

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

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

Abstract—Receive antenna selection (AS) provides many benefits of multiple-antenna systems at drastically reduced hardware costs. In it, the receiver connects a dynamically selected subset of N available antennas to the L available RF chains. Due to the nature of AS, the channel estimates at different antennas, which are required to determine the best subset for data reception, are obtained from different transmissions of the pilot sequence. Consequently, they are outdated by different amounts in a time-varying channel. We show that a linear weighting of the estimates is necessary and 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 then present new “void-filling” methods that exploit these voids and greatly improve the performance of AS. The optimal subset selection rules with void-filling, in which different antennas turn out to have different numbers of estimates, are also explicitly characterized. Closed-form equations for the symbol error probability with and without void-filling are also developed.

Index Terms—Training, antenna selection, diversity methods, fading channels, estimation, error analysis, quadrature phase shift keying, Doppler.

I. INTRODUCTION

RECEIVE antenna selection (AS) is a popular technique to reduce the hardware costs of a multi-antenna receiver in a wireless system [1]–[7]. Given N antennas, it uses $L < N$ radio frequency (RF) chains that process signals from a dynamically selected subset comprising L antennas. The case of $L > 1$ has also been referred to in the literature as generalized diversity combining (GDC) or hybrid-selection/maximal ratio combining (H-S/MRC) [8]–[14]. Selection is advantageous since antennas are typically cheap, while the RF chains are expensive. Consequently, next generation wireless standards such as IEEE 802.11n, Long Term Evolution (LTE), and IEEE 802.16m Advanced WiMax have standardized or are standardizing AS at the transmitter or the receiver, or both.

While AS has received considerable attention in the literature, issues related to pilot-based training for AS, through

which the channel estimates are acquired, have received limited attention. For example, channel state information (CSI) at the receiver was assumed to be perfect in [3], [15]–[17]. In reality, the estimates are imperfect because of noise during estimation and also the time-varying nature of the wireless channel, which makes the estimates partially outdated when they are used for data demodulation. Since the estimates are used to select the antenna subset and also to coherently demodulate the data, imperfections in the estimates lead to inaccurate selection and incorrect demodulation, both of which increase the symbol error probability (SEP).

While hardware limitations motivate AS in the first place, they unfortunately also exacerbate the problem of outdated channel estimates in time-varying channels. Since only L receive antennas can be estimated at any time with L RF chains, the transmitter needs to transmit the pilots multiple times so that the receiver can estimate the channels of all the available antennas and then choose the antenna subset with the best channels. Further, in practice, there can be a significant time delay between successive transmissions. For example, in the IEEE 802.11n standard, a multiple access control (MAC) based training protocol is used for transmit and/or receive AS [18]. In it, ‘training cum data packets’ are successively transmitted to enable the receiver to estimate the channel gains of all its antennas. The pilots embedded in the physical layer header of each training packet help estimate the channel gains. Since the packets, which also carry a data payload, can be several milliseconds long, the training phase also takes several tens of milliseconds. This mechanism was adopted by the standard body as it minimized the changes in the physical layer protocol that were required to accommodate AS in the standard.

Even small training delays can significantly influence the choice of the subset of antennas. For example, selecting the L antennas that have the highest estimated channel gains, as has been done in [19], is no longer optimal. Intuitively, how outdated an antenna’s estimate is should also be taken into account while selecting the antenna. Receive AS with imperfect channel estimates was first explored in [19], [20], but for block fading channels. While [21]–[23] did consider outdated channel estimates, the fact that estimates at different antennas are outdated by different amounts was not accounted for. While [24] considered imperfect estimation and selection in time-varying channels, it only addressed the considerably simpler problem of selecting only one antenna ($L = 1$).

In this paper, 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. We derive the SEP-optimal rule for selecting the antenna subset

Manuscript received September 17, 2010; revised February 3, 2011; accepted February 16, 2011. The associate editor coordinating the review of this paper and approving it for publication was D. Dardari.

V. Kristem is with Broadcom Corporation, Bangalore, India (e-mail: vinod.kristem@gmail.com). He was with the Dept. of Electrical Communication Eng. (ECE), Indian Institute of Science (IISc), Bangalore, India, during the course of this work.

N. B. Mehta is with the ECE Dept. at IISc (e-mail: nbmehta@ece.iisc.ernet.in).

A. F. Molisch is with the Dept. of Electrical Eng., Univ. of Southern California, Los Angeles, CA, USA (e-mail: molisch@usc.edu).

A part of this paper has appeared in IEEE Globecom 2010.

Digital Object Identifier 10.1109/TWC.2011.032411.101649

as a function of the time correlation of the channel. We present general closed-form expressions for the SEP with arbitrary selection weights given imperfect channel estimation and selection in time-varying channels. These expressions are specialized to analyze the performance of the optimal weighted selection rule developed in this paper as well as the conventional no-weighting selection rule. These results significantly generalize those in [24], which considered the simpler case of $L = 1$.

The second major contribution of the paper lies in highlighting and addressing, for the first time, the problem of how to best allocate the available L RF chains when L is not an integer divisor of N . This scenario can easily happen in a practical implementation of AS since the choice of L is primarily driven by hardware cost considerations.

The issues that arise are best explained using Fig. 1, which illustrates the training process for AS with $N = 5$ antennas and $L = 2$ RF chains. To estimate the channels of all 5 antennas, at least 3 pilot transmissions are needed. The first pilot transmission helps estimate the channels of antennas #1 and #2 and the second one helps estimate the channels of antennas #3 and #4. Notably, for the third pilot, only one antenna (#5) remains to be estimated, which requires only one of the two available RF chains. We shall call the unused RF chain as a *training void*. Such voids do not occur when L divides N , as is trivially true for $L = 1$.

Instead of leaving the RF chains unused, we propose connecting them to some of the antennas, which we term *void-filling*. A key question that then arises is which antenna to connect to the unused RF chain: #1, #2, #3, or #4? This choice affects not just selection but also the subsequent data demodulation since the receiver will have two estimates – outdated by different amounts – of the channel gain for the antenna so connected, and only one estimate for the other antennas. In general, with N receive antennas and L RF chains, at least $\lceil \frac{N}{L} \rceil$ training symbols are required, and $L \lceil \frac{N}{L} \rceil - N$ RF chains are unused in the last pilot symbol, where $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ denote the ceil and floor functions, respectively.

Our SEP analysis shows that even random void-filling, which picks the antennas randomly to fill the voids, markedly improves performance. Choosing the antennas intelligently by taking into account the estimates that have already been obtained for some of the antennas improves performance even further. To this end, we develop a near-optimal void-filling criterion. To help understand it, we also compare its performance with two other intuitive, but sub-optimal, void-filling criteria. Of these, the first one, which has been mentioned earlier, picks the antennas randomly and provides analytical insights into the performance benefits from void-filling. Furthermore, given any void-filling criterion, we also derive the corresponding rule for determining the optimal antenna subset that should be selected to receive data. The results are then extended to handle the problem of receiving a packet, which can be coded or uncoded, such that all its symbols are received by the same antenna subset. To the best of our knowledge, this is the first paper that addresses the practical problem of training voids in receive AS.

The outline of this paper is as follows. The AS training and data transmission models are described in Sec. II and

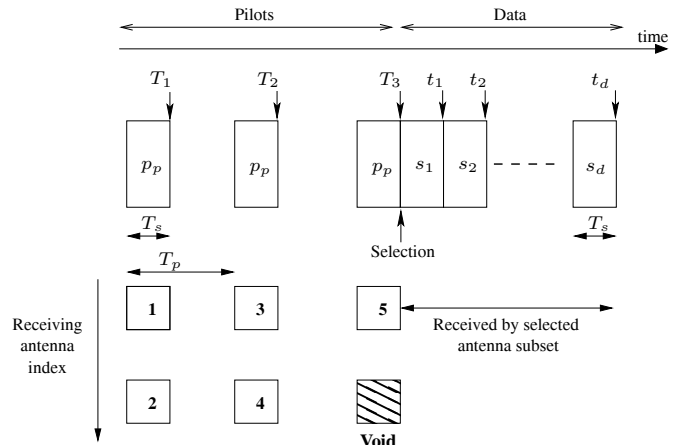


Fig. 1. Training for antenna selection with $N = 5$ antennas and $L = 2$ RF chains.

analyzed in Sec. III. Void-filling is investigated in Sec. IV. Simulation results and conclusions follow in Sec. V and Sec. VI, respectively. Several mathematical derivations are relegated to the Appendix.

II. SYSTEM MODEL

Consider a system with one transmit antenna, N receive antennas, and $1 \leq L < N$ receive RF chains. 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) – abbreviated henceforth as \mathcal{CN} – with unit variance. Furthermore, the channel gains for different receive antennas are assumed to be independent and identically distributed (i.i.d.), which happens when the antennas are at least half a wavelength apart in a rich scattering environment.

A. Channel Estimation

To enable the receiver to estimate the channel gains of all N antennas, the transmitter first sends $\lceil \frac{N}{L} \rceil$ pilots sequentially in time, as illustrated in Fig. 1. The receiver obtains channel estimates of L antennas from each pilot. Two consecutive pilot symbols are separated in time by a duration T_p .¹ The pilots are followed by d data symbols. Each pilot and data symbol has a duration T_s and an average energy E . As discussed, T_p can be much larger than T_s in practice. Since the channel gains are assumed to be i.i.d., the specific order in which the estimates are obtained for different antennas does not matter. We shall, therefore, assume without loss of generality (wlog) that the channel gains of antennas $1, \dots, L$ are estimated at time T_1 , of antennas $L + 1, \dots, 2L$ are estimated at time T_2 , and so on, where $T_2 - T_1 = T_3 - T_2 = \dots = T_{\lceil \frac{N}{L} \rceil} - T_{\lceil \frac{N}{L} \rceil - 1} = T_p$.

The k^{th} receive antenna's channel gain gets estimated from the pilot sent at time $T_{\lceil \frac{k}{L} \rceil}$, for which the received signal is given by

$$r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) = \sqrt{E} p h_k \left(T_{\lceil \frac{k}{L} \rceil} \right) + n_k \left(T_{\lceil \frac{k}{L} \rceil} \right), \quad (1)$$

¹Note that T_p also captures the inter-packet spacing required to switch between antennas. However, the switching time is typically small compared to a packet duration.

where $n_k(t)$ is \mathcal{CN} with variance N_0 and is independent of $h_k(t)$, and p is the pilot symbol; wlog $|p| = 1$. The minimum mean square error (MMSE) channel estimate for the k^{th} receive antenna at time t_i , the time when the i^{th} data symbol is received, can be shown to be

$$\begin{aligned}\widehat{h}_k(t_i) &= \mathbf{E} \left[h_k(t_i) \middle| r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) \right], \\ &= \frac{\sqrt{E} p^* \rho_k^{(i)}}{E + N_0} r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) = \frac{E \rho_k^{(i)}}{E + N_0} h_k \left(T_{\lceil \frac{k}{L} \rceil} \right) + e_k,\end{aligned}\quad (2)$$

where $\rho_k^{(i)} = \mathbf{E} \left[h_k(t_i) h_k \left(T_{\lceil \frac{k}{L} \rceil} \right)^* \right]$ is the correlation between $h_k(t_i)$ and $h_k \left(T_{\lceil \frac{k}{L} \rceil} \right)$ and $e_k = \frac{\sqrt{E} p^* \rho_k^{(i)}}{E + N_0} n_k \left(T_{\lceil \frac{k}{L} \rceil} \right)$ is \mathcal{CN} with variance $\sigma_e^2 = \frac{N_0 E |\rho_k^{(i)}|^2}{(E + N_0)^2}$. The channel correlation coefficient depends on the time difference $t_i - T_{\lceil \frac{k}{L} \rceil}$ and the Doppler spectrum [24].

B. Selection of Antenna Subset and Data Reception Using It

As we shall see, selecting L antennas with the highest estimated channel gains is not optimal when the channel estimates of different antennas are outdated by different amounts. Hence, we propose the following subset selection rule, in which the channel estimates are first weighted and then ordered:

$$w_{[\widehat{1}],i} \left| \widehat{h}_{[\widehat{1}]}(t_i) \right|^2 > \dots > w_{[\widehat{N}],i} \left| \widehat{h}_{[\widehat{N}]}(t_i) \right|^2, \quad (3)$$

where $[\widehat{k}]$ denotes the index of the antenna with k^{th} largest weighted estimated channel gain. *The selected L -element antenna subset*

$$\Omega_L = \left\{ [\widehat{1}], [\widehat{2}], \dots, [\widehat{L}] \right\} \quad (4)$$

is used for receiving data. We will show in Theorem 2 that linear weighting is indeed the SEP-optimal selection strategy.

The pilots are followed by d data symbols s_1, \dots, s_d . When the i^{th} data symbol, s_i , is transmitted at time t_i , the signal received by the antenna subset Ω_L after matched filtering is

$$y_k(t_i) = h_k(t_i) s_i + n_k(t_i), \quad k \in \Omega_L. \quad (5)$$

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

III. SEP ANALYSIS WITH OPTIMAL SUBSET SELECTION

In our analysis, we shall use extensively some standard results on conditional Gaussian RVs and on MMSE estimation. We first summarize our notation and these standard results below.

A. Notation and Standard Gaussian RV Results

For brevity, we denote $\widehat{h}_k(t_i)$ by \widehat{h}_k , $n_k(t_i)$ by n_k , and $y_k(t_i)$ by y_k . x^* denotes the complex conjugate of x , $\{x_i\}_{i=1}^N$ denotes the set $\{x_1, \dots, x_N\}$, and \mathcal{S}^c denotes the complement of a set \mathcal{S} . \mathcal{P}_n is the set of all permutations of the set $\{1, \dots, n\}$. And, $\mathcal{C}(l, n)$ is the set of all l element subsets of the set $\{1, \dots, n\}$. The Hermitian transpose and transpose will be denoted by $(\cdot)^\dagger$ and $(\cdot)^T$, respectively. The probability of an event A and the conditional probability of A given B are

denoted by $\Pr(A)$ and $\Pr(A|B)$, respectively. For an RV X , its probability density function (PDF) is denoted by $p_X(x)$, and $\mathbf{E}[X]$ and $\mathbf{var}[X]$ shall denote its expectation and variance, respectively. The moment generating function (MGF) of X is denoted by $\mathcal{M}_X(s)$ and equals $\int_0^\infty p_X(x) \exp(-sx) dx$. Similarly, $p_{X|A}(x)$, $\mathbf{E}[X|A]$, $\mathbf{var}[X|A]$, and $\mathcal{M}_{X|A}(s)$ shall denote the conditional PDF, expectation, variance, and MGF, respectively, of X given A .

Useful Results: If X and $\mathbf{Y} = [Y_1, \dots, Y_M]^T$ are zero-mean jointly Gaussian, then

$$\mathbf{E}[X|\mathbf{Y}] = \mathbf{E}[X\mathbf{Y}^\dagger] \mathbf{E}[\mathbf{Y}\mathbf{Y}^\dagger]^{-1} \mathbf{Y}, \quad (6)$$

$$\mathbf{var}[X|\mathbf{Y}] = \mathbf{var}[X] - \mathbf{E}[X\mathbf{Y}^\dagger] \mathbf{E}[\mathbf{Y}\mathbf{Y}^\dagger]^{-1} \mathbf{E}[\mathbf{Y}X^\dagger]. \quad (7)$$

The MMSE estimate of X , denoted by \widehat{X} , given the observation \mathbf{Y} is $\widehat{X} = \mathbf{E}[X|\mathbf{Y}]$. It satisfies

$$\mathbf{E}[X|\widehat{X}] = \widehat{X}, \quad (8)$$

$$\mathbf{E}[X\widehat{X}^*] = \mathbf{var}[\widehat{X}], \quad (9)$$

$$\mathbf{var}[X|\widehat{X}] = \mathbf{var}[X] - \mathbf{var}[\widehat{X}], \quad (10)$$

$$\mathbf{var}[\widehat{X}] = \mathbf{E}[X\mathbf{Y}^\dagger] \mathbf{E}[\mathbf{Y}\mathbf{Y}^\dagger]^{-1} \mathbf{E}[\mathbf{Y}X^\dagger]. \quad (11)$$

We now analyze the SEP for an MPSK symbol received at time t_i for receive AS with imperfect and outdated CSI. The analysis can be generalized to MQAM, as well.

B. Decision Variable and its Statistics

Conditioned on \widehat{h}_k and s_i , it follows from (5) that y_k is a complex Gaussian RV. From (8) and (10), its conditional mean and variance are given by

$$\mathbf{E}[y_k | \widehat{h}_k, s_i] = s_i \widehat{h}_k, \quad (12)$$

$$\mathbf{var}[y_k | \widehat{h}_k, s_i] = E \left(1 + \gamma^{-1} - \left| \rho_k^{(i)} \right|^2 (1 + \gamma^{-1})^{-1} \right). \quad (13)$$

Here, $\gamma \triangleq \frac{E}{N_0}$ is the average signal-to-noise-ratio (SNR) per receive antenna. The derivation is relegated to Appendix A. Note that the variance of \widehat{h}_k is different for antennas that are estimated at different times.

The maximum likelihood (ML) decision variable, \mathcal{D}_i , for detecting s_i received at time t_i is based on the observables y_k and \widehat{h}_k , for $1 \leq k \leq N$. Accounting for the different variances of signals received by different antennas, the decision variable \mathcal{D}_i , which is based on maximum ratio combining (MRC), is

$$\mathcal{D}_i = \sum_{k \in \Omega_L} \frac{\widehat{h}_k^* y_k}{(1 + \gamma^{-1})^2 - \left| \rho_k^{(i)} \right|^2}. \quad (14)$$

Therefore, \mathcal{D}_i conditioned on Ω_L , $\{\widehat{h}_k, k \in \Omega_L\}$, and s_i is a complex Gaussian RV. As shown in Appendix B, its conditional mean ($\mu_{\mathcal{D}_i}$) and conditional variance ($\sigma_{\mathcal{D}_i}^2$), are related by

$$\frac{\mu_{\mathcal{D}_i}}{s_i} = \frac{\sigma_{\mathcal{D}_i}^2}{E(1 + \gamma^{-1})^{-1}} = \sum_{k \in \Omega_L} \frac{\left| \widehat{h}_k \right|^2}{(1 + \gamma^{-1})^2 - \left| \rho_k^{(i)} \right|^2}. \quad (15)$$

C. General SEP Expression with Arbitrary Selection Weights

Theorem 1: With outdated and noisy channel estimates, the SEP of the i^{th} MPSK symbol (transmitted at time t_i), when the selection weight for antenna k is $w_{k,i}$, is given by

$$\begin{aligned} \text{SEP}_i(\gamma) &= \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_N} \left[\prod_{k=1}^N \left(\sum_{m=1}^k \frac{\alpha_{\lambda(k)}}{\alpha_{\lambda(m)}} \right)^{-1} \right] \sum_{n=1}^N \left[\prod_{\substack{l=1 \\ l \neq n}}^N \frac{f_n(\lambda)}{f_n(\lambda) - f_l(\lambda)} \right] \\ &\times \left(\frac{M-1}{M} \pi + \sqrt{\frac{f_n(\lambda)}{1+f_n(\lambda)}} \tan^{-1} \left(\sqrt{\frac{1+f_n(\lambda)}{f_n(\lambda)}} \tan \left(\frac{\pi}{M} \right) \right) \right), \end{aligned} \quad (16)$$

where the summation is over all permutations $\lambda \in \mathcal{P}_N$,

$$f_n(\lambda) = \begin{cases} \left(\sum_{m=1}^n \beta_{\lambda(m)} \right) \left[\sum_{p=1}^n \frac{1}{\alpha_{\lambda(p)}} \right]^{-1}, & \text{if } n \leq L \\ \left(\sum_{m=1}^L \beta_{\lambda(m)} \right) \left[\sum_{p=1}^n \frac{1}{\alpha_{\lambda(p)}} \right]^{-1}, & \text{if } n > L \end{cases},$$

$$\beta_k = \frac{(1+\gamma^{-1}) \sin^2 \left(\frac{\pi}{M} \right)}{(1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2} \frac{1}{w_{k,i}}, \text{ and } \alpha_k = \frac{w_{k,i} |\rho_k^{(i)}|^2}{1+\gamma^{-1}}.$$

Proof: The proof is given in Appendix C. ■

Note that the summation involves $N!$ terms and gets increasingly more computationally intensive as N increases. However, only few of these terms contribute significantly to the SEP. Intuitively, the permutations that are aligned with the order of weights are more likely and will contribute more. Hence, the SEP expression can be approximated by fewer terms. However, a systematic procedure to identify the most relevant terms is beyond the scope of this paper.

D. SEP-optimal Subset Selection Rule

We now characterize the optimal weights that minimize the SEP expression derived above.

Theorem 2: Linearly weighting the estimated channel gains, as done in (3), is the SEP-optimal strategy. The optimal selection weights that minimize the SEP of the i^{th} MPSK symbol are

$$w_{k,i}^{\text{opt}}(\gamma) = \frac{1}{(1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2}, \quad 1 \leq k \leq N. \quad (17)$$

Proof: The proof is given in Appendix D. ■

The above result generalizes the optimal selection rule and the SEP expression derived in [24] for single receive antenna selection. Interestingly, the optimal weights do not depend on the subset size, L . The smaller the correlation $\rho_k^{(i)}$, the smaller the weight $w_{k,i}^{\text{opt}}$, which makes intuitive sense. Further, the optimal selection weights – and, hence, the selected antenna – depend on i and can be different for data symbols transmitted at different times. Interestingly, using MMSE channel prediction to handle the time variations leads to $w_{k,i} = 1$. Thus, Theorem 2 also shows that the conventional approach of selecting the antenna subset with the L largest $|\widehat{h}_k|^2$, $1 \leq k \leq N$, is not SEP-optimal even though it uses the MMSE predicted values of the channel gains.

IV. VOID-FILLING: OPTIMAL SUBSET SELECTION AND VOID-FILLING CRITERIA

As mentioned, when L does not divide N , $v = L \lceil \frac{N}{L} \rceil - N$ RF chains are left unused when receiving the last pilot (sent at time $T_{\lceil \frac{N}{L} \rceil}$). Connecting these unused RF chains to some of the receive antennas gives two correlated observations, which can be jointly used to obtain a refined estimate of their time-varying channel gain. Doing so reduces SEP. The following two important questions then naturally arise: (i) which v antennas should be chosen to fill the voids, and (ii) once the voids are filled and estimates are obtained, which L antenna subset to select for receiving the data. We will address these issues in this section.

Let \mathcal{S}_v denote the v -element subset of receive antennas that is chosen to fill the training voids. We first derive the optimal subset selection rule given \mathcal{S}_v , and show that it is different from that in Sec. III-D. This will help us then determine the optimal \mathcal{S}_v itself.

A. Optimal Subset Selection Given a Void-Filling Criterion

For an antenna $k \in \mathcal{S}_v$, the receiver makes two observations – one at time $T_{\lceil \frac{k}{L} \rceil}$ (given by (1)) and the other at time $T_{\lceil \frac{N}{L} \rceil}$. The latter equals

$$r_k \left(T_{\lceil \frac{N}{L} \rceil} \right) = \sqrt{E} p h_k \left(T_{\lceil \frac{N}{L} \rceil} \right) + n_k \left(T_{\lceil \frac{N}{L} \rceil} \right), \quad k \in \mathcal{S}_v. \quad (18)$$

Lemma 1: For $k \in \mathcal{S}_v$, the refined MMSE estimate, denoted by $\widehat{\widehat{h}}_k$, of the k^{th} fading link at time t_i , obtained from the observations $r_k \left(T_{\lceil \frac{k}{L} \rceil} \right)$ and $r_k \left(T_{\lceil \frac{N}{L} \rceil} \right)$, takes the form

$$\begin{aligned} \widehat{\widehat{h}}_k &= \frac{p^*}{\sqrt{E}} r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) \frac{\rho_k^{(i)} - \xi_k \rho_N^{(i)} + \gamma^{-1} \rho_k^{(i)}}{(1+\gamma^{-1})^2 - |\xi_k|^2} \\ &+ \frac{p^*}{\sqrt{E}} r_k \left(T_{\lceil \frac{N}{L} \rceil} \right) \frac{\rho_N^{(i)} - \xi_k \rho_k^{(i)} + \gamma^{-1} \rho_N^{(i)}}{(1+\gamma^{-1})^2 - |\xi_k|^2}, \end{aligned} \quad (19)$$

where ξ_k is the correlation between $h_k \left(T_{\lceil \frac{N}{L} \rceil} \right)$ and $h_k \left(T_{\lceil \frac{k}{L} \rceil} \right)$, and $\rho_N^{(i)}$ is the correlation between $h_k \left(T_{\lceil \frac{N}{L} \rceil} \right)$ and $h_k(t_i)$. Further, $\widehat{\widehat{h}}_k$ is \mathcal{CN} with variance given by

$$\text{var} \left[\widehat{\widehat{h}}_k \right] \triangleq \sigma_k^2 = \frac{(1+\gamma^{-1}) \left(|\rho_k^{(i)}|^2 + |\rho_N^{(i)}|^2 \right) - 2\xi_k \rho_k^{(i)} \rho_N^{(i)}}{(1+\gamma^{-1})^2 - |\xi_k|^2}. \quad (20)$$

Proof: The proof is given in Appendix E. ■

We will now show that, given \mathcal{S}_v , the optimal subset selection rule again linearly weights the channel estimates. However, the optimal weights turn out to be different from those in (17). This is because refining a channel estimate using the second observation changes its variance.

Theorem 3: When the voids are filled with antenna subset \mathcal{S}_v , the optimal antenna subset that minimizes the SEP of the MPSK symbol received at time t_i is given by

$$\Omega'_L = \left\{ \lceil 1 \rceil, \lceil 2 \rceil, \dots, \lceil L \rceil \right\}, \quad (21)$$

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

$$w'_{[\widehat{l}'],i} |\widehat{h}'_{[\widehat{l}']}|^2 > \dots > w'_{[\widehat{N}'],i} |\widehat{h}'_{[\widehat{N}']}|^2. \quad (22)$$

Here, $w'_{k,i} = \frac{1}{1+\gamma^{-1}-\sigma_k'^2}$, $\widehat{h}'_k = \begin{cases} \widehat{h}_k, & \text{if } k \notin \mathcal{S}_v \\ \widehat{\widehat{h}}_k, & \text{if } k \in \mathcal{S}_v \end{cases}$, and

$$\sigma_k'^2 = \begin{cases} \frac{|\rho_k^{(i)}|^2}{1+\gamma^{-1}}, & \text{if } k \notin \mathcal{S}_v \\ \sigma_k^2, & \text{if } k \in \mathcal{S}_v \end{cases}.$$

Proof: The proof is given in Appendix F. ■

Thus, the parameters \widehat{h}'_k and $\sigma_k'^2$ and, consequently, the corresponding linear weights are updated only for the antennas whose channel gain estimates are refined due to void-filling.

B. Void-Filling Criteria

Having found the optimal subset selection rule given \mathcal{S}_v , we now investigate different criteria to determine \mathcal{S}_v itself. It is a subset of $\{1, 2, \dots, L \lfloor \frac{N}{L} \rfloor\}$, which is the set of all antennas that have been sounded exactly once before.

1) *Random Void-Filling:* The simplest strategy is to randomly pick v distinct elements from the set $\{1, 2, \dots, L \lfloor \frac{N}{L} \rfloor\}$. For this void-filling criterion, the SEP is derived in closed-form below.

Theorem 4: With random void-filling and optimal subset selection (as per Theorem 3) the SEP of an MPSK symbol transmitted at time t_i is given by

$$\begin{aligned} \text{SEP}_i(\gamma) &= \frac{1}{\pi} \frac{1}{\binom{L \lfloor \frac{N}{L} \rfloor}{v}} \sum_{\omega \in \mathcal{C}(L \lfloor \frac{N}{L} \rfloor, v)} \sum_{\lambda \in \mathcal{P}_N} \left[\prod_{k=1}^N \left(\sum_{m=1}^k \frac{\Delta_{\lambda(k)}(\omega)}{\Delta_{\lambda(m)}(\omega)} \right)^{-1} \right] \\ &\quad \times \sum_{n=1}^N \left(\prod_{\substack{l=1 \\ l \neq n}}^N \frac{f'_n(\lambda)}{f'_n(\lambda) - f'_l(\lambda)} \right) \\ &\quad \times \left[\frac{M-1}{M} \pi + \sqrt{\frac{f'_n(\lambda)}{1+f'_n(\lambda)}} \tan^{-1} \left(\sqrt{\frac{1+f'_n(\lambda)}{f'_n(\lambda)}} \tan\left(\frac{\pi}{M}\right) \right) \right], \end{aligned} \quad (23)$$

For $n \leq L$, $f'_n(\lambda) = n \left[\sum_{p=1}^n \frac{1}{\Delta_{\lambda(p)}(\omega)} \right]^{-1}$ and, for $n > L$,

$$f'_n(\lambda) = L \left[\sum_{p=1}^n \frac{1}{\Delta_{\lambda(p)}(\omega)} \right]^{-1}; \quad \Delta_k(\omega) = \frac{\sin^2(\frac{\pi}{M}) |\rho_k^{(i)}|^2}{(1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2},$$

for $k \notin \omega$, and $\Delta_k(\omega) = \frac{\sigma_k^2 \sin^2(\frac{\pi}{M})}{1+\gamma^{-1}-\sigma_k^2}$, for $k \in \omega$.

Proof: The proof is given in Appendix G. ■

Thus, random void-filling is an analytically tractable criterion and serves as a good benchmark to compare other void-filling criteria against.

2) *Void-Filling Using Antennas With Largest Linearly Weighted Estimates:* In Sec. III-D, when all N channel estimates are available, we saw that an antenna k got selected for data reception at time t_i depending on its weighted channel estimate $w_{k,i}^{\text{opt}} |\widehat{h}_k|^2$, where $w_{k,i}^{\text{opt}}$ is given by Theorem 2. This motivates the following void-filling criterion that, unlike random void-filling, exploits the channel estimates that are available at the receiver for some of the antennas.

Criterion 1: Choose the v antennas to fill voids that have the v largest weighted channel estimates among $\left\{ w_{k,i}^{\text{opt}} |\widehat{h}_k|^2 \right\}_{k=1}^{L \lfloor N/L \rfloor}$, where $w_{k,i}^{\text{opt}}$ is the SEP-optimal weight for the k^{th} antenna and the i^{th} MPSK symbol.

Intuitively, the antennas that are most likely to be picked based on the $L \lfloor N/L \rfloor$ channel estimates obtained thus far are chosen, and their channel estimates are refined. A closed-form expression for the SEP for this void-filling criterion is analytically intractable since \mathcal{S}_v now depends on $\widehat{h}_1, \widehat{h}_2, \dots, \widehat{h}_{L \lfloor N/L \rfloor}$.

3) *Near-Optimal Void-Filling Criterion:* Having considered two intuitive void-filling criteria, we now determine the optimal criterion for determining \mathcal{S}_v . It is chosen so that when it is followed by its corresponding optimal subset selection (as per Theorem 3), it minimizes the SEP of demodulating the data symbol transmitted at time t_i . For this, we need an expression for the SEP that is conditioned on \mathcal{S}_v and $\{\widehat{h}_k, k \in \mathcal{S}_v\}$. The optimal \mathcal{S}_v is the subset that minimizes this SEP expression. However, deriving such an expression is analytically intractable, as was also the case for the simpler void-filling criterion considered earlier in Sec. IV-B2.

To circumvent this problem, we assume that *the optimal void-filling criterion ensures that an antenna that is used for void-filling will also be selected for data reception.*² The following important result, which leads to a single integral expression for the SEP, then follows.

Theorem 5: The SEP of an MPSK symbol, transmitted at time t_i , conditioned on \mathcal{S}_v and $\{\widehat{h}_k, k \in \mathcal{S}_v\}$ and under the assumption that an antenna used for void-filling is also used for data reception, is given by

$$\begin{aligned} \text{SEP}_i(\mathcal{S}_v, \{\widehat{h}_k, k \in \mathcal{S}_v\}) &= \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_{N-v}} \prod_{k=1}^{N-v} \left(\sum_{m=1}^k \frac{\Lambda_{\lambda(k)}}{\Lambda_{\lambda(m)}} \right)^{-1} \int_0^{\frac{M-1}{M} \pi} \left[\prod_{n=1}^{N-v} \frac{\sin^2 \theta}{\sin^2 \theta + g_n(\lambda)} \right] \\ &\quad \times \prod_{k \in \mathcal{S}_v} \left[1 + \frac{\left(\sigma_k^2 (1+\gamma^{-1}) - |\rho_k^{(i)}|^2 \right) \sin^2(\frac{\pi}{M})}{(1+\gamma^{-1})(1+\gamma^{-1}-\sigma_k^2) \sin^2 \theta} \right]^{-1} \\ &\quad \times \exp \left(\frac{-|\widehat{h}_k|^2}{\left(\sigma_k^2 - \frac{|\rho_k^{(i)}|^2}{1+\gamma^{-1}} + (1+\gamma^{-1}-\sigma_k^2) \frac{\sin^2 \theta}{\sin^2(\frac{\pi}{M})} \right)} \right) d\theta. \end{aligned} \quad (24)$$

For $n \leq L-v$, $g_n(\lambda) = n \left[\sum_{p=1}^n \frac{1}{\Lambda_{\lambda(p)}} \right]^{-1}$ and, for $n > L-v$, $g_n(\lambda) = (L-v) \left[\sum_{p=1}^n \frac{1}{\Lambda_{\lambda(p)}} \right]^{-1}$, and $\Lambda_k = \frac{|\rho_k^{(i)}|^2 \sin^2(\frac{\pi}{M})}{(1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2}$.

Proof: The derivation is relegated to Appendix H. ■

²The intuition behind the assumption is as follows. From Theorem 3, we see that the optimal weight for an antenna with a refined estimate is greater than that for an antenna with one estimate. Hence, antennas in \mathcal{S}_v are more likely to be selected. Given this assumption, the subsequent results are exact, which is why the resultant void-filling criterion is called a *near-optimal* one.

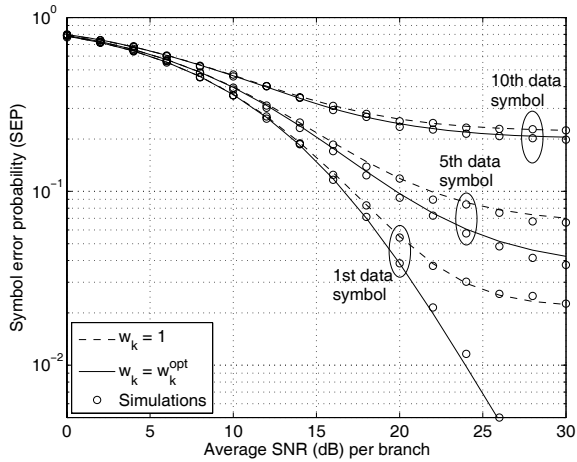


Fig. 2. Effect of the data symbol transmission times and weights (16PSK, $T_p = 10T_s$, $L = 2$, $N = 4$, and $f_d T_p = 0.06$).

The optimal antenna subset, \mathcal{S}_v , that should be used for void-filling can now be determined by computing (24) for different subsets and choosing the one that yields the smallest SEP. Evaluating it is considerably less computationally intensive than using Monte Carlo simulations-based averaging techniques. We also see that the above optimal void-filling criterion does not linearly weight the estimates, unlike the approach of Sec. IV-B2. Note, however, that the corresponding optimal subset selection rule of Theorem 3 (given \mathcal{S}_v) does linearly weight the estimates. It may even select an antenna not in \mathcal{S}_v , though the odds of this are small.

V. SIMULATIONS

We now verify and study the results derived in Secs. III and IV. While our analysis is valid for arbitrary Doppler spectra, we use the classical Jakes spectrum in our simulations. Therefore, the correlation values for $k = 1, 2, \dots, N$ and $i = 1, 2, \dots, d$, are given by $\rho_k^{(i)} = J_0(2\pi f_d ((\lceil \frac{N}{L} \rceil - \lceil \frac{k}{L} \rceil) T_p + iT_s))$ and $\xi_k = J_0(2\pi f_d ((\lceil \frac{N}{L} \rceil - \lceil \frac{k}{L} \rceil) T_p))$, where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind [25] and f_d is the maximum Doppler frequency. The figures are plotted for $T_p = 10T_s$. We first study the case without void-filling, and then investigate the impact of void-filling.

Figure 2 compares the SEPs of the optimal and conventional (no-weighting) subset selection rules. This is done for the 1st, 5th, and 10th data symbols, which are transmitted at times t_1 , t_5 , and t_{10} , respectively. The figure is plotted for 16PSK with $N = 4$ antennas, $L = 2$ RF chains, and a normalized Doppler spread of $f_d T_p = 0.06$. In this case, no voids occur since $L = 2$ is a divisor of $N = 4$. We observe that optimal weighting significantly outperforms no-weighting. For data symbols that are transmitted later, we see that: (i) the SEP increases for both the weighting schemes as the channel estimates become more outdated, and (ii) the relative gains obtained by using the optimal selection weights decrease because the relative variation among the weights of different antennas decreases. Notice also that the analytical and simulation results agree with each other for both the weighting schemes. Therefore, we

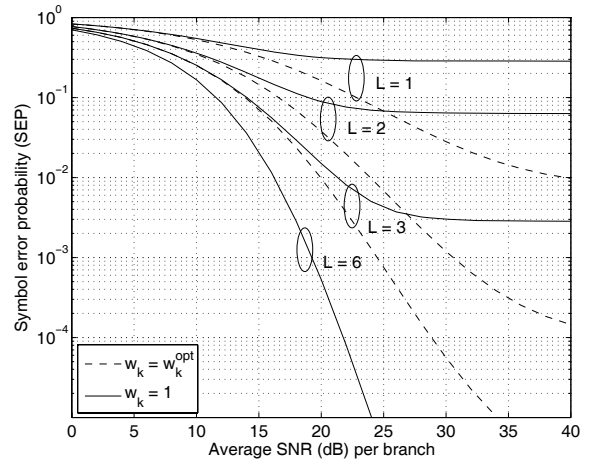


Fig. 3. Effect of the number of receive RF chains and weights (16PSK, $T_p = 10T_s$, $i = 1$, $N = 6$, and $f_d T_p = 0.06$).

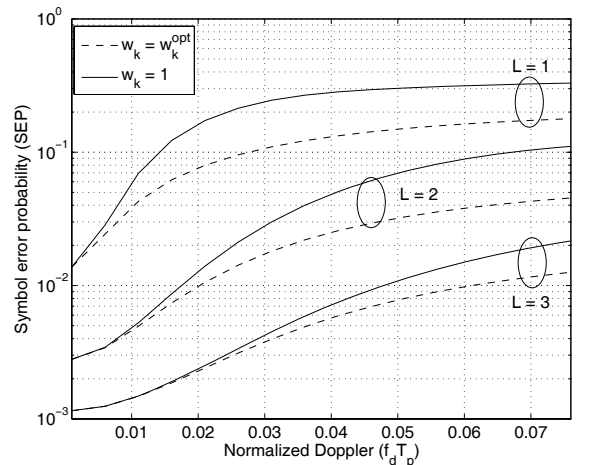


Fig. 4. Effect of the number of receive RF chains and weights (16PSK, $T_p = 10T_s$, $i = 1$, $N = 6$, and SNR = 20 dB).

no longer plot simulation results. Also, since the SEP behavior is qualitatively similar for different t_i , we shall henceforth focus on the SEP of the 1st ($i = 1$) data symbol.

The effect of the number of RF chains, L , on the SEP is investigated in Fig. 3 for different weighted selection schemes. This is done for 16PSK with $N = 6$ antennas and $f_d T_p = 0.06$. Since $f_d T_p > 0$, an error floor occurs due to outdated channel estimates. Optimal weighting significantly reduces this error floor. At a given SNR, the performance gain from using optimal selection weights over no-weighting decreases as L increases. And, when $L = N$, where there is no antenna selection, no performance gain occurs.

Figure 4 plots the SEP as a function of normalized Doppler spread for different selection weights and for different L at an SNR of $\gamma = 20$ dB. As the Doppler spread increases, the SEP increases, as expected, since the channel estimates become more outdated. The gains from using the optimal selection weights are evident at all the Doppler spreads.

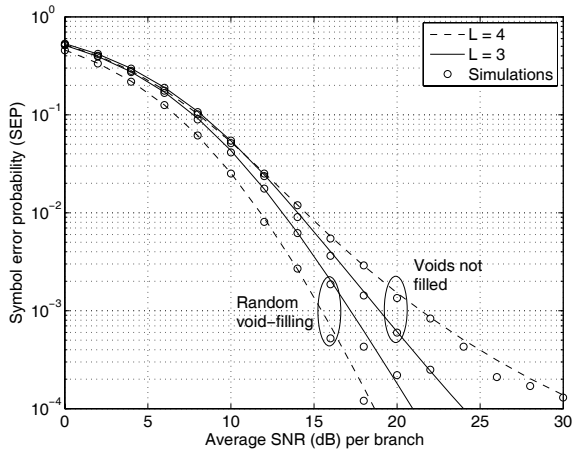


Fig. 5. Effect of number of receive RF chains and random void-filling (8PSK, $f_d T_p = 0.06$, $T_p = 10T_s$, $i = 1$, and $N = 5$).

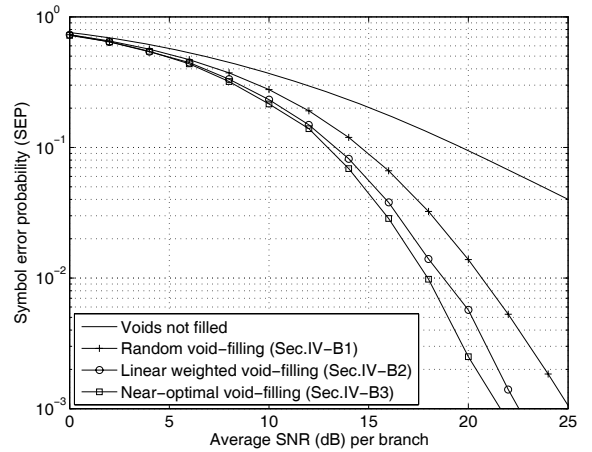


Fig. 6. Performance of different void-filling strategies (16PSK, $L = 3$, $f_d T_p = 0.08$, $T_p = 10T_s$, $i = 1$, and $N = 7$).

A. Impact of Void-Filling

Figure 5 shows the decrease in SEP achieved by random void-filling compared to no void-filling. The SEP is plotted for 8PSK and $N = 5$ with $L = 3$ and $L = 4$. For random void-filling, both analytical and simulation results are plotted, and match very well. The plot brings out the following interesting anomaly that occurs when the voids are not filled. In this case, *at high SNR, the SEP for $L = 4$ is worse than $L = 3$* . This is because how outdated the estimates are affects the SEP the most, at higher SNRs. With $L = 4$, only one out of the 5 antennas gets a fresh (less outdated) estimate from the last (second) pilot and there are three unused RF chains while receiving the second pilot symbol. $L = 3$ does better since two antennas get a fresh estimate from the second (last) pilot and there is one unused RF chain while receiving the second pilot symbol. However, with void-filling, the number of fresh estimates is always L since L antennas are always trained from the last pilot. Therefore, the SEP monotonically decreases as L increases for all SNR.

The performance of all the three void-filling criteria developed in Sec. IV-B is compared in Fig. 6 for $L = 3$ and $N = 7$. In this case, there are two training voids when the last (third) pilot is received. All three void-filling criteria outperform no void-filling. Even random void-filling reduces the SEP by an order of magnitude at an SNR of 20 dB. Furthermore, optimal void-filling outperforms linear weighted void-filling, which, in turn, outperforms random void-filling.

B. Subset Selection and Void-Filling to Receive Packets

In the theory developed thus far, we saw that the subset selection weights and, therefore, the optimal antenna subset depend on the time, t_i , at which the i^{th} data symbol is received, and may be different from one symbol to the other. In practice, the non-zero time required to switch between antennas imposes the additional constraint that all the d data symbols of a packet must be received by the same antenna subset. Motivated by our above results, we propose the following sub-optimal subset selection rule for receiving the entire packet using the same antenna subset.

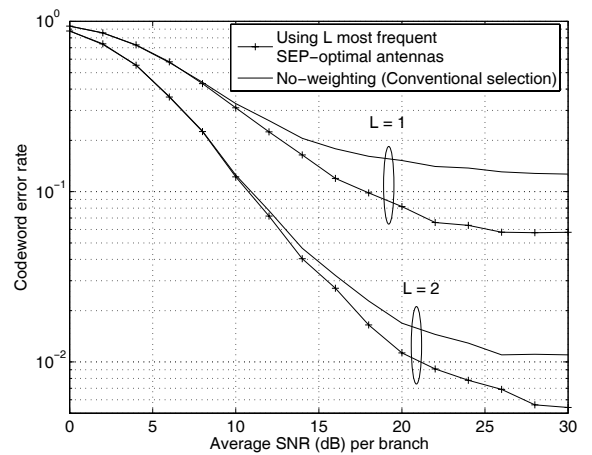


Fig. 7. Performance of weighted selection schemes for packet reception. (8PSK, $T_p = 10T_s$, $d = 50$, $f_d T_p = 0.03$, and $N = 4$).

Criterion 2: Find the SEP-optimal L -antenna subset for each data symbol in the packet (using Theorem 2). Then, choose the L antennas that will be selected for the most number of data symbols.

In addition, when L does not divide N , the subset of antennas chosen to fill training voids must be the same for all the data symbols since the last training symbol (in which voids occur) precedes the reception of the multiple data symbols (see Fig. 1). We propose the following sub-optimal void-filling criterion, which is again motivated by our results for symbol-by-symbol reception.

Criterion 3: For each data symbol of the packet, find the v -antenna subset that should be used for void-filling (as per Sec. IV-B). Then, choose the v antennas that will be selected most often.

Once the voids are filled, the sub-optimal subset selection rule mentioned above, which uses the weights as per Theorem 3, determines the antenna subset that will receive the packet. Note that the above rule can be applied to both coded and uncoded data packets.

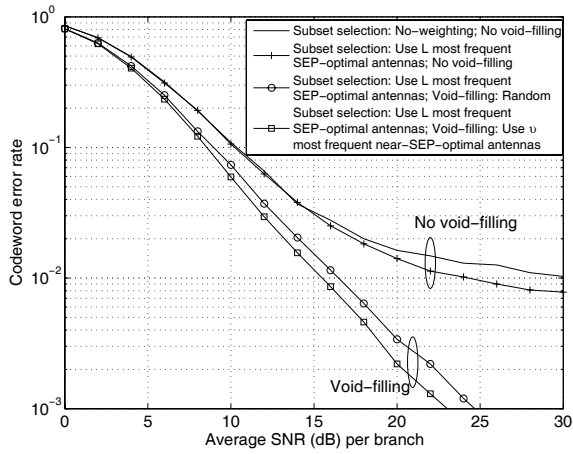


Fig. 8. Performance of void-filling schemes for packet reception (8PSK, $T_p = 10T_s$, $d = 10$, $f_d T_p = 0.03$, $L = 3$, and $N = 4$).

To illustrate the efficacy of the above rules, we consider reception of a packet encoded using a practical rate 1/3 convolutional code, with generator polynomial [133 171 165], which is specified in the 3GPP cellular systems standard [26]. Hard decision Viterbi decoding is performed at the receiver. Figure 7 considers the case where L divides N , and shows that the above subset selection rule outperforms conventional no-weighting based selection. The case where L does not divide N , which requires both void-filling and subset selection, is considered in Fig. 8. Both near-optimal (Sec. IV-B3) and random void-filling (Sec. IV-B1) outperform no void-filling.

VI. CONCLUSIONS

We investigated training and selection criteria for receive antenna subset selection in time-varying channels given the practical AS training constraints imposed by next generation wireless standards. The fact that the channel estimates of different antennas are outdated by different amounts affects the optimal subset selection rule and the overall system performance. 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 selection weights and the SEP of MPSK with optimal subset selection for any L .

We encountered the problem of training voids, which occurs when L does not divide N , when 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 reduces the SEP. We showed that the optimal subset selection rule still uses a linear weighting of the refined channel estimates, albeit with different weights. Even choosing the antennas randomly to fill the voids reduces SEP significantly, and provides an analytically tractable performance benchmark. Taking into account the channel estimates that have been obtained for some of the antennas yields further performance gains. To this end, we derived a near-optimal void-filling rule and an expression for its SEP. We also saw that the theory developed for SEP with symbol-by-symbol

subset selection motivates effective sub-optimal selection and void-filling rules for receiving coded or uncoded data packets under the additional practical constraint that all the data symbols of a packet must be received by the same antenna subset.

APPENDIX

A. Derivation of the Conditional Mean and Variance of y_k in (12) and (13)

The result about the conditional mean directly follows from (8) since \hat{h}_k is the MMSE estimate of h_k and n_k is independent of \hat{h}_k . The conditional variance result is derived as follows. From (1), it follows that $\text{var}[y_k | \hat{h}_k, s_i] = \text{var}[h_k | \hat{h}_k] E + N_0$. From (10), we have $\text{var}[h_k | \hat{h}_k] = 1 - \text{var}[\hat{h}_k]$, and from (2), it can be shown that

$$\text{var}[\hat{h}_k] = |\rho_k^{(i)}|^2 \left(\frac{E}{E+N_0} \right)^2 + \sigma_e^2 = |\rho_k^{(i)}|^2 \frac{1}{1+\gamma^{-1}}. \quad (25)$$

Combining these results yields the desired expression.

B. Derivation of the Conditional Mean and Variance of \mathcal{D}_i

Starting from (14), we get

$$\mu_{\mathcal{D}_i} \triangleq \mathbf{E}[\mathcal{D}_i | \Omega_L, \{\hat{h}_k, k \in \Omega_L\}, s_i], \quad (26)$$

$$= \sum_{k \in \Omega_L} \frac{\hat{h}_k^* \mathbf{E}[y_k | \Omega_L, \{\hat{h}_k, k \in \Omega_L\}, s_i]}{(1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2}. \quad (27)$$

Using (12), this simplifies to $\mu_{\mathcal{D}_i} = s_i \sum_{k \in \Omega_L} \frac{|\hat{h}_k|^2}{(1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2}$.

Similarly, from (14), we get

$$\sigma_{\mathcal{D}_i}^2 \triangleq \text{var}[\mathcal{D}_i | \Omega_L, \{\hat{h}_k, k \in \Omega_L\}, s_i], \quad (28)$$

$$= \sum_{k \in \Omega_L} \frac{|\hat{h}_k|^2 \text{var}[y_k | \Omega_L, \{\hat{h}_k, k \in \Omega_L\}, s_i]}{\left((1+\gamma^{-1})^2 - |\rho_k^{(i)}|^2 \right)^2}. \quad (29)$$

Substituting the expression for the conditional variance from (13) and simplifying gives the desired result.

C. Proof of Theorem 1

Since \mathcal{D}_i conditioned on Ω_L and $\{\hat{h}_k\}_{k \in \Omega_L}$ is a Gaussian RV, the standard SEP expression for MPSK is [27, (40)]

$$\begin{aligned} \text{SEP}_i \left(\Omega_L, \{\hat{h}_k\}_{k \in \Omega_L} \right) &= \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left(\frac{-|\mu_{\mathcal{D}_i}|^2 E \sin^2(\frac{\pi}{M})}{\sigma_{\mathcal{D}_i}^2 \sin^2 \theta} \right) d\theta. \quad (30) \end{aligned}$$

Substituting the expressions in (15) for the conditional mean and variance of \mathcal{D}_i , we get

$$\begin{aligned} \text{SEP}_i \left(\Omega_L, \left\{ \widehat{h}_k \right\}_{k \in \Omega_L} \right) \\ = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left(- \frac{\sin^2 \left(\frac{\pi}{M} \right)}{\sin^2 \theta} \sum_{k \in \Omega_L} \frac{(1+\gamma^{-1}) \left| \widehat{h}_k \right|^2}{(1+\gamma^{-1})^2 - \left| \rho_k^{(i)} \right|^2} \right) d\theta. \end{aligned} \quad (31)$$

Now, writing the above expression in terms of β_k (defined in the theorem statement) and the weights $\{w_{k,i}\}_{k=1}^N$, we get

$$\text{SEP}_i \left(\Omega_L, \left\{ \widehat{h}_k \right\}_{k \in \Omega_L} \right) = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left(- \sum_{k \in \Omega_L} \frac{\beta_k w_{k,i} \left| \widehat{h}_k \right|^2}{\sin^2 \theta} \right) d\theta. \quad (32)$$

The integrand in the SEP expression in (32) is of the form $\exp \left(- \frac{1}{\sin^2 \theta} \sum_{k \in \Omega_L} \beta_k X_k^{(i)} \right)$, where $X_k^{(i)} \triangleq w_{k,i} \left| \widehat{h}_k \right|^2$ is an exponential RV. From (25), we get $\mathbf{E} \left[X_k^{(i)} \right] \triangleq \alpha_k = w_{k,i} \left| \rho_k^{(i)} \right|^2 (1+\gamma^{-1})^{-1}$. Furthermore, $X_1^{(i)}, \dots, X_N^{(i)}$ are mutually independent, but they are not identically distributed. When sorted in descending order, we get $X_{[1]}^{(i)} > X_{[2]}^{(i)} > \dots > X_{[N]}^{(i)}$. The selection rule then selects the antenna subset $\Omega_L = \left\{ [1], \dots, [L] \right\}$.

Let $Y \triangleq \sum_{k=1}^L \beta_{[k]} X_{[k]}^{(i)}$. Therefore, the SEP in (32) is equivalent to $\text{SEP}_i \left(\Omega_L, \left\{ \widehat{h}_k \right\}_{k \in \Omega_L} \right) = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left(\frac{-Y}{\sin^2 \theta} \right) d\theta$. For a given permutation λ of the set $\{1, 2, \dots, N\}$, let A_λ denote the event that it gives the sorted order above, *i.e.*, $[1] = \lambda(1), [2] = \lambda(2), \dots, [N] = \lambda(N)$. The expression for SEP, when averaged over fading and when conditioned over all possible A_λ , becomes

$$\text{SEP}_i(\gamma) = \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_N} \Pr(A_\lambda) \int_0^{\frac{M-1}{M}\pi} \mathcal{M}_{Y|A_\lambda} \left(\frac{1}{\sin^2 \theta} \right) d\theta. \quad (33)$$

Using the virtual branch combining (VBC) technique of [28], it can be shown that

$$\Pr(A_\lambda) = \prod_{k=1}^N \left[\sum_{m=1}^k \frac{\alpha_{\lambda(k)}}{\alpha_{\lambda(m)}} \right]^{-1}, \quad \mathcal{M}_{Y|A_\lambda}(y|A_\lambda) = \prod_{n=1}^N \frac{1}{1+yf_n(\lambda)},$$

where $f_n(\cdot)$ is as defined in the theorem statement. Substituting these in (33) results in

$$\begin{aligned} \text{SEP}_i(\gamma) = \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_N} \left(\prod_{k=1}^N \left[\sum_{m=1}^k \frac{\alpha_{\lambda(k)}}{\alpha_{\lambda(m)}} \right]^{-1} \right) \\ \times \int_0^{\frac{M-1}{M}\pi} \prod_{n=1}^N \frac{\sin^2 \theta}{\sin^2 \theta + f_n(\lambda)} d\theta. \end{aligned} \quad (34)$$

The partial fraction expansion of the integrand in (34) is

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

The integral of each of the above terms can be written in closed-form using the following identity, which follows from [25, (2.562)]:

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

Substituting the above identity in (34) yields the desired result.

D. Proof of Theorem 2

From (31), the SEP conditioned on Ω_L and $\left\{ \widehat{h}_l \right\}_{l \in \Omega_L}$ is given by

$$\begin{aligned} \text{SEP}_i \left(\Omega_L, \left\{ \widehat{h}_k \right\}_{k \in \Omega_L} \right) \\ = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left(- \frac{1}{\sin^2 \theta} \sum_{k \in \Omega_L} \frac{(1+\gamma^{-1}) \sin^2 \left(\frac{\pi}{M} \right)}{(1+\gamma^{-1})^2 - \left| \rho_k^{(i)} \right|^2} \left| \widehat{h}_k \right|^2 \right) d\theta. \end{aligned}$$

From the above equation, it follows that the subset Ω_L that minimizes the SEP of MPSK is

$$\Omega_L = \arg \max_{\omega \subset \{1, \dots, N\}} \sum_{k \in \omega} \frac{1}{(1+\gamma^{-1})^2 - \left| \rho_k^{(i)} \right|^2} \left| \widehat{h}_k \right|^2, \quad (35)$$

ω is an L -element subset of $\{1, \dots, N\}$. *The key thing to note is that the function to be maximized in (35) is additive in k .* This implies that the best antenna subset is simply the L antennas with the L largest values of $w_{k,i}^{\text{opt}} \left| \widehat{h}_k \right|^2$, for $k = 1, \dots, N$, where $w_{k,i}^{\text{opt}}$ is given in (17). Hence, the result follows.

E. Proof of Lemma 1

Since $h_k(t_i)$ and $\left[r_k \left(T_{\lceil \frac{N}{L} \rceil} \right) \quad r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) \right]^T$ are zero-mean jointly Gaussian, it follows from (6) that the MMSE estimate of $h_k(t_i)$, conditioned on $r_k \left(T_{\lceil \frac{N}{L} \rceil} \right)$ and $r_k \left(T_{\lceil \frac{k}{L} \rceil} \right)$, is

$$\widehat{h}_k = \begin{bmatrix} \sqrt{E} p \rho_N^{(i)} \\ \sqrt{E} p \rho_k^{(i)} \end{bmatrix}^\dagger \begin{bmatrix} E(1+\gamma^{-1}) & \xi_k E \\ \xi_k E & E(1+\gamma^{-1}) \end{bmatrix}^{-1} \begin{bmatrix} r_k \left(T_{\lceil \frac{N}{L} \rceil} \right) \\ r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) \end{bmatrix}. \quad (36)$$

This is on account of the following results:

$$\begin{aligned} \mathbf{E} \left[h_k(t_i) r_k^* \left(T_{\lceil \frac{k}{L} \rceil} \right) \right] &= \sqrt{E} p^* \rho_k^{(i)}, \\ \mathbf{E} \left[h_k(t_i) r_k^* \left(T_{\lceil \frac{N}{L} \rceil} \right) \right] &= \sqrt{E} p^* \rho_N^{(i)}, \\ \mathbf{E} \left[r_k \left(T_{\lceil \frac{N}{L} \rceil} \right) r_k \left(T_{\lceil \frac{k}{L} \rceil} \right)^* \right] &= \xi_k E. \end{aligned}$$

Further,

$$\begin{aligned} \mathbf{E} \left[\left| r_k \left(T_{\lceil \frac{N}{L} \rceil} \right) \right|^2 \right] &= \mathbf{E} \left[\left| r_k \left(T_{\lceil \frac{k}{L} \rceil} \right) \right|^2 \right], \\ &= E \left(1 + \gamma^{-1} \right). \end{aligned}$$

Hence, (19) follows.

Since $r_k \left(T_{\lceil \frac{N}{L} \rceil} \right)$ and $r_k \left(T_{\lceil \frac{k}{L} \rceil} \right)$ are \mathcal{CN} , so is \widehat{h}_k . Therefore, using (11), the variance of \widehat{h}_k can be derived.

F. Proof of Theorem 3

Let Ω'_L denote the antenna subset selected for data reception given \mathcal{S}_v . Recall from (5) that the signal received at time t_i is $y_k = h_k(t_i)s_i + n_k(t_i)$, for $k \in \Omega'_L$.

From (8) and (10), we can show that, for $k \in \mathcal{S}_v$,

$$\mathbf{E} \left[y_k \mid \widehat{h}_k, s_i \right] = s_i \widehat{h}_k, \quad (37)$$

$$\text{var} \left[y_k \mid \widehat{h}_k, s_i \right] = E \left(1 + \gamma^{-1} - \sigma_k^2 \right). \quad (38)$$

From (12) and (13), we also know that, for $k \notin \mathcal{S}_v$,

$$\mathbf{E} \left[y_k \mid \widehat{h}_k, s_i \right] = s_i \widehat{h}_k, \quad (39)$$

$$\text{var} \left[y_k \mid \widehat{h}_k, s_i \right] = E \left(1 + \gamma^{-1} - \frac{|\rho_k^{(i)}|^2}{1 + \gamma^{-1}} \right). \quad (40)$$

When MRC is employed at the receiver, it follows from the above set of equations that the decision variable, \mathcal{D}_i , for data decoding is

$$\mathcal{D}_i = \sum_{k \in \Omega'_L} \frac{(\widehat{h}'_k)^* y_k}{1 + \gamma^{-1} - \sigma_k'^2}, \quad (41)$$

where the summation is over the selected antenna subset (Ω'_L),

$$\widehat{h}'_k = \begin{cases} \widehat{h}_k & \text{if } k \notin \mathcal{S}_v \\ \widehat{h}_k & \text{if } k \in \mathcal{S}_v \end{cases}, \quad (42)$$

and

$$\sigma_k'^2 = \begin{cases} \frac{|\rho_k^{(i)}|^2}{1 + \gamma^{-1}} & \text{if } k \notin \mathcal{S}_v \\ \sigma_k^2 & \text{if } k \in \mathcal{S}_v \end{cases}. \quad (43)$$

Conditioned on $\mathcal{S}_v, \Omega'_L, \{\widehat{h}'_k, k \in \Omega'_L\}$, and s_i , the decision variable \mathcal{D}_i is a complex Gaussian RV. Using (37), (38), (39), and (40) the conditional mean and variance of \mathcal{D}_i can be shown, along the lines of Appendix B, to be

$$\begin{aligned} \mathbf{E} \left[\mathcal{D}_i \mid \mathcal{S}_v, \Omega'_L, \{\widehat{h}'_k, k \in \Omega'_L\}, s_i \right] &= \sum_{k \in \Omega'_L} \frac{s_i |\widehat{h}'_k|^2}{1 + \gamma^{-1} - \sigma_k'^2}, \\ \text{var} \left[\mathcal{D}_i \mid \mathcal{S}_v, \Omega'_L, \{\widehat{h}'_k, k \in \Omega'_L\}, s_i \right] &= \sum_{k \in \Omega'_L} \frac{E |\widehat{h}'_k|^2}{1 + \gamma^{-1} - \sigma_k'^2}. \end{aligned}$$

Conditioned on \mathcal{S}_v, Ω'_L , and $\{\widehat{h}'_k\}_{k \in \Omega'_L}$, the SEP of an MPSK symbol transmitted at time t_i is, therefore,

$$\begin{aligned} \text{SEP}_i \left(\mathcal{S}_v, \Omega'_L, \{\widehat{h}'_k\}_{k \in \Omega'_L} \right) \\ = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left(- \sum_{k \in \Omega'_L} \frac{|\widehat{h}'_k|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k'^2) \sin^2 \theta} \right) d\theta. \quad (44) \end{aligned}$$

Thus, the SEP-optimal antenna subset, Ω'_L , should maximize

$$\sum_{k \in \Omega'_L} \frac{1}{1 + \gamma^{-1} - \sigma_k'^2} |\widehat{h}'_k|^2,$$

which also is additive in k . The separability argument, as used earlier in Appendix D, then leads to the desired result.

G. Proof of Theorem 4

Using (44), the SEP of an MPSK symbol transmitted at time t_i can be written as

$$\begin{aligned} \text{SEP}_i \left(\mathcal{S}_v, \Omega'_L, \left\{ \widehat{h}_k, k \in \mathcal{S}_v \cap \Omega'_L \right\}, \left\{ \widehat{h}_k, k \in \mathcal{S}_v^c \cap \Omega'_L \right\} \right) \\ = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left[- \sum_{k \in \mathcal{S}_v \cap \Omega'_L} \frac{|\widehat{h}_k|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k^2) \sin^2 \theta} \right. \\ \left. - \sum_{k \in \mathcal{S}_v^c \cap \Omega'_L} \frac{|\widehat{h}_k|^2 (1 + \gamma^{-1}) \sin^2 \left(\frac{\pi}{M} \right)}{\left((1 + \gamma^{-1})^2 - |\rho_k^{(i)}|^2 \right) \sin^2 \theta} \right] d\theta. \quad (45) \end{aligned}$$

Considering all possible random v -element subsets of the set $\{1, \dots, L \lfloor \frac{N}{L} \rfloor\}$, the SEP averaged over all the subsets \mathcal{S}_v is

$$\begin{aligned} \text{SEP}_i \left(\Omega'_L, \left\{ \widehat{h}_k, k \in \Omega'_L \right\}, \left\{ \widehat{h}_k \right\}_{k=1}^N \right) &= \frac{1}{\pi} \frac{1}{\binom{L \lfloor \frac{N}{L} \rfloor}{v}} \\ &\times \sum_{\omega \in \mathcal{C} \left(L \lfloor \frac{N}{L} \rfloor, v \right)} \int_0^{\frac{M-1}{M}\pi} \exp \left[- \sum_{k \in \omega \cap \Omega'_L} \frac{|\widehat{h}_k|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k^2) \sin^2 \theta} \right. \\ &\left. - \sum_{k \in \omega^c \cap \Omega'_L} \frac{|\widehat{h}_k|^2 (1 + \gamma^{-1}) \sin^2 \left(\frac{\pi}{M} \right)}{\left((1 + \gamma^{-1})^2 - |\rho_k^{(i)}|^2 \right) \sin^2 \theta} \right] d\theta. \end{aligned}$$

The integrand in the above SEP expression is of the form $\exp \left(- \frac{1}{\sin^2 \theta} \sum_{k \in \Omega'_L} Z_k(\omega) \right)$, where $Z_k(\omega)$ is an exponential RV. Let $\mathbf{E} [Z_k(\omega)] \triangleq \Delta_k(\omega)$. From (20) and (25), we have

$$\Delta_k(\omega) = \begin{cases} \frac{\sigma_k^2 \sin^2 \left(\frac{\pi}{M} \right)}{1 + \gamma^{-1} - \sigma_k^2}, & \text{for } k \in \omega \\ \frac{|\rho_k^{(i)}|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1})^2 - |\rho_k^{(i)}|^2}, & \text{for } k \notin \omega \end{cases}. \quad (46)$$

Using VBC, as done in Theorem 1, and simplifying produces the desired result.

H. Proof of Theorem 5

When the antenna subset that is used for void-filling is also used for data reception, we have $\mathcal{S}_v \subset \Omega'_L$. Let Θ denote the $L-v$ element subset of the other antennas that are selected. For these antennas, only one observation is available for channel estimation. From Theorem 3, these antennas have the $L-v$ largest weighted channel estimates in the subset $\{k \in \{1, 2, \dots, N\} \setminus \mathcal{S}_v\}$, with the weights given by (17). The SEP expression in (45) then simplifies to

$$\begin{aligned} \text{SEP}_i \left(\mathcal{S}_v, \Theta, \left\{ \widehat{h}_k, k \in \mathcal{S}_v \right\}, \left\{ \widehat{h}_k, k \in \Theta \right\} \right) \\ = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \exp \left[- \sum_{k \in \mathcal{S}_v} \frac{|\widehat{h}_k|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k^2) \sin^2 \theta} \right. \\ \left. - \sum_{k \in \Theta} \frac{|\widehat{h}_k|^2 (1 + \gamma^{-1}) \sin^2 \left(\frac{\pi}{M} \right)}{\left((1 + \gamma^{-1})^2 - |\rho_k^{(i)}|^2 \right) \sin^2 \theta} \right] d\theta. \end{aligned}$$

Given \mathcal{S}_v , the sets $\{\widehat{h}_k, k \in \Theta\}$ and $\{\widehat{h}_k, k \in \mathcal{S}_v\}$ are independent since the channel gains are i.i.d. and $\Theta \cap \mathcal{S}_v$ is a null set. Hence, upon averaging over $\{\widehat{h}_k, k \in \Theta\}$, we get

$$\begin{aligned} \text{SEP}_i \left(\mathcal{S}_v, \left\{ \widehat{h}_k, k \in \mathcal{S}_v \right\}, \Theta \right) = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \mathcal{M}_{\sum_{k \in \Theta} X'_k} \left(\frac{-1}{\sin^2 \theta} \right) \\ \times \exp \left(- \sum_{k \in \mathcal{S}_v} \frac{|\widehat{h}_k|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k^2) \sin^2 \theta} \right) d\theta, \end{aligned}$$

where $X'_k = \frac{|\widehat{h}_k|^2 (1 + \gamma^{-1}) \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1})^2 - |\rho_k^{(i)}|^2}$ is an exponential RV with mean

$$\mathbf{E}[X'_k] \triangleq \Lambda_k = \frac{|\rho_k^{(i)}|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1})^2 - |\rho_k^{(i)}|^2}.$$

Note that Θ is obtained by picking the largest $L-v$ elements of $\{X'_k, k \in \{1, 2, \dots, N\} \setminus \mathcal{S}_v\}$. Using VBC, the above SEP expression upon averaging over Θ simplifies to

$$\begin{aligned} \text{SEP}_i \left(\mathcal{S}_v, \left\{ \widehat{h}_k, k \in \mathcal{S}_v \right\} \right) \\ = \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_{N-v}} \left[\prod_{k=1}^{N-v} \left[\sum_{m=1}^k \frac{\Lambda_{\lambda(k)}}{\Lambda_{\lambda(m)}} \right]^{-1} \right] \int_0^{\frac{M-1}{M}\pi} \left[\prod_{n=1}^{N-v} \frac{\sin^2 \theta}{\sin^2 \theta + g_n(\lambda)} \right] \\ \times \exp \left(- \sum_{k \in \mathcal{S}_v} \frac{|\widehat{h}_k|^2 \sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k^2) \sin^2 \theta} \right) d\theta, \quad (47) \end{aligned}$$

where $g_n(\cdot)$ is as defined in the theorem statement. Conditioned on \mathcal{S}_v , the estimates $\{\widehat{h}_k, k \in \mathcal{S}_v\}$ are mutually independent. Hence, upon averaging the above expression over

$\{\widehat{h}_k, k \in \mathcal{S}_v\}$ and conditioning on $\{\widehat{h}_k, k \in \mathcal{S}_v\}$, we get

$$\begin{aligned} \text{SEP}_i \left(\mathcal{S}_v, \left\{ \widehat{h}_k, k \in \mathcal{S}_v \right\} \right) \\ = \frac{1}{\pi} \sum_{\lambda \in \mathcal{P}_{N-v}} \prod_{k=1}^{N-v} \left(\sum_{m=1}^k \frac{\Lambda_{\lambda(k)}}{\Lambda_{\lambda(m)}} \right)^{-1} \int_0^{\frac{M-1}{M}\pi} \left[\prod_{n=1}^{N-v} \frac{\sin^2 \theta}{\sin^2 \theta + g_n(\lambda)} \right] \\ \times \prod_{k \in \mathcal{S}_v} \mathcal{M}_{|\widehat{h}_k|^2 | \widehat{h}_k} \left(\frac{\sin^2 \left(\frac{\pi}{M} \right)}{(1 + \gamma^{-1} - \sigma_k^2) \sin^2 \theta} \right) d\theta. \quad (48) \end{aligned}$$

Also, \widehat{h}_k conditioned on \widehat{h}_k is a Gaussian RV. Using (6) and (7), it can be shown that

$$\mathbf{E} \left[\widehat{h}_k | \widehat{h}_k \right] = \widehat{h}_k, \quad (49)$$

$$\mathbf{var} \left[\widehat{h}_k | \widehat{h}_k \right] = \sigma_k^2 - |\rho_k^{(i)}|^2 (1 + \gamma^{-1})^{-1}. \quad (50)$$

Hence, $|\widehat{h}_k|^2$ conditioned on \widehat{h}_k is a Chi-square RV whose conditional MGF equals [29]

$$\begin{aligned} \mathcal{M}_{|\widehat{h}_k|^2 | \widehat{h}_k} (x) = \left(1 + x \left(\sigma_k^2 - |\rho_k^{(i)}|^2 (1 + \gamma^{-1})^{-1} \right) \right)^{-1} \\ \times \exp \left(\frac{-|\widehat{h}_k|^2 x}{1 + x \left(\sigma_k^2 - |\rho_k^{(i)}|^2 (1 + \gamma^{-1})^{-1} \right)} \right). \quad (51) \end{aligned}$$

Substituting the MGF in (48) and simplifying further yields the desired result.

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," *MIMO System Technology for Wireless Communications*, G. Tsoulos, editor, 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] A. B. Narasimhamurthy and C. Tepedelenlioglu, "Antenna selection for MIMO OFDM systems with channel estimation error," in *Proc. Globecom*, pp. 3290-3294, 2007.
- [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] 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.
- [8] A. Conti, W. M. Gifford, M. Z. Win, and M. Chiani, "Optimized simple bounds for diversity systems," *IEEE Trans. Commun.*, vol. 57, pp. 2674-2685, Sep. 2009.
- [9] A. Conti, M. Z. Win, and M. Chiani, "Slow adaptive M-QAM with diversity in fast fading and shadowing," *IEEE Trans. Commun.*, vol. 55, pp. 895-905, May 2007.
- [10] M. Z. Win, R. K. Mallik, and G. Chrisikos, "Higher order statistics of antenna subset diversity," *IEEE Trans. Commun.*, vol. 2, pp. 871-875, Sep. 2003.
- [11] M. Z. Win, N. C. Beaulieu, L. A. Shepp, B. F. Logan, Jr., and J. H. Winters, "On the SNR penalty of MPSK with hybrid selection/maximal ratio combining over i.i.d. Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 1012-1023, June 2003.

- [12] A. F. Molisch, M. Z. Win, and J. H. Winters, "Reduced-complexity transmit/receive-diversity systems," *IEEE Trans. Signal Process.*, vol. 51, pp. 2729-2738, Nov. 2003.
- [13] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio-combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, pp. 1773-1776, Dec. 1999.
- [14] M. Simon and M. S. Alouini, *Digital Communication over Fading Channels.*, 2nd edition. Wiley-Interscience, 2005.
- [15] X. N. Zeng and A. Ghayeb, "Performance bounds for space-time block codes with receive antenna selection," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2130-2137, Sep. 2004.
- [16] 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.
- [17] Z. Chen, J. Yuan, and B. Vucetic, "Analysis of transmit antenna selection/maximal-ratio combining in Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 54, pp. 1312-1321, July 2005.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.*, 2006.
- [23] 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.
- [24] 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, July 2010.
- [25] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products.*, 4th edition. Academic Press, 1980.
- [26] "Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); multiplexing and channel coding (release 8)," tech. rep. 36.212 (v8.2.0), 3rd Generation Partnership Project (3GPP), 2008.
- [27] 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, Sep. 1999.
- [28] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining of diversity branches with unequal SNR in Rayleigh fading," in *Proc. VTC*, 1999.
- [29] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th edition. Dover, 1972.



Vinod Kristem received his Bachelor of Technology degree in Electronics and Communications Engineering from the National Institute of Technology (NIT), Warangal in 2007. He received his Master of Engineering degree in Telecommunications from Indian Institute of Science, Bangalore, India in 2009. Since then, he has been with Beceem Communications Pvt. Ltd., Bangalore, India (which was recently acquired by Broadcom Corp.) working on channel estimation and physical layer measurements for WiMAX and LTE. His research interests include

the design and analysis of algorithms for wireless communication networks, MIMO systems, and cooperative communications.



Neelesh B. Mehta (S'98-M'01-SM'06) received his Bachelor of Technology degree in Electronics and Communications Engineering from the Indian Institute of Technology (IIT), Madras in 1996, and his M.S. and Ph.D. degrees in Electrical Engineering from the California Institute of Technology, Pasadena, CA in 1997 and 2001, respectively. He was a visiting graduate student researcher at Stanford University in 1999 and 2000. He is now an Assistant Professor at the Dept. of Electrical Communication Engineering, Indian Institute of Science (IISc), Bangalore, India. Until 2002, he was a research scientist in the Wireless Systems Research group in AT&T Laboratories, Middletown, NJ. In 2002-2003, he was a Staff Scientist at Broadcom Corp., Matawan, NJ, and was involved in GPRS/EDGE cellular handset development. From 2003-2007, he was at the Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, USA.

His research includes work on link adaptation, multiple access protocols, WCDMA downlinks, system-level performance analysis of cellular systems, MIMO and antenna selection, and cooperative communications. He was also actively involved in radio access network physical layer (RAN1) standardization activities in 3GPP. He has served on several TPCs. He was a TPC co-chair for WISARD 2010 and 2011, National Conference on Communications (NCC) 2011, the Transmission technologies track of VTC 2009 (Fall), the Frontiers of Networking and Communications symposium of Chinacom 2008, and the tutorials co-chair for SPCOM 2010. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and an executive committee member of the IEEE Bangalore Section and the Bangalore chapter of the IEEE Signal Processing Society.



Andreas F. Molisch (S'89, M'95, SM'00, F'05) received the Dipl. Ing., Dr. techn., and habilitation degrees from the Technical University Vienna (Austria) in 1990, 1994, and 1999, respectively. From 1991 to 2000, he was with the TU Vienna, becoming an associate professor there in 1999. From 2000-2002, he was with the Wireless Systems Research Department at AT&T (Bell) Laboratories Research in Middletown, NJ. From 2002-2008, he was with Mitsubishi Electric Research Labs, Cambridge, MA, USA, most recently as Distinguished Member of

Technical Staff and Chief Wireless Standards Architect. Concurrently he was also Professor and Chairholder for radio systems at Lund University, Sweden. Since 2009, he is Professor of Electrical Engineering at the University of Southern California, Los Angeles, CA, USA.

Dr. Molisch has done research in the areas of SAW filters, radiative transfer in atomic vapors, atomic line filters, smart antennas, and wideband systems. His current research interests are measurement and modeling of mobile radio channels, UWB, cooperative communications, and MIMO systems. Dr. Molisch has authored, co-authored or edited four books, among them the textbook *Wireless Communications*, 2nd edition (Wiley-IEEE Press, 2011), fourteen book chapters, some 130 journal papers, and numerous conference contributions, as well as more than 70 patents.

Dr. Molisch is an Area editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS for Antennas and Propagation and co-editor of special issues of several journals. He has been member of numerous TPCs, vice-chair of the TPCs of VTC 2005 spring and VTC 2010 spring, general chair of ICUWB 2006, TPC co-chair of the wireless symposium of Globecom 2007, TPC chair of Chinacom2007, general chair of Chinacom 2008, TPC co-chair of the Commun. Theory Workshop 2009, tutorial co-chair of WCNC 2009, and workshop co-chair at ICC 2010. He has participated in the European research initiatives "COST 231", "COST 259", and "COST273", where he was chairman of the MIMO channel working group, he was chairman of the IEEE 802.15.4a channel model standardization group. From 2005-2008, he was also chairman of Commission C (signals and systems) of URSI (International Union of Radio Scientists), and in 2009-10, he was the Chair of the Radio Communications Committee of the IEEE Communications Society. Dr. Molisch is a Fellow of the IEEE, a Fellow of the IET, an IEEE Distinguished Lecturer, and recipient of several awards.