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

On Training and Training Voids for Receive Antenna Subset Selection in Time-Varying Channels

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

***Abstract*—Antenna selection (AS) provides most of the benefits of multiple-antenna systems at drastically reduced hardware costs. In receive AS, the receiver connects a dynamically selected subset of N available antennas to the L available RF chains. The "best" subset to be used for data reception is determined by means of channel estimates acquired using training sequences. Due to the nature of AS, the channel estimates at different antennas are obtained from different transmissions of the pilot sequence, and are, thus, outdated by different amounts in a time-varying channel. We show that a linear weighting of the estimates is optimum for the subset selection process, where the weights are related to the temporal correlation of the channel variations. When L is not an integer divisor of N, we highlight a new issue of "training voids", in which the last pilot transmission is not fully exploited by the receiver. We present a "void-filling" method for fully exploiting these voids, which essentially provides more accurate training for some antennas, and derive the optimal subset selection rule for any void-filling method. We also derive new closed-form equations for the performance of receive AS with optimal subset selection.**

I. Introduction

Antenna selection (AS) is a popular technique to reduce the hardware costs at the transmitter or receiver of a wireless system [1]–[4]. It uses a small number, L, of radio frequency (RF) chains and a larger number, N, of antennas, and only processes signals from a dynamically selected subset of antennas. This is advantageous since antenna elements are typically cheap, while the RF chains are expensive. Consequently, many next generation wireless communications standards such as IEEE 802.11n, third generation partnership project (3GPP) long term evolution (LTE), and the IEEE 802.16m Advanced WiMax have standardized or are standardizing AS at the transmitter or the receiver, or both.

We focus on receive AS in this paper, and consider a transmitter with one antenna. While receive AS has received attention in the literature, the topic of training – i.e., the estimation of channel states to determine the optimum antenna subset that should be used for reception – is often treated cursorily and/or based on highly idealized assumptions. For example, several papers assume perfect channel state information (CSI) at the receiver. In practice, the CSI is acquired using a pilot-based training scheme, and is imperfect because of outdated and noisy channel estimates. Imperfect CSI can lead to inaccurate selection and imperfect data decoding, both of which increase the symbol error probability (SEP). While [5], [6] explored receive AS with imperfect channel estimates, they did not consider the effect of outdated estimates in time-varying channels. While [7]–[9] did consider outdated channel estimates as well, they did not account for the fact that estimates at different antennas are outdated by different amounts, as detailed in the next paragraph.

The basic operation of AS imposes the constraint that only L antennas channel gains 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 antenna subset with the best channels. For example, in the IEEE 802.11n standard, a multiple access control (MAC) based training protocol is used for transmit and/or receive AS, in which a sequence of consecutive 'training packets' are transmitted to obtain the CSI of all antennas [10]. The primary reason for doing this is to ensure that the physical layer protocols in the standard do not have to be modified considerably to accommodate AS training. The pilots embedded in the physical layer header of each training packet help estimate the channel gains. Each training packet, which carries a physical layer header and a data payload, can be several milliseconds long. For this reason, the pilots sent across the training packets are also spaced several milliseconds apart. Thus, the CSI at different antennas is outdated by different amounts. This fact significantly impacts how the antenna subset should be selected, as was shown for single receive AS ($L = 1$) in [11].[1]

In this paper, we consider the more general case of optimal subset selection for $L > 1$. Considering subset selection also leads to new interesting questions about how to allocate the available L RF chains at the receiver for training when L is not an integer divisor of N. With N receive antennas and L RF chains, the minimum number of training symbols required is $\lceil \frac{N}{L} \rceil$, and $N - L \lfloor \frac{N}{L} \rfloor$ RF chains are unused in the last pilot symbol, where $\lceil . \rceil$ and $\lfloor . \rfloor$ denote the ceil and floor functions, respectively. We shall call the unused RF chains as *training voids*. Such voids do not occur when L divides N, which occurs, for example, for $L = 1$.

As an example, consider Fig. 1 with $N = 5$ and $L = 2$. To

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, molisch@usc.edu

[1]For example, sorting estimated channel gains in decreasing order and selecting the first L of them, as is done for time-invariant channels, is no longer optimal. Intuitively, the antenna with the highest channel gain estimate may not be the best one to select in case its estimate is very outdated.

978-1-4244-5638-3/10/$26.00 ©2010 IEEE

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

estimate the channels of all 5 antennas, at least 3 pilot transmissions are needed. The first pilot helps estimate channels of antennas #1 and #2, the second pilot helps estimate channels of antennas #3 and #4. For the third pilot, only one antenna's (#5) channel remains to be estimated. Consequently, one of the two RF chains is not needed for receiving the third pilot.

Instead of leaving the RF chains unused, we propose connecting them to some of the antennas. In our example above, this would imply connecting the second RF chain during the third pilot symbol to antenna #1, #2, #3, or #4. For the antenna so connected, the receiver now has two estimates of the channel to the transmitter, and can, thus, reduce the estimation error for it. We shall call this *void-filling*, and shall show that it significantly improves performance.

Our specific contributions are the following:

- We analyze and optimize the performance of antenna subset selection ($L > 1$) over time-varying Rayleigh fading channels given a practical training model for AS. The model leads to noisy channel estimates, which affect both the subset selection and demodulation. We derive, for the first time, optimal criteria for selecting the best antenna subset as a function of the time correlation of the channel. We also derive closed-form expressions for the SEP with optimal subset selection. These results generalize the results of [11], which considered only $L = 1$.
- We develop the optimal subset selection rule for any given void-filling rule. Thereafter, we develop a new random void-filling method that fully utilizes all the available RF chains at the receiver all the time and yields better performance. To the best of our knowledge, this is the first paper that introduces the practical problem of training voids in receive AS and proposes an effective and analytically tractable method to address it.

The outline of this paper is as follows. The AS training is described in Sec. II and analyzed in Sec. III. Void-filling is investigated in Sec. IV. Simulation results and conclusions follow in Sec. V and Sec. VI, respectively. Mathematical details are relegated to the Appendix.

II. System Model

Consider a system with one transmit antenna, N receive antennas, and L RF chains at the receiver. 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. Furthermore, the channel gains for different receive antennas are assumed to be independent and identically distributed (i.i.d.).

A. Channel Estimation

To enable the receiver to estimate the channel gains of all N antennas, the transmitter first transmits the pilot symbol p_p (of energy E and duration T_s) $\lceil N/L \rceil$ times, as illustrated in Fig. 1. Two consecutive pilot symbols are separated in time

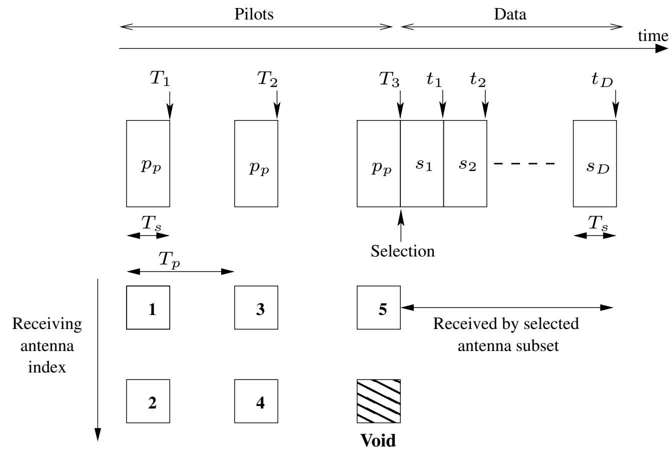

Fig. 1. Illustration of training for receive AS with $N = 5$ antennas and $L = 2$ RF chains. The illustration shows how $\lceil \frac{5}{2} \rceil = 3$ pilots are transmitted, and the sequence in which the receiver estimates channel gains.

by a duration T_p.[2] The pilots are followed by D data symbols, each of duration T_s and average energy E.

The pilot symbol received by the k^{th} receive antenna at time $T_{\lceil k/L \rceil}$ is

$$r_k\left(T_{\lceil k/L \rceil}\right) = p_p h_k\left(T_{\lceil k/L \rceil}\right) + n_k\left(T_{\lceil k/L \rceil}\right), \quad (1)$$

where the noise $n_k(t)$ is an additive white circularly symmetric complex Gaussian noise process, which is independent of $h_k(t)$, and has a variance of N_0. Based on $r_k\left(T_{\lceil k/L \rceil}\right)$, the channel estimate at the k^{th} receive antenna is

$$\hat{h}_k\left(T_{\lceil k/L \rceil}\right) = \frac{r_k\left(T_{\lceil k/L \rceil}\right)}{p_p} = h_k\left(T_{\lceil k/L \rceil}\right) + e_k, \quad (2)$$

where $e_k = \frac{n_k\left(T_{\lceil k/L \rceil}\right)}{p_p}$ is the noise-induced channel estimation error with variance $\sigma_e^2 = \frac{N_0}{E} \triangleq \gamma^{-1}$. ($\gamma$ is the average signal-to-noise-ratio (SNR) per receive antenna.)

B. Outdated channel estimates

Due to the time-varying nature of the wireless links, the N channels will have changed by the time data transmission starts. Specifically, the channel for receive antenna i at time $t + \delta$ can be written in terms of its channel at time t as [12], [13][3]

$$h_i(t+\delta) = \rho_i(\delta) h_i(t) + \sqrt{1 - |\rho_i(\delta)|^2}\, n_i'(t+\delta), \quad (3)$$

where $\rho_i(\delta)$ is the channel correlation coefficient. For all $i = 1, 2, \ldots, N$, the variation $n_i'(t+\delta)$ is a circularly symmetric

[2] In order to reduce overhead IEEE 802.11n usually transmits payload data together with each pilot tone. These payload data *have* to be received with the same antenna for which the pilot tone is intended. Thus, the error probability for the payload does not depend on any selection algorithm, and is irrelevant for the purposes of this paper. Note that T_p also captures the inter-packet spacing required to switch between antennas. However, the switching time is small compared to a packet duration.

[3] The model in (3) assumes that the channel realizations at different times can be computed based only on the correlation with the channel at time $t = 0$, and not as a realization of a stochastic process with a continuous correlation function. This approximation is good so long as the maximum Doppler frequency times δ is small, and shall be verified in Sec. V.

complex Gaussian RV with unit variance that is independent of $h_i(t)$. The channel correlation coefficient depends on the time difference δ and the Doppler spectrum. This accounts for channel variations during training as well as data transmission.

C. Data Reception

Let $\widehat{\Omega}_L$ denote the antenna subset selected for data reception based on the estimates $\hat{h}_k\left(T_{\lceil k/L \rceil}\right)$, for $1 \le k \le N$. When the i^{th} data symbol, s_i, is transmitted, the signal received by the antenna subset $\widehat{\Omega}_L$ at time t_i, after matched filtering, is

$$y_k(t_i) = h_k(t_i) s_i + n_k(t_i), \quad k \in \widehat{\Omega}_L. \tag{4}$$

The data symbols are drawn from the MPSK constellation with equal probability.

III. Optimal Subset Selection and SEP Analysis

Henceforth, we simplify our notation as follows: we denote $\hat{h}_k\left(T_{\lceil k/L \rceil}\right)$ by $\hat{h}_k$, $n'_k(t)$ by n'_k, $n_k(t_i)$ by n_k, and $y_k(t_i)$ by y_k. The expectation and variance of a RV A are denoted by $\mathbf{E}[A]$ and $\mathbf{var}[A]$, respectively. Similarly, $\mathbf{E}[A|B]$ and $\mathbf{var}[A|B]$ denote the conditional expectation and variance, respectively, of A given B. The probability of event A and the probability of A given B are denoted by $\Pr(A)$ and $\Pr(A|B)$, respectively. x^* denotes the complex conjugate of x, and $\{x_i\}_{i=1}^N$ denotes the set $\{x_1, \ldots, x_N\}$.

A. Decision Variable and its Statistics

Conditioned on $\hat{h}_k$ and s_i, y_k is a complex Gaussian RV with conditional mean and variance given by

$$\mathbf{E}\left[y_k \,\middle|\, \hat{h}_k, s_i\right] = s_i \hat{h}_k \rho_k^{(i)} (1+\gamma^{-1})^{-1}, \tag{5}$$

$$\mathbf{var}\left[y_k \,\middle|\, \hat{h}_k, s_i\right] = E\left(1+\gamma^{-1} - \left|\rho_k^{(i)}\right|^2 (1+\gamma^{-1})^{-1}\right), \tag{6}$$

where $\rho_k^{(i)}$ is the correlation between $h_k(t_i)$ and $h_k\left(T_{\lceil k/L \rceil}\right)$.

The decision variable, $\mathcal{D}$, at a maximum likelihood receiver for detecting the data received at time t_i is based on the observables y_k and $\hat{h}_k$, for $1 \le k \le N$. Accounting for the dependence of the variance of y_k on k (as per (6)), $\mathcal{D}$ can be shown to be:

$$\mathcal{D} = \sum_{k \in \widehat{\Omega}_L} \frac{\left(\hat{h}_k \rho_k^{(i)}\right)^* y_k}{1+\gamma^{-1} - \left|\rho_k^{(i)}\right|^2 (1+\gamma^{-1})^{-1}}. \tag{7}$$

Making use of the above results, we now derive the optimal subset selection rule.

B. Optimal Subset Selection Rule

Theorem *1:* Let

$$w_{k,i}^{\text{opt}} = \frac{\left|\rho_k^{(i)}\right|^2}{(1+\gamma^{-1})^2 - \left|\rho_k^{(i)}\right|^2}, \tag{8}$$

Let the weighted estimates $w_{k,i}^{\text{opt}} \left|\hat{h}_k\right|^2$, for $1 \le k \le N$, be sorted in descending order as

$$w_{[\hat{1}],i}^{\text{opt}} \left|\hat{h}_{[\hat{1}]}\right|^2 > w_{[\hat{2}],i}^{\text{opt}} \left|\hat{h}_{[\hat{2}]}\right|^2 > \cdots > w_{[\hat{N}],i}^{\text{opt}} \left|\hat{h}_{[\hat{N}]}\right|^2, \tag{9}$$

where $[\hat{k}]$ denotes the index of the antenna with the k^{th} largest weighted channel estimate. Then, the optimal antenna subset that minimizes the SEP of the MPSK symbol received at time t_i is given by[4]

$$\widehat{\Omega}_L = \left\{[\hat{1}], [\hat{2}], \cdots, [\hat{L}]\right\}. \tag{10}$$

Proof: The proof is given in Appendix A. ■

The above result is intuitively satisfying because the weight for an antenna's channel estimate depends on how outdated it is. The smaller the correlation $\rho_k^{(i)}$, the smaller the weight $w_{k,i}^{\text{opt}}$.

C. SEP of MPSK with Optimal Subset Selection

Theorem *2:* With outdated and noisy channel estimates and optimal weights, the SEP of an MPSK symbol transmitted at time t_i, is given by

$$\begin{aligned} P_i^{\text{MPSK}}(\gamma) = \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_N} \Bigg(\Bigg(\prod_{k=1}^{N} \left[\sum_{m=1}^{k} \frac{\Gamma_{\lambda(k)}}{\Gamma_{\lambda(m)}} \right]^{-1} \Bigg) \\ \times \sum_{n=1}^{N} \Bigg(\prod_{\substack{l=1 \\ l \neq n}}^{N} \frac{f_n(\lambda)}{f_n(\lambda) - f_l(\lambda)} \Bigg) \Bigg[\frac{M-1}{M}\pi \\ - \sqrt{\frac{f_n(\lambda)}{1+f_n(\lambda)}} \tan^{-1}\Bigg(\sqrt{\frac{1+f_n(\lambda)}{f_n(\lambda)}} \tan\left(\frac{M-1}{M}\pi\right) \Bigg) \Bigg] \Bigg), \end{aligned} \tag{11}$$

where $\mathcal{P}_N$ is the set of all permutations of $\{1, 2, \ldots, N\}$ and the summation is over all permutations $\lambda \in \mathcal{P}_N$. For $n \le L$, $f_n(\lambda) = n\left[\sum_{p=1}^{n} \frac{1}{\Gamma_{\lambda(p)}}\right]^{-1}$ and, for $n > L$, $f_n(\lambda) = L\left[\sum_{p=1}^{n} \frac{1}{\Gamma_{\lambda(p)}}\right]^{-1}$, and $\Gamma_k = \frac{\left|\rho_k^{(i)}\right|^2}{(1+\gamma^{-1})^2 - \left|\rho_k^{(i)}\right|^2} \sin^2\left(\frac{\pi}{M}\right)$.

Proof: The proof is given in Appendix B. ■

IV. Optimal Subset Selection with Void-Filling

As mentioned, when L does not divide N, some of the RF chains are left unused when receiving the last pilot. In a system with N receive antennas and L RF chains, the number of voids (unused RF chains) that need to be filled is $\upsilon = N - L\left\lfloor \frac{N}{L} \right\rfloor$. Without loss of generality, let antennas $1, \ldots, L$ get estimated using the first pilot, antennas $L+1, \ldots, 2L$ get estimated using the second pilot, and so on. The specific order in which antennas are trained does not matter since the channel gains of different antennas are assumed to be i.i.d.

[4] We introduce a '$\hat{\cdot}$' in $[\hat{k}]$ in order to emphasize the fact that subset selection is done using noisy channel estimates and is not perfect.

A. Optimal Subset Selection Given a Specific Void-Filling Rule

Let $\mathcal{S}_\upsilon$ denote the υ-element subset of receive antennas chosen to fill the training voids that occur when receiving the last pilot symbol at time $T_{\lceil N/L \rceil}$. For these antennas, in addition to an observation at $T_{\lceil k/L \rceil}$ (given by (1)), the receiver also makes another observation at time $T_{\lceil N/L \rceil}$, which equals

$$r_k\left(T_{\lceil N/L \rceil}\right) = p_p h_k\left(T_{\lceil N/L \rceil}\right) + n_k\left(T_{\lceil N/L \rceil}\right), k \in \mathcal{S}_\upsilon. \quad (12)$$

The correlation between the two observations can be exploited to refine the channel estimates for these antennas and, hence, to improve the subset selection accuracy. Let ξ_k denote the correlation between $h_k(T_{\lceil N/L \rceil})$ and $h_k\left(T_{\lceil k/L \rceil}\right)$. Then, the refined MMSE estimate of the k^{th} fading link at time $T_{\lceil N/L \rceil}$, denoted by $\hat{\hat{h}}_k$, is given by

$$\hat{\hat{h}}_k = \frac{p_p^*}{E} r_k(T_{\lceil N/L \rceil}) \frac{1+\gamma^{-1} - |\xi_k|^2}{(1+\gamma^{-1})^2 - |\xi_k|^2} + \frac{p_p^*}{E} r_k\left(T_{\lceil k/L \rceil}\right) \frac{\xi_k^* \gamma^{-1}}{(1+\gamma^{-1})^2 - |\xi_k|^2}. \quad (13)$$

Therefore, $\hat{\hat{h}}_k$ is a circular symmetric complex Gaussian RV with variance given by $\mathbf{var}\left[\hat{\hat{h}}_k\right] \triangleq \sigma_k^2 = \frac{(1+\gamma^{-1})^3 - |\xi_k|^2 (1+\gamma^{-1})(2-\gamma^{-2}) + |\xi_k|^4 (1-\gamma^{-1})}{\left((1+\gamma^{-1})^2 - |\xi_k|^2\right)^2}$. We will now show that, given $\mathcal{S}_\upsilon$, the optimal subset selection rule again linearly weights the channel estimates, but the optimal weights are different from those in (8). This is because some of the antennas have refined estimates.

Theorem *3:* When voids are filled with antenna subset $\mathcal{S}_\upsilon$, the optimal antenna subset that minimizes the SEP of the i^{th} MPSK symbol, received at time t_i, is given by

$$\Omega_L' = \left\{[\hat{l}]\right\}_{l=1}^{L}, \quad (14)$$

where $[\hat{i}]$ is the index of the antenna with the i^{th} largest entry in the following ordering of weighted channel estimates:

$$w'_{[\hat{1}],i} \left|\hat{h}'_{[\hat{1}]}\right|^2 > w'_{[\hat{2}],i} \left|\hat{h}'_{[\hat{2}]}\right|^2 > \cdots > w'_{[\hat{N}],i} \left|\hat{h}'_{[\hat{N}]}\right|^2. \quad (15)$$

Here, $w'_{k,i} = \frac{\left|\zeta_k^{(i)}\right|^2}{{\sigma'_k}^2 \left({\sigma'_k}^2 (1+\gamma^{-1}) - \left|\zeta_k^{(i)}\right|^2\right)}$,

$$\hat{h}'_k = \begin{cases} \hat{h}_k, & \text{if } k \notin \mathcal{S}_\upsilon \\ \hat{\hat{h}}_k, & \text{if } k \in \mathcal{S}_\upsilon \end{cases}, \quad (16)$$

$$\zeta_k^{(i)} = \begin{cases} \rho_k^{(i)}, & \text{if } k \notin \mathcal{S}_\upsilon \\ \frac{\rho_N^{(i)} (1 - |\xi_k|^2 + \gamma^{-1}) + \rho_k^{(i)} \xi_k^* \gamma^{-1}}{(1+\gamma^{-1})^2 - |\xi_k|^2}, & \text{if } k \in \mathcal{S}_\upsilon \end{cases}, \quad (17)$$

$${\sigma'_k}^2 = \begin{cases} 1+\gamma^{-1}, & \text{if } k \notin \mathcal{S}_\upsilon \\ \sigma_k^2, & \text{if } k \in \mathcal{S}_\upsilon \end{cases}. \quad (18)$$

Proof: The proof is given in Appendix C. ■

It can be clearly seen that the parameters $\hat{h}'_k$, $\zeta_k^{(i)}$, and ${\sigma'_k}^2$, and consequently the corresponding linear weights, are updated only for antennas whose channels are estimated twice.

B. Random Void-Filling Rule

We propose below a simple void-filling rule, in which the receiver randomly picks υ distinct antennas from the set $\left\{1, 2, \ldots, L\left\lfloor \frac{N}{L} \right\rfloor\right\}$. Note that this set excludes antennas that have an RF chain allocated to them during the last pilot's transmission. Its optimal subset selection rule is derived in Theorem 3. Its SEP can be derived along the lines of Theorem 2, and is given in [14]. Thus, random void-filling is an analytically tractable rule and serves as a good benchmark to compare against. In general, more sophisticated criteria to fill voids can be developed that exploit the channel estimates of the antennas that have already been estimated in the previous $\lceil \frac{N}{L} \rceil - 1$ pilots. These are developed in [14]. One such rule linearly weights the channel estimates based on the outdatedness of the CSI.

V. Simulations

We now present graphically the results derived in Sec. III and Sec. IV, and study the effect of the number of receive RF chains (L) on the SEP. While our analysis is valid for arbitrary Doppler spectra, we use the classical Jakes spectrum in our simulations. Therefore, the correlation values for $k = 1, \ldots, N$, $i = 1, 2, \ldots, D$, equal $\rho_k^{(i)} = J_0\left(2\pi f_d \left(\left(\lceil \frac{N}{L} \rceil - \lceil \frac{k}{L} \rceil\right) T_p + i T_s\right)\right)$ and $\xi_k = J_0\left(2\pi f_d \left(\left(\lceil \frac{N}{L} \rceil - \lceil \frac{k}{L} \rceil\right) T_p\right)\right)$, where $J_0(.)$ is the zeroth order Bessel function of the first kind [15] and f_d is the maximum Doppler frequency. The figures are plotted for $T_p = 10T_s$ and for the first data symbol of the packet. The SEP behavior for subsequent data symbols is similar [14].

Figure 2 plots the SEP as a function of SNR for MPSK with $N = 4$ antennas and $L = 2$ RF chains. In this case, no voids occur since 2 is an integer divisor of 4. We see that while the SEP decreases to 0 as the SNR increases when $f_d T_p = 0$, an error floor exists when $f_d T_p > 0$. The error floor increases as $f_d T_p$ increases. We see that optimal weighting significantly outperforms no-weighting even when $f_d T_p$ is as small as 0.06. As expected, no-weighting is optimal only when $f_d T_p \approx 0$. Notice the excellent match between analytical and simulation results for both the weighting schemes.

Figure 3 compares the SEP of the random void-filling rule with the case where the voids are not filled. The SEP is plotted for 8PSK and $N = 5$ with $L = 3$ and $L = 4$. It brings out an interesting scenario where, at high SNR, the SEP for $L = 4$ is worse than $L = 3$ when the single void is not filled. This is because at high SNR, how outdated the estimates are affects SEP the most. With $L = 4$, only one out of the 5 antennas gets a fresh (less outdated) estimate from the last (second) pilot and three RF chains are not used for receiving the second pilot symbol. Where as, with $L = 3$, two antennas get a fresh (less outdated) estimate from the second (last) pilot and one RF chain is not used to receive the second pilot symbol. With void filling, L antennas are always trained in the last pilot transmission. Therefore, the number of fresh estimates is L, and the SEP monotonically decreases as L increases for all SNR.

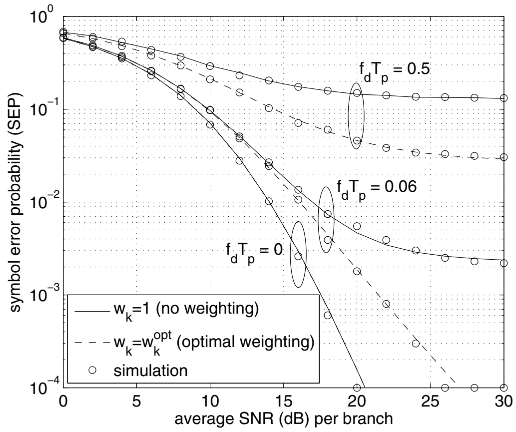

Fig. 2. Effect of normalized Doppler spread and selection weights (8PSK, $T_p = 10T_s$, $L = 2$, and $N = 4$).

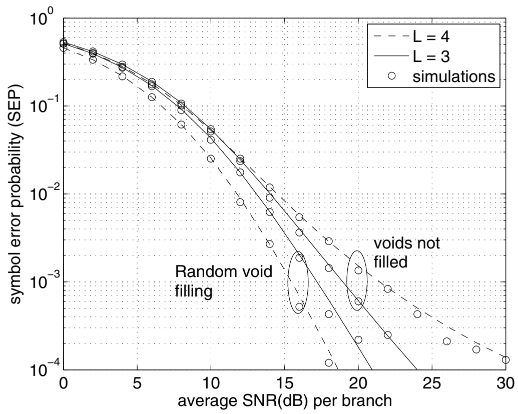

Fig. 3. Effect of number of receive RF chains, L (8PSK, $f_dT_p = 0.06$, $T_p = 10T_s$, and $N = 5$).

VI. Conclusions

We investigated training and selection criteria for receive antenna subset selection in time-varying channels given the practical constraints imposed by next generation wireless standards. In a practical training scheme for AS, the training delays are considerable and the channel estimates of different antennas are outdated by different amounts. We saw that this changes the criterion used to select the antennas subset and significantly affects the overall SEP. We showed that, in order to minimize the SEP, the optimal subset of antennas should be selected on the basis of a linear weighting of the channel estimates. We developed closed-form expressions for the weights and the SEP with optimal subset selection for any L.

We discovered the problem of training voids, which occurs when L does not divide N as some of the RF chains remained unused in receiving the last pilot. We showed that filling the voids, *i.e.*, using the unused RF chains to refine the estimates already obtained for some of the antennas, significantly improves the performance. We showed that the optimal subset selection rule still uses a linear weighting of the refined channel estimates, albeit with different weights. We saw that filling the voids by randomly chosen antennas improves performance and serves as a useful benchmark for void-filling schemes, in general. We even saw that when the number of RF chains increases the SEP can increase without void-filling, and that random void-filling resolves this anomaly.

Appendix

A. Proof of Theorem 1

Since $\mathcal{D}$, conditioned on $\widehat{\Omega}_L$ and $\left\{\hat{h}_l\right\}_{l\in\widehat{\Omega}_L}$, is a Gaussian RV, the standard SEP expression for MPSK is [16, (40)]

$$P_i\left(\widehat{\Omega}_L, \left\{\hat{h}_l\right\}_{l\in\widehat{\Omega}_L}, s_i\right) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \exp\left(\frac{-\left|\mu_{\mathcal{D}}\right|^2 \sin^2\left(\frac{\pi}{M}\right)}{\sigma_{\mathcal{D}}^2 \sin^2\theta}\right) d\theta. \tag{19}$$

Using (5) and (6) and standard results on moments of conditional Gaussian RVs, the conditional first and second moments of $\mathcal{D}$ required in (19) can be shown to be related as

$$\frac{\mu_{\mathcal{D}}}{s_i(1+\gamma^{-1})^{-1}} = \frac{\sigma_{\mathcal{D}}^2}{E} = \sum_{k\in\widehat{\Omega}_L} \frac{\left|\hat{h}_k\right|^2 \left|\rho_k^{(i)}\right|^2}{1+\gamma^{-1} - \left|\rho_k^{(i)}\right|^2 (1+\gamma^{-1})^{-1}}.$$

Substituting the above result in (19) yields

$$P_i\left(\widehat{\Omega}_L, \left\{\hat{h}_l\right\}_{l\in\widehat{\Omega}_L}\right) = \frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi} \exp\left(\sum_{k\in\widehat{\Omega}_L} \frac{-\Gamma_k\left|\hat{h}_k\right|^2}{(1+\gamma^{-1})\sin^2\theta}\right) d\theta, \tag{20}$$

where Γ_k is defined in Theorem 2. From the above equation, it is clear that the optimal antenna subset, $\widehat{\Omega}_L$, that minimizes the SEP of MPSK should maximize $\sum_{k\in\widehat{\Omega}_L} \Gamma_k \left|\hat{h}_k\right|^2$. Hence, the result follows.

B. Proof of Theorem 2

To conserve space, we outline the key steps below. A detailed derivation is given in [14]. Let $X_k \triangleq \Gamma_k \left|\hat{h}_k\right|^2 (1+\gamma^{-1})^{-1}$. It is an exponential RV with mean $\mathbf{E}[X_k] \triangleq \Gamma_k$. Furthermore, $X_1, \ldots, X_N$ are mutually independent (but non-identical). When sorted in decreasing order, we get $X_{[1]} > X_{[2]} > \cdots > X_{[N]}$. Let $Y \triangleq \sum_{k=1}^{L} X_{[k]}$.

For a given permutation λ of the set $\{1, 2, \cdots, N\}$, let A_λ denote the event that it gives the sorted order above, *i.e.*, $[1] = \lambda(1), [2] = \lambda(2), \cdots, [N] = \lambda(N)$. Since the X_is are independent exponential RVs, the probability of A_λ can be shown to be equal to [17]

$$\Pr(A_\lambda) = \prod_{k=1}^{N} \left[\sum_{m=1}^{k} \frac{\Gamma_{\lambda(k)}}{\Gamma_{\lambda(m)}}\right]^{-1}. \tag{21}$$

Furthermore, we also have [17]

$$\mathcal{M}_{Y|A_\sigma}(y|A_\lambda) = \prod_{n=1}^{N} \frac{1}{1+y f_n(\lambda)}. \tag{22}$$

Therefore, the SEP expression in (20), when averaged over fading, can be written as

$$P_i^{\mathrm{MPSK}}(\gamma)=\frac{1}{\pi}\sum_{\lambda\in\mathcal{P}_N}\Pr\left(A_\lambda\right)\int_0^{\frac{M-1}{M}\pi}\mathcal{M}_{Y|A_\lambda}\left(\frac{1}{\sin^2\theta}\right)d\theta, \tag{23}$$

where $\mathcal{P}_N$ is the set of all permutations of $\{1,2,\ldots,N\}$. where f_n is as defined in the theorem statement. Substituting (21) and (22) in (23), we get

$$P_i^{\mathrm{MPSK}}(\gamma)=\frac{1}{\pi}\sum_{\lambda\in\mathcal{P}_N}\left(\prod_{k=1}^{N}\left[\sum_{m=1}^{k}\frac{\Gamma_{\lambda(k)}}{\Gamma_{\lambda(m)}}\right]^{-1}\right)\times\int_0^{\frac{M-1}{M}\pi}\prod_{n=1}^{N}\frac{\sin^2\theta}{\sin^2\theta+f_n(\lambda)}\,d\theta. \tag{24}$$

The partial fraction expansion of the integrand in (24) is

$$\prod_{n=1}^{N}\frac{\sin^2\theta}{\sin^2\theta+f_n(\lambda)}=\sum_{n=1}^{N}\left[\prod_{\substack{k=1\\k\neq n}}^{N}\frac{f_n(\lambda)}{f_n(\lambda)-f_k(\lambda)}\right]\frac{\sin^2\theta}{\sin^2\theta+f_n(\lambda)}.$$

Furthermore, the integral of each term above can be written in closed-form using the following identity [15, (2.562)]:

$$\int_0^{\frac{M-1}{M}\pi}\frac{\sin^2\theta}{\sin^2\theta+f_n(\lambda)}d\theta=\frac{M-1}{M}\pi-\sqrt{\frac{f_n(\lambda)}{1+f_n(\lambda)}}\tan^{-1}\left(\sqrt{\frac{1+f_n(\lambda)}{f_n(\lambda)}}\tan\left(\frac{M-1}{M}\pi\right)\right). \tag{25}$$

Substituting (25) in (24) yields the desired result.

C. Proof of Theorem 3

The proof is along the lines of Theorem 1. Therefore, we outline only the key steps to conserve space. Proceeding in a manner similar to (5) and (6), the new decision variable for maximum likelihood detection of s_i is given by

$$\mathcal{D}'=\sum_{k\in\Omega'_L}\frac{\left(\zeta_k^{(i)}\hat{h}'_k\right)^* y_k}{{\sigma'_k}^2-\left|\zeta_k^{(i)}\right|^2+\gamma^{-1}{\sigma'_k}^2}, \tag{26}$$

where $\hat{h}'_k$, $\zeta_k^{(i)}$, and ${\sigma'_k}^2$ are as given in (16), (17), and (18), respectively. Note that $\mathcal{D}'$ differs slightly from the decision variable $\mathcal{D}$ in (7) because the channel estimates of the receive antennas that filled voids have been refined.

Consequently, for $\mathcal{D}'$, its conditional mean, $\mu_{\mathcal{D}'}\triangleq\mathbf{E}\left[\mathcal{D}'\,\middle|\,\mathcal{S}_\upsilon,\Omega'_L,\left\{\hat{h}'_k\right\}_{k\in\Omega'_L},s_i\right]$, and its conditional variance, $\sigma^2_{\mathcal{D}'}\triangleq\mathbf{var}\left[\mathcal{D}'\,\middle|\,\mathcal{S}_\upsilon,\Omega'_L,\left\{\hat{h}'_k\right\}_{k\in\Omega'_L},s_i\right]$, are related as

$$\frac{\mu_{\mathcal{D}'}}{s_i}=\frac{\sigma^2_{\mathcal{D}'}}{E}=\sum_{k\in\Omega'_L}\frac{\left|\zeta_k^{(i)}\right|^2\left|\hat{h}'_k\right|^2}{{\sigma'_k}^2\left(\left(1+\gamma^{-1}\right){\sigma'_k}^2-\left|\zeta_k^{(i)}\right|^2\right)}.$$

Conditioned on $\mathcal{S}_\upsilon$, Ω'_L, and $\left\{\hat{h}'_k\right\}_{k\in\Omega'_L}$, the SEP of the MPSK symbol transmitted at time t_i is

$$P_i\left(\mathcal{S}_\upsilon,\Omega'_L,\left\{\hat{h}'_k\right\}_{k\in\Omega'_L}\right)=\frac{1}{\pi}\int_0^{\frac{M-1}{M}\pi}\exp\left(-\sum_{k\in\Omega'_L}\frac{\left|\zeta_k^{(i)}\right|^2\left|\hat{h}'_k\right|^2\sin^2\left(\frac{\pi}{M}\right)/\sin^2\theta}{{\sigma'_k}^2\left((1+\gamma^{-1}){\sigma'_k}^2-\left|\zeta_k^{(i)}\right|^2\right)}\right)d\theta.$$

From the above equation its clear that the optimal antenna subset, Ω'_L, that minimizes the SEP should maximize $\sum_{k\in\Omega'_L}w'_{k,i}\left|\hat{h}'_k\right|^2$, with $w'_{k,i}$, as defined in theorem statement. Hence, the desired result follows.

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] N. B. Mehta and A. F. Molisch, "Antenna selection in MIMO systems," in *MIMO System Technology for Wireless Communications* (G. Tsulos, ed.), ch. 6, CRC Press, 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] D. A. Gore and A. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Signal Process.*, vol. 50, pp. 2580–2588, Oct. 2002.

[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] K. Zhang and Z. Niu, "Adaptive receive antenna selection for orthogonal space-time block codes with channel estimation errors with antenna selection," in *Proc. Globecom*, pp. 3314–3318, 2005.

[7] S. Han and C. Yang, "Performance analysis of MRT and transmit antenna selection with feedback delay and channel estimation error," in *Proc. WCNC*, pp. 1135–1139, 2007.

[8] W. Xie, S. Liu, D. Yoon, and J.-W. Chong, "Impacts of Gaussian error and Doppler spread on the performance of MIMO systems with antenna selection," in *Proc. Int. Conf. Wireless. Commun., Netw., Mobile Comput.*, pp. 1–4, 2006.

[9] R. Annavajjala and L. B. Milstein, "Performance analysis of optimum and suboptimum selection diversity schemes on Rayleigh fading channels with imperfect channel estimates," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 1119–1130, May 2007.

[10] H. Zhang, A. Molisch, and J. Zhang, "Applying antenna selection in WLANs for achieving broadband multimedia communications," *IEEE Trans. Broadcast.*, vol. 52, pp. 475–482, Dec. 2006.

[11] 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.

[12] D. V. Marathe and S. Bhashyam, "Power control for multi-antenna Gaussian channels with delayed feedback," in *Proc. Asilomar Conf. on Signals, Syst., and Comput.*, pp. 1598–1602, 2005.

[13] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1650–1662, Jul. 2005.

[14] V. Kristem, N. B. Mehta, and A. F. Molisch, "Training and training voids in receive antenna subset selection in time-varying channels," *To be submitted to IEEE Trans. Wireless Commun.*, 2010.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, 4^{th} ed., 1980.

[16] M.-S. Alouini and A. 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.

[17] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining of diversity branches with unequal SNR in Rayleigh fading," *Proc. VTC*, 1999.