra pilot symbol is transmitted with energy βE_2, and is received by the *selected L antennas*. This additional pilot helps in refining the channel estimates of the selected L antennas as explained below. Note that, in general, $\beta \neq \alpha$. The received signal for the extra pilot is

$$r'_k = \sqrt{\beta E_2} p h_k + n'_k, \quad k \in \hat{\Omega}_L. \tag{5}$$

The channel estimate of a selected antenna $k \in \hat{\Omega}_L$ can be refined using the observation r'_k in addition to r_k. The refined MMSE estimate, $\hat{\hat{h}}_k$, that uses both r_k and r'_k equals [11]

$$\hat{\hat{h}}_k = \frac{\sqrt{\alpha E_2} p^* r_k + \sqrt{\beta E_2} p^* r'_k}{(\alpha + \beta) E_2 + N_0} = a_2 h_k + e'_k, \tag{6}$$

where $a_2 \triangleq \frac{(\alpha+\beta)E_2}{(\alpha+\beta)E_2 + N_0}$ and the zero-mean Gaussian noise term $e'_k \triangleq \frac{\sqrt{\alpha E_2} p^* n_k + \sqrt{\beta E_2} p^* n'_k}{(\alpha+\beta)E_2 + N_0}$ has a variance of $\sigma_{e'}^2 = \frac{(\alpha+\beta)E_2 N_0}{((\alpha+\beta)E_2 + N_0)^2}$. Note that e'_k and e_k are correlated.

Data Reception: The pilot symbols are followed by D data symbols, each transmitted with energy E_2. The data symbols are drawn from the MPSK constellation with equal probability.

They are all received by the antenna subset $\hat{\Omega}_L$. The received signal for a data symbol (symbol index not shown) is

$$y_k = h_k s + n_k'', \quad k \in \hat{\Omega}_L. \tag{7}$$

The ML decision variable, $\mathcal{D}$, for data decoding is $\mathcal{D} = \sum_{k\in\hat{\Omega}_L} \hat{\hat{h}}_k^* y_k$. The total energy constraint now becomes

$$E_T = (a\alpha + \beta + D)\, E_2. \tag{8}$$

Let $\gamma \triangleq \frac{E_T}{N_0}$. Hence, $\frac{E_2}{N_0} = \frac{\gamma}{a\alpha+\beta+D}$. Now, both α and β affect the energy E_2 allocated to a data symbol. The conventional method is a special case of the proposed method when $\beta = 0$.

A corresponding scheme can also be developed for multiple transmit antennas; it is beyond the scope of this paper.

III. SEP Analysis and Optimization

We now analyze the fading-averaged SEP for the MPSK for receive AS with imperfect CSI for the conventional and proposed AS training methods, and optimize their parameters to minimize the SEP. A corresponding analysis for MQAM can also be developed [12].

Henceforth, $\mathbf{E}[A]$ and $\mathbf{var}[A]$ shall denote the expectation and variance, respectively, of RV A. Similarly, $\mathbf{E}[A|B]$ and $\mathbf{var}[A|B]$ will denote the conditional expectation and variance given B. x^* denotes the complex conjugate of x.

A. Conventional Method Optimization

We discuss the conventional method only briefly given the analysis in [5]. The main contribution of this section lies in determining the optimal value of ε that minimizes the SEP subjected to total energy constraint. This will provide a fair benchmark for our new method, analyzed next.

In terms of the notation of the current paper, the SEP expression for MPSK, denoted by $P_{\text{MPSK}}^{\varepsilon}(\gamma)$, is [5]:

$$P_{\text{MPSK}}^{\varepsilon}(\gamma) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \left(\frac{\sin^2\theta}{\sin^2\theta + c}\right)^L \prod_{n=L+1}^{N} \frac{\sin^2\theta}{\sin^2\theta + \frac{cL}{n}}\, d\theta, \tag{9}$$

where $c \triangleq \frac{\varepsilon\gamma^2 \sin^2\left(\frac{\pi}{M}\right)}{(a\varepsilon+D)((\varepsilon+1)\gamma+a\varepsilon+D)}$. It follows that the optimal value of ε, denoted by $\varepsilon^\star(\gamma)$, that minimizes $P_{\text{MPSK}}^{\varepsilon}(\gamma)$ is

$$\varepsilon^\star(\gamma) = \sqrt{\frac{D(\gamma + D)}{a(\gamma + a)}}. \tag{10}$$

When $\gamma \to \infty$, we have $\lim_{\gamma\to\infty} \varepsilon^\star(\gamma) \triangleq \varepsilon_\infty^\star = \sqrt{\frac{D}{a}}$.

B. SEP Analysis and Optimization of Proposed Method

The following result about $\mathcal{D}$, which follow from standard results on conditional Gaussians, shall be useful in deriving the SEP of the proposed method.

Lemma *1:* Conditioned on $\left\{\hat{h}_l, \hat{\hat{h}}_l\right\}_{l\in\hat{\Omega}_L}$ and s, $\mathcal{D}$ is a complex Gaussian RV with conditional mean, $\mu_{\mathcal{D}}$, and variance, $\sigma_{\mathcal{D}}^2$, given by

$$\mu_{\mathcal{D}} \triangleq \mathbf{E}\left[\mathcal{D} \,\Big|\, \left\{\hat{h}_l, \hat{\hat{h}}_l\right\}_{l\in\hat{\Omega}_L}, s\right] = s \sum_{k\in\hat{\Omega}_L} \left|\hat{\hat{h}}_k\right|^2,$$

$$\sigma_{\mathcal{D}}^2 \triangleq \mathbf{var}\left[\mathcal{D} \,\Big|\, \left\{\hat{h}_l, \hat{\hat{h}}_l\right\}_{l\in\hat{\Omega}_L}, s\right] = ((1-a_2)E_2 + N_0) \sum_{k\in\hat{\Omega}_L} \left|\hat{\hat{h}}_k\right|^2,$$

where $a_2 = \frac{(\alpha+\beta)E_2}{(\alpha+\beta)E_2+N_0}$.

The expression for SEP given α and β now follows.

Theorem *1:* With noisy channel estimates, the fading-averaged SEP of MPSK is:

$$P_{\text{MPSK}}^{\alpha-\beta}(\gamma) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \left(\frac{\sin^2\theta}{a_2 b + \sin^2\theta}\right)^L \times \prod_{n=L+1}^{N} \left(1 + \frac{a_1 bL/n}{(a_2 - a_1)b + \sin^2\theta}\right)^{-1} d\theta, \tag{11}$$

where $a_1 = \frac{\alpha\gamma}{(a+\gamma)\alpha+\beta+D}$, $a_2 = \frac{(\alpha+\beta)\gamma}{(a+\gamma)\alpha+(\gamma+1)\beta+D}$, and $b = \frac{\gamma\sin^2\left(\frac{\pi}{M}\right)}{(1-a_2)\gamma+a\alpha+\beta+D}$.

Proof: The proof is relegated to the Appendix. ■

Closed-form expressions can be derived from (11) using [13, (5A.42), (5A.56)]. However, they are quite lengthy, and not shown here. The optimal values of α and β can now be found numerically using gradient search. However, the asymptotic energy regime considered below provides considerable insight.

C. Asymptotics ($\frac{E_T}{N_0} = \gamma \to \infty$)

The SEP in (11) can be upper bounded by replacing $\sin^2\theta$ with 1 [13]. As $\gamma \to \infty$, the bound becomes

$$P_{\text{MPSK}}^{\alpha-\beta,\infty}(\gamma) \le \gamma^{-N} \frac{(L+1)\,(L+2)\cdots(N)}{L^{N-L}(M/(M-1))} \times \left(\frac{(a\alpha+\beta+D)(\alpha+\beta+1)}{(\alpha+\beta)\sin^2\left(\frac{\pi}{M}\right)}\right)^N \left(1 + \frac{\beta\sin^2\left(\frac{\pi}{M}\right)}{\alpha(\alpha+\beta+1)}\right)^{N-L}. \tag{12}$$

The above expression shows that the diversity order is N for any $\alpha > 0$ and $\beta \ge 0$. Let $\alpha_\infty^\star$ and $\beta_\infty^\star$ denote the optimal values of α and β, respectively, that minimize the SEP upper bound. $\alpha_\infty^\star$ and $\beta_\infty^\star$ satisfy the following properties:

Theorem *2:* 1) $\alpha_\infty^\star$ and $\beta_\infty^\star$ are related by

$$\beta_\infty^\star = -\alpha_\infty^\star + \sqrt{D - (\alpha_\infty^\star)^2\,(a-1)}. \tag{13}$$

2) $\alpha_\infty^\star$ is unique and lies in the following range:

$$0 < \alpha_\infty^\star \le \sqrt{\frac{D(N-L)}{N(a-1)}} \sin\left(\frac{\pi}{M}\right) \le \varepsilon_\infty^\star = \sqrt{\frac{D}{a}}. \tag{14}$$

The equalities hold only when $M = 2$ (BPSK) and L divides N.

Proof: The proof uses fundamental results of functional calculus and is detailed in [12]. ■

For large constellation sizes ($M \to \infty$), the optimal values $\alpha_\infty^\star$ and $\beta_\infty^\star$ can, in fact, be determined in closed-form:

Theorem *3:*

$$\lim_{M\to\infty} \frac{\alpha_\infty^\star}{\sin\left(\frac{\pi}{M}\right)} = \sqrt{\frac{D(N-L)}{N(a-1)}}. \tag{15}$$

Proof: The proof is omitted due to space constraints, and is given in [12]. ■

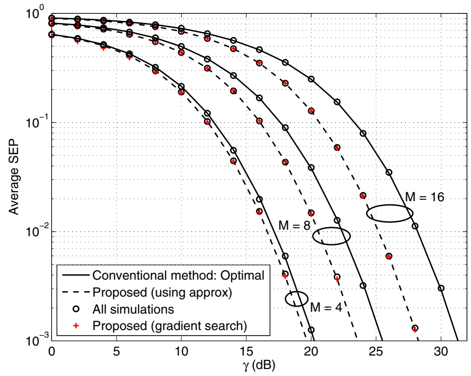

Fig. 2. Comparison of proposed and conventional AS training methods: Effect of constellation size on average MPSK SEP ($N = 6$, $L = 1$, $D = 1$).

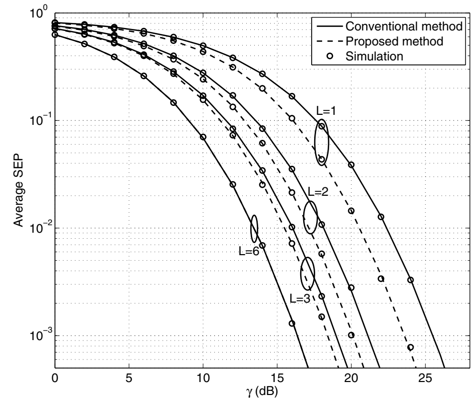

Fig. 3. Comparison of proposed and conventional AS training methods: Effect of number of RF chains on average SEP (8PSK, $N = 6$, and $D = 1$).

It follows from (15) and (13), that, for large M,

$$\alpha_\infty^\star \approx \sqrt{\frac{D(N-L)}{N(a-1)}}\sin\left(\frac{\pi}{M}\right),$$

$$\beta_\infty^\star \approx -\sqrt{\frac{D(N-L)}{N(a-1)}}\sin\left(\frac{\pi}{M}\right)+\sqrt{D-D\left(1-\frac{L}{N}\right)\sin^2\left(\frac{\pi}{M}\right)}. \quad (16)$$

These asymptotic expressions provide several insights:

1) Sensitivity to System Parameters: While $\varepsilon_\infty^\star$ in the conventional method depends only on D and the number of pilot symbols, a; in the proposed scheme, $\alpha_\infty^\star$ depends on *all* system parameters N, L, D, M, and a.

2) Energy Allocation: Using (3) and (8), it can be shown that $E_1 \leq E_2$ and $\alpha_\infty^\star + \beta_\infty^\star \geq \varepsilon_\infty^\star$ [12]. *Thus, the proposed method not only allocates more energy to each data symbol, it also ensures that the quality of estimates, which is related to $\alpha_\infty^\star + \beta_\infty^\star$ (refer (6)) is better than that of the conventional method, which is related to $\varepsilon_\infty^\star$ (refer (2)).*

IV. Simulation Results And Comparisons

We now plot the analytical results derived in Section III and validate them with Monte Carlo simulations. We also study the effect of all the system parameters, N, L, D, and M, on the optimal SEPs of the conventional and proposed methods.

Figures 2 plots the SEP as a function of the normalized total energy, γ, for MPSK, for a given L and N. While the SEP of both methods expectedly increases with M, the performance gain of the proposed method increases with M. One reason is that $\varepsilon_\infty^\star$ of the conventional method is insensitive to M. Notice the excellent match between the analytical and simulation results. The optimal α and β for the proposed method are found using a gradient search over the SEP formulae of Theorem 1 and using the closed-form asymptotic approximation from Theorem 3. The approximation is accurate even at low values of γ.

Figure 3 shows the effect of the antenna subset size, L, on the SEP. As L increases, the number of pilots that need to be transmitted decreases. Hence, relatively less energy is spent on training. Consequently, for both methods, the SEP improves for all γ. The proposed method does better for all $L < N$.

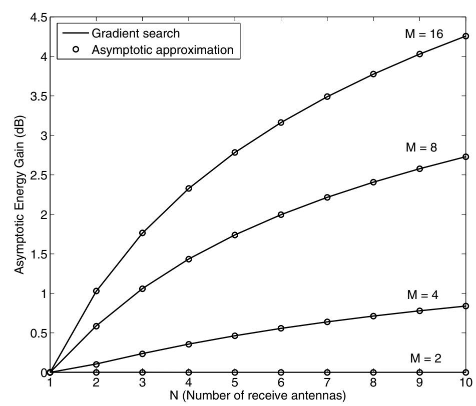

Fig. 4. Effect of number of receive antennas on asymptotic energy gain of proposed method over conventional method ($D = 1$ and $L = 1$).

Figure 4 studies the joint impact of N and M on the asymptotic energy gain, which measures the savings in total energy achieved by the proposed method over the conventional optimized method for the same SEP. The energy gain increases with M and N. The latter occurs because as N increases: (i) relatively more energy is allocated to data in the proposed training method compared to conventional method ($i.e.$, E_2/E_1 increases with N), (ii) the relative quality of estimates ($i.e.$, $(\alpha_\infty^\star + \beta_\infty^\star)E_2/(\varepsilon_\infty^\star E_1)$) increases with N.

Figure 5 studies the impact of D on the asymptotic energy gain. We can see that the energy gain decreases as D increases because the relative energy allocated to data symbols decreases since the ratio E_2/E_1 decreases with D. Notice that the energy gain exceeds 1 dB even when $D \leq 25$ for $M \geq 8$. The above two figures again show that the closed-form approximations are extremely accurate.

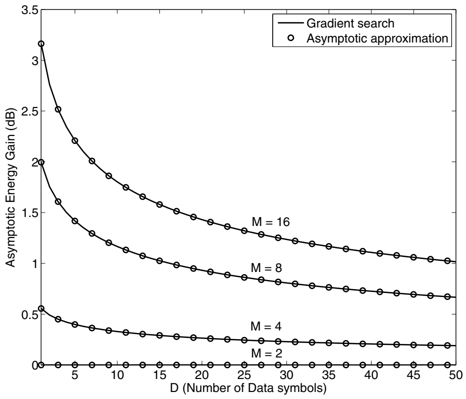

Fig. 5. Effect of number of data symbols, D, and constellation size, M, on the asymptotic energy gain of proposed method over conventional method ($N = 6$ and $L = 1$).

V. Conclusions

We proposed a new training method tailored for antenna selection that exploited the observation that subset selection is robust to channel estimation errors, while coherent demodulation is not. The method assigned less energy to 'selection pilots', which are used to select the antenna subset. It instead allocated more energy to an extra last pilot to refine the channel estimates for the antenna subset actually used for data reception. We derived closed-form equations for the SEP of MPSK, and showed analytically that this approach always a provides lower SEP than the 'conventional' method that uses the same pilots for selection and channel estimation.

The analysis gave considerable insight about the behavior of the training methods. While the conventional method's optimal energy allocation depended only on D and the number of pilot symbols, a, the proposed method's optimal parameters depended on N, L, D, and the constellation size, M. A highly accurate closed-form expression for the optimal parameters was also derived. The energy gain increased as the constellation size or the number of pilots transmitted increased.

It is worth noting that the proposed scheme is uniquely tailored to AS. It effectively exploits the fact that selection is more robust to estimation errors than coherent demodulation [14]. For example, such an insertion of an extra pilot would yield no benefit under a total energy constraint for a system with as many RF chains as receive antennas. On the other hand, in AS, it leads to significant benefits.

Appendix

We outline the key steps of the proof of Theorem 1 below. The SEP for MPSK when $\mathcal{D}$ is a Gaussian RV [13] is

$$P_{\text{MPSK}}\left(\left\{\hat{h}_l, \hat{\hat{h}}_l\right\}_{l\in\hat{\Omega}_L}\right) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \exp\left(\frac{-\left|\mu_{\mathcal{D}}\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\sigma_{\mathcal{D}}^2 \sin^2\theta}\right) d\theta.$$

Note that $\hat{\Omega}_L$ depends on $\hat{h}_l$. Using Lemma 1 and averaging over $\left\{\hat{\hat{h}}_l\right\}_{l\in\hat{\Omega}_L}$, we get

$$P_{\text{MPSK}}\left(\left\{\hat{h}_l\right\}_{l\in\hat{\Omega}_L}\right) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \mathcal{M}_{Y \,|\, \left\{\hat{h}_l\right\}_{l\in\hat{\Omega}_L}}\left(\frac{-b}{\sin^2\theta}\right) d\theta, \quad (17)$$

where $b = \frac{E_2 \sin^2\left(\frac{\pi}{M}\right)}{(1-a_2)E_2 + N_0}$ and $Y \triangleq \sum_{k\in\hat{\Omega}_L} \left|\hat{\hat{h}}_k\right|^2$ is a non-central Chi-square RV whose conditional moment generating function (MGF) is given in [15]. Averaging over $\left\{\hat{h}_l\right\}_{l\in\hat{\Omega}_L}$, we get

$$P_{\text{MPSK}} = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \left(1 + \frac{(a_2 - a_1)b}{\sin^2\theta}\right)^{-L} \times \mathcal{M}_{\sum\limits_{k\in\hat{\Omega}_L} \left|\hat{h}_k\right|^2}\left(\frac{-b}{\sin^2\theta + (a_2 - a_1)b}\right) d\theta. \quad (18)$$

where $\mathcal{M}_{\sum_{k\in\hat{\Omega}_L} \left|\hat{h}_k\right|^2}(.)$ is the MGF of $Y = \sum_{k\in\hat{\Omega}_L} \left|\hat{h}_k\right|^2$. It equals [5]

$$\mathcal{M}_{\sum\limits_{k\in\hat{\Omega}_L} \left|\hat{h}_k\right|^2}(x) = (1 - a_1 x)^{-L} \prod_{n=L+1}^{N} \left(1 - \frac{a_1 L x}{n}\right)^{-1}.$$

Simplifying (18) further gives the desired expression.

References

[1] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Mag.*, vol. 5, pp. 46–56, Mar. 2004.

[2] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, pp. 68–73, Oct. 2004.

[3] "Draft amendment to wireless LAN media access control (MAC) and physical layer (PHY) specifications: Enhancements for higher throughput," Tech. Rep. P802.11n/D0.04, IEEE, Mar. 2006.

[4] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Process. Mag.*, pp. 12–25, Nov. 2004.

[5] W. M. Gifford, M. Z. Win, and M. Chiani, "Antenna subset diversity with non-ideal channel estimation," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1527–1539, May 2008.

[6] T. Gucluoglu and E. Panayirci, "Performance of transmit and receive antenna selection in the presence of channel estimation errors," *IEEE Commun. Lett.*, vol. 12, pp. 371–373, May 2008.

[7] P. Polydorou and P. Ho, "Error performance of MPSK with diversity combining in non-uniform Rayleigh fading and non-ideal channel estimation," in *Proc. VTC (Spring)*, pp. 627–631, May 2000.

[8] Y. Peng, S. Cui, and R. You, "Optimal pilot-to-data power ratio for diversity combining with imperfect channel estimation," *IEEE Commun. Lett.*, vol. 10, pp. 97–99, Feb. 2006.

[9] A. Maaref and S. Aissa, "Optimized rate-adaptive PSAM for MIMO MRC systems with transmit and receive CSI imperfections," *IEEE Trans. Commun.*, vol. 57, pp. 821–830, Mar 2009.

[10] A. F. Molisch, *Wireless Communications*. Wiley-IEEE Press, 2005.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1. Prentice Hall Signal Processing Series, 1 ed., 1993.

[12] V. Kristem, N. B. Mehta, and A. F. Molisch, "A novel, balanced, and energy-efficient training method for receive antenna selection," *Submitted to IEEE Trans. Wireless Commun.*, 2009.

[13] M. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*. Wiley-Interscience, 2 ed., 2005.

[14] A. F. Molisch, M. Z. Win, and J. H. Winters, "Performance of reduced-complexity transmit/receive-diversity systems," in *Proc. PIMRC*, pp. 739–742, Oct. 2002.

[15] M. Abramowitz and I. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover, 9 ed., 1972.