Energy-efficient Training for Antenna Selection in Time-varying Channels

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

***Abstract*—Training for receive antenna selection (AS) differs from that for conventional multiple antenna systems because of the limited hardware usage inherent in AS. We analyze and optimize the performance of a novel energy-efficient training method tailored for receive AS. In it, the transmitter sends not only pilots that enable the selection process, but also an extra pilot that leads to accurate channel estimates for the selected antenna that actually receives data. For time-varying channels, we propose a novel antenna selection rule and prove that it minimizes the symbol error probability (SEP). We also derive closed-form expressions for the SEP of MPSK, and show that the considered training method is significantly more energy-efficient than the conventional AS training method.**

I. Introduction

Single receive antenna selection (AS) provides a low hardware complexity solution for exploiting the spatial diversity benefits of receiving with multiple antennas [1]. It dynamically selects the antenna with the 'best' instantaneous channel gain to the transmitter, and only processes the signal received from it. AS enables a receiver with N antennas to employ only one radio frequency (RF) chain instead of N RF chains. Consequently, AS has been adopted in next generation wireless systems such as IEEE 802.11n [2]. Despite its lower hardware complexity, AS with perfect channel state information (CSI) achieves the full diversity order [3], [4].

In practice, in order to select the best antenna, the CSI of all the antennas must be acquired at the receiver using a pilot-based training scheme. The low hardware complexity, which is a key motivator for AS, constrains how training gets done for AS since only one antenna can receive at any time instant. Consequently, the transmitter needs to send a pilot symbol N times to enable the receiver to *sequentially* estimate the channel gains of its N antennas to the transmitter [5]; depending on the system design, the pilots can be several milliseconds apart. The receive antenna is then selected based on these estimates.

An improved training scheme for AS was proposed in [6], assuming that the channels do not vary over the training phase and the subsequent data transmission phase. In it, the transmitter sends an extra pilot symbol after the first N pilots, which we shall refer to as selection pilots. While the selection pilots are used to select the best antenna, the extra pilot helps in refining the channel estimates of the selected antenna that will be used for data reception. The robustness of AS to selection errors enables the transmitter to significantly reduce energy it allocates to the selection pilots and instead use it to boost the energy of the extra pilot and the data symbols, and improve overall performance. Henceforth, we shall refer to this scheme as the *extra pilot* scheme.

V. Kristem and A. F. Molisch are with the Dept. of Electrical Eng. at the Univ. of Southern California, CA, USA. N. B. Mehta is with the Dept. of Electrical Communication Eng. at the Indian Institute of Science (IISc), Bangalore, India.

Emails: vinod.kristem@gmail.com, nbmehta@ece.iisc.ernet.in, molisch@usc.edu

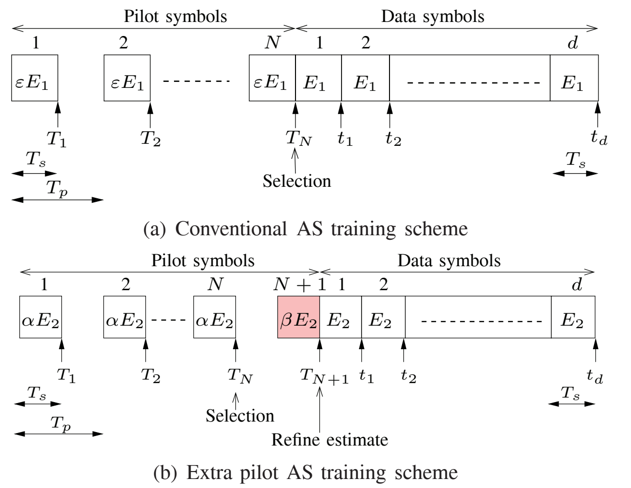

Fig. 1. Training for receive antenna selection: illustration of conventional and extra pilot schemes.

In this paper, we analyze the performance of the extra pilot scheme in *time-varying* channels and demonstrate the significant energy-efficiency gains its offers. We show that the larger the Doppler spread, the greater is the energy-efficiency gain from the scheme. Under a total energy constraint, we derive a novel SEP-optimal antenna selection rule for it. As per the rule, an affine function of the estimated channel gain is evaluated for each antenna, and the one with the largest value is selected. This rule differs from the ones considered in [5], [7], which did not consider an extra pilot, and the no-weighting rule assumed in several other papers in the literature, e.g., [8]–[10]. We also derive the closed-form expressions for the fading-averaged SEP of the scheme for MPSK constellations. The optimal AS rule and the SEP analysis take into account the imperfection of the channel estimates due to noise and training delays and their impact on both selection and data demodulation. They also explicitly account for the fact that the channel estimates of different antennas, which are obtained at different times, are outdated by different amounts by the time the data is demodulated.

II. System Model

Consider a system with one transmit antenna, N receive antennas, and one receive RF chain. Let $h_k(t)$ denote the frequency-flat channel between the transmitter and the k^{th} receive antenna at time t. It is modeled as a circularly symmetric complex Gaussian random variable (RV) with unit variance. The channel gains of different receive antennas are assumed to be independent and identically distributed (i.i.d.). Let $\rho(t_1, t_2)$ denote the correlation of $h_k(t)$ at times t_1 and t_2. Noise is modeled as circularly symmetric Gaussian noise with variance N_0, and is independent across antennas and time.

We shall use the following notation. $\Pr(A)$, $\mathbf{E}[A]$, and $\mathbf{var}[A]$ denote the probability, expectation, and variance of A, respectively. Similarly, $\Pr(A|B)$, $\mathbf{E}[A|B]$, and $\mathbf{var}[A|B]$ denote the conditional probability, expectation, and variance of A given B, respectively. The complex conjugate of x is denoted by x^*. The Hermitian transpose and transpose are denoted by $(\cdot)^H$ and $(\cdot)^T$, respectively. The cardinality of a set A is denoted by $|A|$. The covariance matrix of two random vectors X and Y is denoted by $\Sigma_{XY} \triangleq \mathbf{E}\left[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])^H\right]$.

We now describe the conventional and extra pilot AS training schemes.

A. Conventional AS Training Scheme

To enable the receiver to estimate the channel gains of all N antennas, the transmitter first transmits N pilot symbols, each of energy εE_1 and duration T_s, as illustrated in Fig. 1(a). Here, E_1 is the energy allocated to a data symbol and $\varepsilon \geq 0$ is the energy scaling factor for pilots.

The signal, $r_k(T_k)$, received by the k^{th} receive antenna at time T_k is given by

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

where p is a complex pilot symbol with $|p| = 1$ and n_k is Gaussian noise. The minimum mean square error (MMSE) channel estimate, $\widehat{h}_k(t_i)$, of $h_k(t_i)$ is given by

$$\widehat{h}_k(t_i) = \frac{\sqrt{\varepsilon E_1} p^* \rho(T_k, t_i)}{\varepsilon E_1 + N_0} r_k(T_k). \tag{2}$$

Two consecutive pilot symbols are separated in time by a duration T_p. As shown in [5], the SEP-optimal selection rule weights $\widehat{h}_k(t_i)$ and selects the antenna with the largest weighted channel gain estimate. The weights are given in [5, Theorem 3], and are not repeated here. The pilots are followed by d data symbols, all of which are received by the same selected antenna. The i^{th} data symbol is transmitted at time t_i and with energy E_1.

The constraint on the total energy, E_T, then takes the form

$$(N\varepsilon + d)E_1 = E_T. \tag{3}$$

Let $\gamma \triangleq \frac{E_T}{N_0}$. Notice that due to the total energy constraint, ε affects the energy allocated, E_1, to each data symbol. The SEP of MPSK, $P_{\text{MPSK}}^{\varepsilon}(\gamma)$, as a function of ε and γ is given by [5]

$$P_{\text{MPSK}}^{\varepsilon}(\gamma) = \frac{1}{\pi} \sum_{k=1}^{N} \sum_{l=0}^{N-1} \sum_{m=1}^{\left|\mathcal{S}_l^k\right|} \int_0^{\frac{M-1}{M}\pi} \frac{(-1)^l}{\sigma_k^2} \times \left(\frac{1}{\sigma_k^2} + \frac{\sin^2\left(\frac{\pi}{M}\right) / \sin^2\theta}{1-\sigma_k^2+\gamma^{-1}} + \sum_{z \in \mathcal{S}_k^l(m)} \frac{1-\sigma_z^2+\gamma^{-1}}{\left(1-\sigma_k^2+\gamma^{-1}\right)\sigma_z^2} \right)^{-1} d\theta, \tag{4}$$

where M is the constellation size, $\sigma_z^2 \triangleq \mathbf{var}\left[\widehat{h}_z\right] = \frac{\varepsilon|\rho(T_z,t_i)|^2}{\varepsilon+\gamma^{-1}}$, $\mathcal{S}_l^k$ denotes the set of all l-element subsets formed from the set $\{1, 2, \ldots, N\} \setminus \{k\}$, and $\mathcal{S}_l^k(m)$ is the m^{th} l-element subset.

B. Extra Pilot AS Training Scheme

The extra pilot scheme is shown in Fig. 1(b). As before, the transmitter now first sequentially transmits N pilot symbols so that all the N channels can be estimated. However, each pilot symbol is now transmitted with energy αE_2, where E_2 is the energy per data symbol and $\alpha \geq 0$ is the energy scaling factor for the selection pilots.

The pilot signal received by the k^{th} receive antenna at time T_k is given by

$$r_k(T_k) = \sqrt{\alpha E_2} p h_k(T_k) + n_k, \tag{5}$$

where n_k is Gaussian noise. As in (2), the MMSE channel estimate, $\widehat{h}_k$, of $h_k(t_i)$ is given by

$$\widehat{h}_k = \frac{\sqrt{\alpha E_2} p^* \rho(T_k, t_i)}{\alpha E_2 + N_0} r_k(T_k). \tag{6}$$

Note that $\widehat{h}_k$ is a circularly symmetric complex Gaussian RV with variance $\sigma_k^2 \triangleq \mathbf{var}\left[\widehat{h}_k\right] = \frac{\alpha|\rho(T_k,t_i)|^2}{\alpha+\gamma^{-1}}$.

Based on the channel estimates $\widehat{h}_1, \ldots, \widehat{h}_N$, the receiver selects the antenna that will be used to receive data. The selection rule is derived in Sec. III-C. Let $\widehat{[1]}$ denote the index of the selected antenna.

Extra Pilot Transmission: In this scheme, an extra pilot symbol is transmitted at time T_{N+1} with energy βE_2, and is received only by the *selected antenna*. The signal received by the selected antenna is

$$r_{\widehat{[1]}}(T_{N+1}) = \sqrt{\beta E_2} p h_{\widehat{[1]}}(T_{N+1}) + n', \tag{7}$$

where n' is Gaussian noise.

Data Reception: The pilot symbols are followed by d data MPSK symbols, each transmitted with energy E_2. They are all received by the selected antenna. The received signal for the i^{th} data symbol is

$$y_{\widehat{[1]}}(t_i) = h_{\widehat{[1]}}(t_i) s + n'', \tag{8}$$

where n'' is Gaussian noise. The total energy constraint now becomes

$$E_T = (N\alpha + \beta + d) E_2. \tag{9}$$

As before, let $\gamma \triangleq \frac{E_T}{N_0}$. Now, both α and β affect the energy, E_2, allocated to a data symbol.

III. SEP Analysis & Optimization of Extra Pilot Scheme

We now derive the optimal receive antenna selection rule and the SEP with optimal selection for the extra pilot scheme.

A. Refined Channel Estimate

Unlike the other antennas, two observations – one at time $T_{\widehat{[1]}}$ (given by (6)) and the other at time T_{N+1} (given by (7)) – are available at the receiver for the selected antenna, $\widehat{[1]}$. As shown below, they can be used to refine its channel estimate.

Lemma *1:* Let $k = \widehat{[1]}$. The refined MMSE channel estimate, $\widehat{\widehat{h}}_k$, of the k^{th} antenna's channel gain at time t_i, obtained from the observations $r_k(T_k)$ and $r_k(T_{N+1})$, is given by

$$\widehat{\widehat{h}}_k = \left(\alpha\beta\left(1-\rho\left(T_k,T_{N+1}\right)^2\right) + (\alpha+\beta)\gamma^{-1} + \gamma^{-2}\right)^{-1} \times \Bigg[\frac{p^* r_k(T_k)}{\sqrt{E}}\Big\{\beta\sqrt{\alpha}\left(\rho\left(T_k,t_i\right) - \rho\left(T_{N+1},t_i\right)\rho\left(T_k,T_{N+1}\right)\right) + \frac{\sqrt{\alpha}}{\gamma}\rho\left(T_k,t_i\right)\Big\} + \frac{r_k(T_{N+1})p^*}{\sqrt{E}}\Big\{\sqrt{\beta}\gamma^{-1}\rho\left(T_{N+1},t_i\right) + \alpha\sqrt{\beta}\left(\rho\left(T_{N+1},t_i\right) - \rho\left(T_k,t_i\right)\rho\left(T_k,T_{N+1}\right)\right)\Big\}\Bigg]. \quad (10)$$

Further, $\widehat{\widehat{h}}_k$ is a circularly symmetric complex Gaussian RV, whose variance, ${\sigma'_k}^2$, is given by

$${\sigma'_k}^2 = \left(\alpha\beta\left(1-\rho(T_k,T_{N+1})^2\right) + (\alpha+\beta)\gamma^{-1} + \gamma^{-2}\right)^{-1} \times \left[\alpha\gamma^{-1}\rho(T_k,t_i)^2 + \beta\gamma^{-1}\rho(T_{N+1},t_i)^2 + \alpha\beta\left(\rho(T_k,t_i)^2 + \rho(T_{N+1},t_i)^2 - 2\rho(T_{N+1},t_i)\rho(T_k,T_{N+1})\rho(T_k,t_i)\right)\right]. \quad (11)$$

Proof: The proof is given in Appendix A. ■

B. Decision Variable and Statistics

The maximum likelihood (ML) decision variable, $\mathcal{D}$, for detecting the symbol received at time t_i is

$$\mathcal{D} = \widehat{\widehat{h}}^*_{\widehat{[1]}} y_{\widehat{[1]}}(t_i). \quad (12)$$

Conditioned on $\widehat{[1]}$, $\widehat{\widehat{h}}_k$, and s_i, $\mathcal{D}$ is a circularly symmetric complex Gaussian RV with first and second order moments given by

$$\mathbf{E}\left[\mathcal{D}\,\middle|\,\widehat{[1]}=k, \widehat{\widehat{h}}_k, s_i\right] = \left|\widehat{\widehat{h}}_k\right|^2 s_i, \quad (13)$$

$$\mathbf{var}\left[\mathcal{D}\,\middle|\,\widehat{[1]}=k, \widehat{\widehat{h}}_k, s_i\right] = \left|\widehat{\widehat{h}}_k\right|^2 \left(\left(1-{\sigma'_k}^2\right)E + N_0\right). \quad (14)$$

C. Optimal Selection Rule and SEP

With the above results on the conditional mean and variance of the decision variable, the optimal AS rule is as follows.

Theorem *1:* The optimal receive AS rule that minimizes the SEP of the MPSK symbol transmitted at time t_i is

$$\widehat{[1]} = \arg\max_{1\le k\le N}\left(a_k + b_k\left|\widehat{h}_k\right|^2\right), \quad (15)$$

where $\widehat{h}_k$ is the MMSE channel estimate obtained from the k^{th} pilot symbol, $a_k = \ln\left(1+c_k\right)$, $b_k = \frac{\sin^2\left(\frac{\pi}{M}\right)}{(1+c_k)\left(1-{\sigma'_k}^2+\gamma^{-1}\right)}$, and $c_k = \left({\sigma'_k}^2 - \sigma_k^2\right)\frac{\sin^2\left(\frac{\pi}{M}\right)}{1-{\sigma'_k}^2+\gamma^{-1}}$.

Proof: The proof is relegated to Appendix B. ■

Theorem *2:* The SEP of the MPSK data symbol transmitted at time t_i with the above optimal AS rule is given by

$$P_{\text{MPSK}}^{\alpha-\beta}(\gamma) = \frac{1}{\pi}\sum_{k=1}^{N}\sum_{l=0}^{N-1}\frac{(-1)^l}{\sigma_k^2}\sum_{m=0}^{\left|\mathcal{S}_l^k\right|}\exp\left(-\sum_{z\in\mathcal{S}_l^k(m)}\frac{a_k-a_z}{b_z\sigma_z^2}\right) \times \int_0^{\frac{M-1}{M}\pi}\left(1+\frac{c_k}{\sin^2\theta}\right)^{-1}\frac{\exp\left(-g_k f\left(\mathcal{S}_l^k(m)\right)\right)}{f\left(\mathcal{S}_l^k(m)\right)}\,d\theta, \quad (16)$$

where $g_k = \max_{l=1,\ldots,N, l\neq k}\frac{a_l-a_k}{b_k}$ and $f\left(\mathcal{S}_l^k(m)\right) = \frac{1}{\sigma_k^2} + \sum_{z\in\mathcal{S}_l^k(m)}\frac{b_k}{b_z\sigma_z^2} + \frac{\sin^2\left(\frac{\pi}{M}\right)}{\left(\sin^2\theta+c_k\right)\left(1-{\sigma'_k}^2+\gamma^{-1}\right)}$.

Proof: The proof is given in Appendix C. ■

The optimal value of ε, denoted by ε^*, for the conventional scheme and the optimal values of α and β, denoted by α^* and β^*, for the extra pilot scheme can now be found numerically by means of a gradient search over the corresponding closed-form SEP expressions in (4) and (16).

IV. Simulations Results

We now plot the analytical results derived in Sec. II and validate them with Monte Carlo simulations that use 10^4 samples for each point. The modified Jakes simulator of [11] is used to generate the time-varying Rayleigh channels. Therefore, the correlation values for $k = 1, 2, \cdots, N+1$ and $i = 1, 2, \cdots, d$ equal $\rho\left(T_k, t_i\right) = J_0(2\pi f_d((N+1-k)T_p + iT_s))$. We shall focus on the case with one data symbol ($d = 1$) in this paper due to space limitations. As we shall see, even this case provides valuable insights about the effect of time-variations on the training for AS.

Figure 2 plots the SEP of the conventional and extra pilot schemes as a function of $\gamma = \frac{E_T}{N_0}$ for optimal values of their respective parameters. Two different values of Doppler spread are considered. We see that for the same SEP, the extra pilot scheme requires 2-4 dB less total energy than the conventional scheme, which is a significant improvement. Furthermore, the gains from the extra pilot scheme increase as the Doppler spread increases. This is because the extra pilot ensures that the estimate of the selected antenna, which is used for coherent demodulation, is less outdated than the estimate used in the conventional scheme, where any of the N estimates – and not just the one from the extra pilot – could be used for coherent reception. This effect becomes more pronounced as the Doppler spread increases. Notice also the excellent match between the analytical and simulation curves. We, therefore, show only the analytical curves henceforth.

Figure 3 compares the SEPs of the conventional and extra pilot schemes for 4PSK, 8PSK, and 16PSK for $f_dT_p = 0.01$. We see that the extra pilot scheme outperforms the conventional scheme for all the constellations. Further, the gains

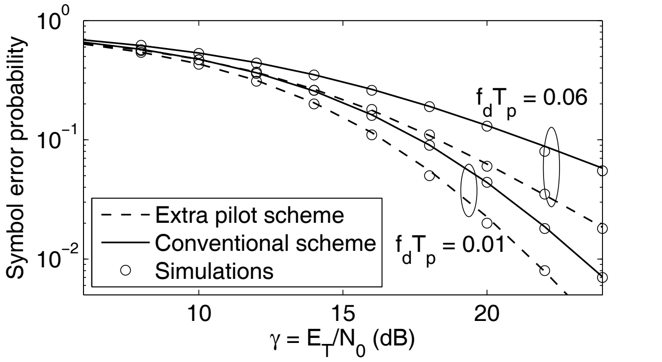

Fig. 2. SEP comparison of conventional and extra pilot AS training schemes for different Doppler spreads (8PSK, $N = 4$, $T_p = 10T_s$, and $d = 1$).

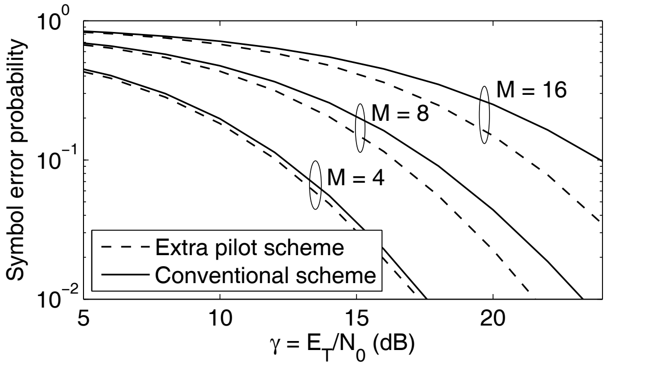

Fig. 3. SEP comparison of conventional and extra pilot AS training schemes for different constellation sizes ($N = 4$, $f_dT_p = 0.01$, $T_p = 10T_s$, and $d = 1$).

increase with the constellation size. This is because the larger constellations are more sensitive to channel estimation errors, which are reduced significantly by the extra pilot scheme.

Figure 4 plots the optimal energy allocation (ε^* for the conventional scheme and α^* and β^* for the extra pilot scheme) as a function of the total energy for two different values of f_dT_p. Interestingly, for the conventional scheme, ε^* does not depend on the Doppler spread. It also increases monotonically with γ. However, in the extra pilot scheme, α^* and β^* do depend on the Doppler spread. For low γ, α^* is close to zero and β^* is large. This is because the energy is insufficient to enable accurate selection. Consequently, very little energy is used to get the estimates for AS. Instead, energy is allocated to the extra pilot that is used for coherent demodulation. Once γ crosses a threshold, the system allocates more energy to the selection pilots to benefit from selection diversity. This triggers a corresponding decrease in β^* since the selection pilot and the last pilot are both used for coherent demodulation.

V. Conclusions

We analyzed a novel training scheme that is tailored for receive antenna selection in time-varying channels. The scheme introduces an extra pilot that helps the receiver selectively refine its channel estimate of the selected antenna. By exploiting

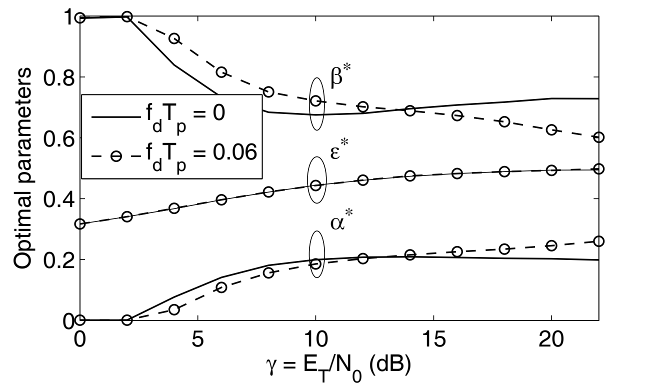

Fig. 4. Optimal values of the parameters of the conventional and extra pilot AS training schemes (8PSK, $N = 4$, $T_p = 10T_s$, and $d = 1$).

the fact that antenna selection is robust to channel estimation errors, it assigns less energy to pilots that are used by the receiver to select an antenna. We saw that in time-varying channels, the optimal antenna selection rule is affine in nature, and is different from both the no-weighting and the linear-weighting based selection rules that have been considered in the literature. Closed-form expressions for the SEP of MPSK facilitated an optimization of the allocation of energy across pilots and data. For different constellation sizes, we saw that the extra pilot training scheme is more energy-efficient than the conventional training scheme, and the energy-efficiency gains increase as the Doppler spread increases.

Appendix

A. Proof of Lemma 1

The MMSE estimate of $X \triangleq h_k(t_i)$, which is obtained from the observations $\mathbf{Y} \triangleq [r_k(T_k)\ r_k(T_{N+1})]^T$, is given by

$$\widehat{\widehat{h}}_k = \Sigma_{X\mathbf{Y}}\Sigma_{\mathbf{YY}}^{-1}(\mathbf{Y} - \mathbf{E}\left[\mathbf{Y}\right]). \tag{17}$$

Now substituting $\mathbf{E}\left[h_k(s)h_k(t)^*\right] = \rho(s,t)$, $\mathbf{var}\left[r_k(T_k)\right] = \alpha E_2 + N_0$, and $\mathbf{var}\left[r_k(T_{N+1})\right] = \beta E_2 + N_0$ yields (10).

Since $\widehat{\widehat{h}}_k$ is linear combination of circularly symmetric complex Gaussian RVs, the resulting RV is also a circularly symmetric complex Gaussian. Its variance is

$${\sigma'_k}^2 = \mathbf{var}\left[X\right] - \Sigma_{X\mathbf{Y}}\Sigma_{\mathbf{YY}}^{-1}\Sigma_{\mathbf{Y}X}. \tag{18}$$

Again substituting the mean and variance of the observations in (18), and simplifying further yields (11).

B. Proof of Theorem 1

Starting from [12, (40)], it can be shown that the SEP conditioned on $\widehat{[1]}$, $\widehat{h}_{\widehat{[1]}}$, and the transmitted symbol s_i, is

$$P_{\text{MPSK}}^{\alpha-\beta}(\gamma \,\Big|\, \widehat{[1]} = k, \widehat{\widehat{h}}_k, s_i)$$
$$= \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \exp\left(\frac{-\left|\widehat{\widehat{h}}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(1 - {\sigma'_k}^2 + \gamma^{-1}\right)\sin^2\theta}\right) d\theta. \tag{19}$$

Unconditioning the SEP over $\widehat{\widehat{h}}_k$ given $\widehat{h}_k$, we get

$$P^{\alpha-\beta}_{\text{MPSK}}(\gamma \,\Big|\, \widehat{[1]}=k, \widehat{h}_k, s_i) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \mathcal{M}_{\left|\widehat{\widehat{h}}_k\right|^2 \,\big|\, \widehat{h}_k} \left(\frac{-\sin^2\left(\frac{\pi}{M}\right)}{\left(1-{\sigma'_k}^2+\gamma^{-1}\right)\sin^2\theta} \right) d\theta, \quad (20)$$

where $\mathcal{M}_{\left|\widehat{\widehat{h}}_k\right|^2 \,\big|\, \widehat{h}_k}(.)$ is the moment generating function (MGF) of $\left|\widehat{\widehat{h}}_k\right|^2$ conditioned on $\widehat{h}_k$.

It is easy to see that $\widehat{\widehat{h}}_k$ conditioned on $\widehat{h}_k$ is a complex Gaussian RV with $\mathbf{E}\left[\widehat{\widehat{h}}_k \,\Big|\, \widehat{h}_k\right] = \widehat{h}_k$ and $\mathbf{var}\left[\widehat{\widehat{h}}_k \,\Big|\, \widehat{h}_k\right] = {\sigma'_k}^2 - \sigma_k^2$. Hence,

$$\mathcal{M}_{\left|\widehat{\widehat{h}}_k\right|^2 \,\big|\, \widehat{h}_k}(x) = \frac{1}{1-x\left({\sigma'_k}^2-\sigma_k^2\right)} \exp\left(\frac{\left|\widehat{h}_k\right|^2 x}{1-x\left({\sigma'_k}^2-\sigma_k^2\right)} \right).$$

Substituting this in (20) and simplifying further yields

$$P^{\alpha-\beta}_{\text{MPSK}}(\gamma \,\Big|\, \widehat{[1]}=k, \widehat{h}_k) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \left(1+\frac{c_k}{\sin^2\theta}\right)^{-1} \times \exp\left(-\frac{\left|\widehat{h}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(c_k+\sin^2\theta\right)\left(1-{\sigma'_k}^2+\gamma^{-1}\right)} \right) d\theta, \quad (21)$$

where c_k is defined in the theorem statement.

The above expression is in the form of a single integral, minimizing which is analytically intractable. In order to derive the optimal selection rule, we, therefore, focus on its Chernoff upper bound. Replacing θ with $\pi/2$ yields the upper bound

$$P^{\alpha-\beta}_{\text{MPSK}}(\gamma \,\Big|\, \widehat{[1]}=k, \widehat{h}_k) \leq \frac{M-1}{M}\left(1+c_k\right)^{-1} \exp\left(\frac{-\left|\widehat{h}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(1+c_k\right)\left(1-{\sigma'_k}^2+\gamma^{-1}\right)} \right). \quad (22)$$

Therefore, the SEP-optimal antenna is

$$\widehat{[1]} = \arg\min_{k=1,\ldots,N} \left(1+c_k\right)^{-1} \exp\left(\frac{-\left|\widehat{h}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(1+c_k\right)\left(1-{\sigma'_k}^2+\gamma^{-1}\right)} \right).$$

Simplifying this yields (15).

C. Proof of Theorem 2

From (21), the SEP conditioned on the selected antenna and the estimates $\hat{h}_i$, $1 \leq i \leq N$, is given by

$$P^{\alpha-\beta}_{\text{MPSK}}(\gamma \,\Big|\, \widehat{[1]}=k, \widehat{h}_k) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \left(1+\frac{c_k}{\sin^2\theta}\right)^{-1} \times \exp\left(-\frac{\left|\widehat{h}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(c_k+\sin^2\theta\right)\left(1-{\sigma'_k}^2+\frac{1}{\gamma}\right)} \right) d\theta. \quad (23)$$

Averaging over the channel estimates, we get

$$P^{\alpha-\beta}_{\text{MPSK}}(\gamma) = \frac{1}{\pi}\sum_{k=1}^{N}\int_0^{\frac{M-1}{M}\pi}\int_{\left|\widehat{h}_1\right|^2=0}^{\infty}\cdots\int_{\left|\widehat{h}_N\right|^2=0}^{\infty}\left(1+\frac{c_k}{\sin^2\theta}\right)^{-1} \times \exp\left(-\frac{\left|\widehat{h}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(c_k+\sin^2\theta\right)\left(1-{\sigma'_k}^2+\frac{1}{\gamma}\right)} \right) \frac{1}{\sigma_k^2}\exp\left(\frac{-\left|\widehat{h}_k\right|^2}{\sigma_k^2} \right) \times \prod_{l=1,l\neq k}^{N} \Pr\left(a_k + b_k\left|\widehat{h}_k\right|^2 > a_l + b_l\left|\widehat{h}_l\right|^2 \right) d\left|\widehat{h}_1\right|^2 \cdots d\left|\widehat{h}_N\right|^2 d\theta.$$

Integrating out $N-1$ channel estimates yields

$$P^{\alpha-\beta}_{\text{MPSK}}(\gamma) = \frac{1}{\pi}\sum_{k=1}^{N}\sum_{l=0}^{N-1}(-1)^l\sum_{m=1}^{\left|S_l^k\right|}\int_0^{\frac{M-1}{M}\pi}\int_{g_k}^{\infty}\left(1+\frac{c_k}{\sin^2\theta}\right)^{-1} \times \exp\left(-\frac{\left|\widehat{h}_k\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\left(c_k+\sin^2\theta\right)\left(1-{\sigma'_k}^2+\frac{1}{\gamma}\right)} \right) \frac{1}{\sigma_k^2}\exp\left(\frac{-\left|\widehat{h}_k\right|^2}{\sigma_k^2} \right) \times \exp\left(-\sum_{z\in S_l^k(m)} \left(\frac{a_k-a_z}{b_z\sigma_z^2} + \frac{b_k}{b_z\sigma_z^2}\left|\widehat{h}_k\right|^2 \right) \right) d\left|\widehat{h}_k\right|^2 d\theta,$$

where g_k is defined in the theorem statement. Evaluating the integral over $\left|\widehat{h}_k\right|^2$ and simplifying yields (16).

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] "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.

[3] A. Ghrayeb and T. M. Duman, "Performance analysis of MIMO systems with antenna selection over quasi-static fading channels," *IEEE Trans. Veh. Technol.*, vol. 52, pp. 281–288, Mar. 2003.

[4] M. Z. Win and J. H. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 49, pp. 1926–1934, Nov. 2001.

[5] V. Kristem, N. B. Mehta, and A. F. Molisch, "Optimal receive antenna selection in time-varying fading channels with practical training constraints," *IEEE Trans. Commun.*, vol. 58, pp. 2023–2034, Jul. 2010.

[6] V. Kristem, N. B. Mehta, and A. F. Molisch, "A novel, balanced, and energy-efficient training method for receive antenna selection," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2742–2753, Sep. 2010.

[7] V. Kristem, N. B. Mehta, and A. F. Molisch, "Training and voids in receive antenna subset selection in time-varying channels," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 1992–2003, Jun. 2011.

[8] 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.

[9] A. B. Narasimhamurthy and C. Tepedelenlioglu, "Antenna selection for MIMO OFDM systems with channel estimation error," in *Proc. Globecom*, pp. 3290–3294, 2007.

[10] 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.

[11] Y. Li and Y. Guan, "Modified Jakes model for simulating multiple uncorrelated fading waveforms," in *Proc. ICC*, pp. 46–49, 2000.

[12] M.-S. Alouini and A. J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1324–1334, Sept. 1999.