

Effects of Channel Prediction for Transmit Antenna Selection with Maximal Ratio Combining in Rayleigh Fading

Shiva Prakash, *Student Member, IEEE*, and Ian McLoughlin, *Senior Member, IEEE*

Abstract—Antenna selection is a low-complexity method for pragmatically exploiting spatial diversity in wireless systems. It has potentially reduced hardware cost compared to space-time or MIMO (multiple-input and multiple-output) coding, due to the reduction in the amount of radio frequency (RF) hardware required. Whilst receive antenna selection is perhaps more common, transmit antenna selection also offers several advantages, particularly for hardware-costly transmit schemes such as those requiring linearisation. In use, transmit antenna selection (TAS) requires at least partial channel knowledge at the transmitter in order to perform selection. This knowledge usually comes in the form of an index to the best set of antenna/antennas fed back from the receiver; which implies a delay between the channel being sampled (at the receiver) and this knowledge being acted upon (at the transmitter). In this paper, performance degradation due to outdated channel knowledge is determined analytically, and related to channel characteristics. A predictive scheme is then developed to mitigate against delay-induced degradation. Several factors relating to TAS system performance under different channel scenarios both with and without mitigation are explored. Closed form expressions for performance metrics such as bit error rate (BER), outage probability, average signal-to-noise ratio (SNR) gain and higher order moments of output SNR are derived and verified by simulations. The impact of prediction is analyzed for different TAS setups and channel prediction scenarios, as are various system design parameters.

I. INTRODUCTION

MIMO (multiple-input and multiple-output) is often cited as a solution for achieving the high data rate demands of future wireless networks through increased spectral efficiency and link reliability [1]. Future systems are expected to further optimise performance by being capable of adapting to varying propagation conditions during operation – switching or adjusting parameters to suit dynamic channel characteristics such as channel state matrix, signal-to-noise ratio (SNR) and the presence of interference. Based on whether or not channel knowledge is available at the transmitter, MIMO systems can be classed as either open or closed loop. Traditional space time coding (STC) techniques are inherently open loop, and Orthogonal Space Time Block Codes (OSTBC) [2] can achieve full diversity while enjoying just simple maximum-likelihood (ML) decoding. The low code rate disadvantage of this scheme when used with more than two transmit antennas and complex modulations, can be rectified by a trade-off between spectral efficiency and higher decode complexity. However, there is still a need for extra transmit chains. Further performance gains can be obtained in closed loop systems by providing a channel state information (CSI) feedback path,

creating a system in which transmission is on the Eigenmodes of the transmit antenna correlation matrix. The optimal design of the resulting MIMO system is for transmission in the direction of the strongest channel covariance matrix eigenvector [1]. These schemes generally perform better than STC, and can improve reliability and increase throughput via both array and diversity gain. However, the improvement comes at the expense of greater hardware cost and increased CSI feedback rates, and may be sensitive to feedback delay. MIMO system costs are also higher in that transmit/receive RF electronics paths are required for every additional antenna, and algorithmic processing consumes computational resources.

Transmit selection diversity or transmit antenna selection (TAS) is a cost-effective and low complexity closed loop technique in which a receiver periodically advises the transmitter to select a subset of transmit antennas, chosen to maximise SNR. Compared to transmit beamforming, this can substantially reduce hardware costs, since fewer RF hardware chains are required: just multiplexers to select a subset of transmit antennas. At the same time, diversity and array gain benefits are maintained. Furthermore, antenna index feedback needs only a low bandwidth channel (compared to full CSI feedback). In this paper we analyse a TAS scheme employing maximal ratio combining (MRC) at the receiver, where a single best antenna is selected for transmission based on the knowledge of fading statistics at the receiver. MRC, as an optimal combining scheme irrespective of channel fading statistics, is suitable for most amplitude and phase modulated signals [3], and has been adopted at the receiver. At high SNRs a TAS/MRC scheme such as this can achieve full diversity order, and can approach the performance of some complex space-time codes of the same spectral efficiency [4].

Feedback delay, leading to outdated channel knowledge at the transmitter, is particularly troublesome for TAS: acting on degraded information to select a transmit antenna, effectively reduces antenna diversity, adversely affecting symbol error probability (SEP) and bit error rate (BER). It is, of course, more significant in fast fading channels where a return link delay might render any channel information completely outdated by the time the transmit antennas switch. This is especially true if TAS is used in mobile systems operating at S-band frequencies of around 2 to 3.7 GHz, such as in WiMAX (worldwide interoperability for microwave access), where Doppler shift is an important consideration. Furthermore, such adaptive techniques when combined with wireless systems that have restricted uplink feedback rates,

such as W-CDMA (wideband code-division multiple-access) or 3GPP (third-generation partnership project – with a 1.5kbps uplink bandwidth) are more likely to experience large reverse channel feedback delays. Other factors typically contributing to switching delay include decode latency, ARQ (automatic repeat request) handling and block buffering.

A. Related work

Channel prediction was proposed initially for a DECT (Digital European Cordless Telecommunications) framework to improve transmitter combining and selection diversity gain [5], however the analysis is limited to a single receive antenna at the mobile. The importance of long range channel prediction beyond coherence time in mitigating the effect of deep fades was investigated by Eyceoz and Duel-Hallen [6], while [7], [8] applied channel and power prediction to improve transmit diversity performance in frequency selective W-CDMA channels. However, the analysis was restricted to a two antenna, single receiver system, where an upper bound for BER with perfect prediction was the performance benchmark. With regard to channel prediction, an extensive analysis of linear prediction for mobile radio channel coefficients has been developed by Ekman [9], [10], where efficient noise reduction using Wiener smoothers, as well as an unbiased power predictor in SISO (single-input and single-output) channels is proposed to improve performance. Øien et al. [11] employed prediction using pilot symbol assisted modulation (PSAM) for combatting delay in an adaptive modulation receive diversity system. Similarly Zhou [12] dealt with channel prediction for transmit beam-forming with MRC. Linear channel prediction was applied for two transmit selective Alamouti STBC schemes [13], however this did not include a complete mathematical analysis for the prediction process. Joint transmit and receive antenna selection was considered in [14], [15] where the impact of feedback delay on BER was been analysed over flat fading Rayleigh channels. Cai and Giannakis [16] considered the performance of an ideal combined transmit selection diversity and receive generalized selection combining over Rayleigh fading channels. Moving beyond this, a spatial diversity analysis of an ideal TAS/MRC scheme was published by Chen *et al.* [4], where performance metrics such as BER for binary phase shift keying (BPSK), outage probability and gains of the TAS/MRC scheme for flat Rayleigh fading channels were derived. Finally, the authors recently introduced a similar TAS scheme limited to a two transmit, two receive configuration [17]. With only partial analysis, this did not consider outage probability, average SNR gains, amount of fading, antenna placement trade-offs, or the equal-performance threshold between predictive and non-predictive TAS.

B. Contributions and organisation of paper

To mitigate issues related to delay between channel measurement and switching, this paper considers a predictive-TAS (TASP) scheme allied with MRC at the receiver, which employs a power predictor to exploit temporal channel correlation. This predictor is designed for TAS in future transmission

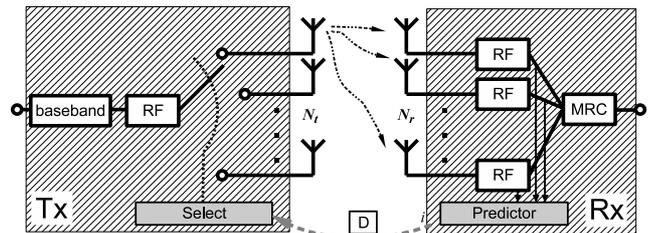


Fig. 1: Block diagram showing transmit antenna selection based upon information from a feedback path from MRC receive antennas and a predictor.

blocks based upon current and historical channel information [17]. The benefits of using long range prediction (LRP) of channel values for TASP decisions are evaluated and compared with a non-predictive (TASD) scheme acting on outdated channel information, for a range of system configurations. Several performance metrics related to BER, outage probability and SNR gains, which will be useful for efficient system design, have been derived and analysed. A block fading model facilitating LRP is incorporated into the TAS setup, and TASD is compared with TASP. Several design trade-off possibilities, that may be of use to system implementers, will be discussed. In all cases we assume that noiseless estimates are available for receiver demodulation, so that feedback delay effects, with and without prediction, can be isolated.

The paper is organized as follows. Section II presents the system structure and channel models, where in Section II-E, the fading probability density function (PDF) of the TASP scheme is derived. In Section III, the closed form BER equation is derived and the merits of using a predictor are discussed with respect to various TAS/MRC configurations. Section IV derives the outage probability of the TASP scheme and compares it with the delayed version. Next in Section V, the fading statistic of the TASP system is derived for use in recognizing TAS/MRC systems of similar performance. Section VI then discusses the investigations of TASP performance w.r.t channel power prediction performance and their relevant operating characteristics, then Section VII will conclude the paper.

II. SYSTEM MODEL AND ASSUMPTIONS

In the following introductory paragraphs, we will first describe the TAS/MRC system and the antenna selection process in an ideal system with no feedback delay constraints. Next we briefly describe the issue of feedback delay in TAS/MRC. In Section II-B we review the well known Wiener channel predictor. In II-C we derive statistics for the power prediction process, the results of which will be used in II-D to derive TAS/MRC fading distributions.

A block flat fading SIMO (single-input multiple-output) channel with N_t possible transmit and N_r receive antennas is considered in a TAS/MRC system, as shown in Fig.1. Based upon an index l , $1 \leq l \leq N_t$, fed-back to the transmit end, a single best transmit antenna is selected from N_t candidates for block data transmission. This arrangement is denoted as $(N_t, 1; N_r)$. Diversity reception with maximal ratio combining

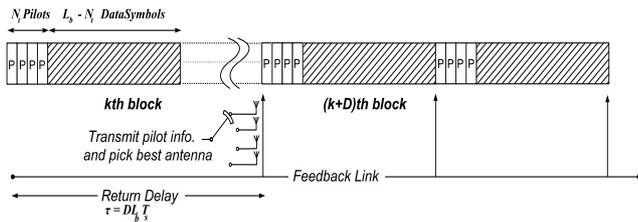


Fig. 2: Transmission frame, showing pilot symbols at the start of each transmission block, and delay between pilot transmission and the antenna switching point.

and coherent demodulation is employed at the receiver, and a block stationary channel assumed. We adopt the well known pilot symbol aided modulation (PSAM) technique that was formulated in [18]. Our PSAM block structure essentially amounts to a SIMO version of the one in [12], which extends PSAM plus diversity [11] (consisting of a single transmit and multiple receive antennas), into a full MIMO beam-forming system.

As shown in Fig. 2, periodic pilot symbols from each antenna are inserted at the beginning of every block of length L_b in a turn-by-turn fashion, so as to enable channel estimation for each transmit-receive pair. The pilot symbols constitute the first N_t entries of the block and aid in channel prediction and estimation. For data transmission, only one active transmit antenna is used for each block. The channel matrix of size $(N_r \times N_t)$ for the k^{th} block is $\mathbf{H}(k)$ with the complex fading coefficient $h_{ij}(k)$, $1 \leq i \leq N_t$, $1 \leq j \leq N_r$ as its entry, denoting the true fading coefficient of the channel between the i^{th} transmit antenna and the j^{th} receive antenna.

The channels $h_{ij}(k)$ are characterised as spatially independent and identically distributed (i.i.d)¹ complex Gaussian random variables with distribution $\mathcal{CN}(0, \sigma_h^2)$ that follow Jakes model [19], [20] with Doppler spread f_d (assumed equal for all channels). Also they are considered to change slowly so that multiple training signals can be transmitted without significant frame overhead [21], making antenna selection for block fading a practical solution. Essentially, we will assume that channel coefficients are constant over a block, and temporally correlated between blocks. For each block instant k , the transmitter receives switching information from the receiver: an integer l , denoting the index of the best transmit antenna to be used. The criterion for selection is to pick an antenna which provides the highest power or SNR gain amongst all transmitters. In an ideal case, where there is no feedback delay, the receiver picks the transmit antenna with signal power p_{max} to maximize the post processing SNR at the output of the maximal ratio combiner.

$$p_{max}(k) = \max_{1 \leq i \leq N_t} p_i(k) \quad (1)$$

$$p_i(k) = \sum_{j=1}^{N_r} |h_{ij}(k)|^2 \quad (2)$$

¹The assumption of i.i.d is valid since large antenna separation at the base station is feasible, while for the mobile receiver the separation distance can be made small because of a rich scattering environment.

The index, l , of p_{max} is communicated to the transmitter through a feedback channel. Then for subsequent block k the received signal vector will be expressed as:

$$\mathbf{y}(k; m) = \sqrt{E_s} \mathbf{h}_l(k) x(k; m) + \mathbf{z}(k; m) \quad (3)$$

with $k \geq 0$, $N_t \leq m \leq L_b - 1$, and $x(k; m)$ representing a data symbol at time instant $(kL_b T_s + mT_s)$ transmitted from the selected antenna with E_s as the power per data symbol, in the k^{th} block. T_s is the symbol duration in seconds and block length is L_b symbols. $\mathbf{h}_l(k)$ is a $(N_r \times 1)$ channel vector, the l^{th} column of the channel matrix $\mathbf{H}(k)$, and assumed to be perfectly estimated. $\mathbf{z}(k)$ is the additive white Gaussian noise (AWGN) vector with distribution $\mathcal{CN}(0, \sigma_z^2 \mathbf{I}_{N_r})$ with $\sigma_z^2 = N_0$, the receiver AWGN noise power for a given receive branch. \mathbf{I}_{N_r} is an identity matrix of dimension N_r . Perfect channel estimation is assumed to enable coherent receiver demodulation. A common way to achieve this in practice, used in adaptive systems such as [22], [12], [11], is through noncausal channel smoothing Wiener interpolation filters [18], [23]. We also assume that index l is received error-free.

A. TAS with feedback delay (TASD)

Even with perfect channel estimates being used for antenna selection, the performance of TAS/MRC will degrade as a result of acting on outdated knowledge at the transmitter. Large delays render feedback information useless, effectively breaking the feedback loop. Delays are commonly caused by error coding/decoding, reverse channel blocking, ARQ, bandwidth restrictions or higher layer protocol processing. Given a \mathcal{D} block delay, the transmitter relies on switching information sent by the receiver \mathcal{D} blocks earlier. Thus (1) and (2) now become:

$$p'_{max}(k) = \max_{1 \leq i \leq N_t} p'_i(k) \quad (4)$$

$$p'_i(k) = \sum_{j=1}^{N_r} |h_{ij}(k - \mathcal{D})|^2 \quad (5)$$

The channel coefficients at the present block k differ from those at block $(k - \mathcal{D})$, with the relationship [24], [25]:

$$h_{ij}(k) = \rho_d \sigma_h^2 h_{ij}(k - \mathcal{D}) + \sqrt{\sigma_h^2 - \rho_d^2} n_{ij}(k) \quad (6)$$

where the correlation between the true and delayed or outdated fading gain is given as $\rho_d = \mathbb{E}[h_{ij}(k) h_{ij}^*(k - \tau)] = \sigma_h^2 J_0(2\pi f_d \tau)$ [19]. $\mathbb{E}[\cdot]$ is the expectation operator, $J_0(\cdot)$ is the zeroth order Bessel function of the first kind and τ is the feedback time delay length given by $\tau = \mathcal{D} L_b T_s$. $n_{ij}(k)$ is AWGN with zero mean and unit variance. Delay causes channel mismatch, hence performance degradation occurs as a result of poor antenna selection. In addition, since $f_d \propto v f_c$, where v is the vehicle speed and f_c is the operating carrier frequency, the normalised delay term $f_d \tau$ becomes larger for higher vehicle speeds and higher frequencies. Thus the channel correlation ρ_d is dependent on feedback delay, velocity and carrier frequency. For very small values of normalised feedback delay $f_d \tau \ll 1$, sufficient channel correlation exists so that $p'_{max}(k) \approx p_{max}(k)$. Therefore switching information is

relevant, and the BER degradation may be small. For a large delay, ρ_d tends to zero; essentially resulting in the transmitter acting on completely outdated channel knowledge, causing incorrect antenna selection. Intuitively one would then expect the TASD/MRC system to behave like an open loop system: simple MRC with one transmit antenna. In general, despite using a clear estimate for antenna selection, delay degrades the BER of TAS/MRC, an effect that will be explored in Section II-D and beyond.

B. Wiener channel prediction

To combat delay we aim to predict channel coefficients ahead of time, to improve antenna selection performance. The channel prediction filter here is strictly causal in nature, and to enable prediction, channel estimates for each block are obtained using PSAM [18]. For a block of data, channel estimation is carried out independently for all channels, with the entries of channel matrix $\mathbf{H}(k)$ estimated as:

$$\tilde{h}_{ij}(k) = h_{ij}(k) + v_{ij}(k) \quad (7)$$

where $\tilde{h}_{ij}(k)$ is the channel estimate while $v_{ij}(k)$ is the AWGN channel estimation error with distribution $\mathcal{CN}(0, \sigma_v^2)$, with $\sigma_v^2 = N_0/E_p$, E_p being the power of the pilot symbol. Thus the variance of the estimated channel amplitude is given by $\sigma_h^2 = \sigma_h^2 + \sigma_v^2$, since it is assumed that $h_{ij}(k)$ and $v_{ij}(k)$ are statistically independent. Note that the same estimates will aid in channel smoothing for perfect receiver demodulation as assumed before, therefore $\tilde{h}_{ij}(k) = h_{ij}(k)$, for receiver symbol detection. For channel prediction, a vector of delayed estimates obtained from (7) are used. E_p/N_0 here is assumed to be at least equal to the average receive SNR, $\bar{\gamma}$. In practice, the pilot power could be made large enough to achieve channel smoothing for demodulation purposes [26].

The Wiener-Hopf equation for the \mathcal{D} block ahead predicted channel is $\hat{h}_{ij}(k+\mathcal{D}) = \mathbf{w}_{opt}^H \mathbf{h}_{ij}$, where \mathbf{h}_{ij} is the complex vector of estimated fading amplitudes, corresponding to prediction length L given by $\mathbf{h}_{ij} = [h_{ij}(k), \tilde{h}_{ij}(k-1), \dots, \tilde{h}_{ij}(k-(L-1))]^T$ and \mathbf{w}_{opt} is the optimal complex coefficient vector given by $\mathbf{w}_{opt} = [\mathbf{R}]^{-1} \mathbf{r}$ where $[\mathbf{R}]_{\varphi\vartheta} = \sigma_h^2 J_0(2\pi f_d |\varphi - \vartheta| L_b T_s) + \sigma_v^2 \delta(\varphi - \vartheta)$, and $\mathbf{r}_\varphi = \sigma_h^2 J_0(2\pi f_d |D + \varphi - 1| L_b T_s)$ $\varphi, \vartheta = 1, 2, \dots, L$. When $\sigma_h^2 = 1$, the normalised correlation coefficient between the true and the predicted channel is given as $\hat{\rho}_{h\hat{h}} = \sqrt{\mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}}$ which is bounded: $0 \leq \hat{\rho}_{h\hat{h}} \leq 1$, a value of one meaning perfect prediction and zero meaning no correlation between predicted and actual channel. The prediction error for any channel is given by $\epsilon_c(k+\mathcal{D}) = h(k+\mathcal{D}) - \hat{h}(k+\mathcal{D})$ with the mean square error (MSE) being minimized when the optimal coefficient vector $\mathbf{w} = \mathbf{w}_{opt}$ is used. Then the MSE is given by $\min_{\mathbf{w}_{opt}} \sigma_{\epsilon_c}^2 = \sigma_h^2 - \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$ and is bounded by $0 \leq \sigma_{\epsilon_c}^2 \leq \sigma_h^2$. Then the true channel can be written as:

$$h_{ij}(k+\mathcal{D}) = \hat{h}_{ij}(k+\mathcal{D}) + \sqrt{\sigma_h^2 - \hat{\rho}_{h\hat{h}}^2} n_{ij}(k+\mathcal{D}) \quad (8)$$

where $n_{ij}(k+\mathcal{D})$ is AWGN with zero mean and unit variance. The predicted channel amplitude is also a Gaussian random variable with variance $\sigma_{\hat{h}}^2 = \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$.

C. Prediction of channel power

In this section we develop a framework for the MRC power predictor, to determine fading statistics for predicted power and post-processing SNR at the receiver. This will be useful in deriving fading statistics for the TASP/MRC system discussed in II-D. Based on the \mathcal{D} block ahead predicted channel coefficients, at instant k , the receiver computes the corresponding predicted channel power for each antenna and selects a transmit antenna, l , corresponding to the maximum power gain. Mathematically this is:

$$\hat{p}_i(k+\mathcal{D}) = \sum_{j=1}^{N_r} \left| \hat{h}_{ij}(k+\mathcal{D}) \right|^2 \quad (9)$$

$$\hat{p}_{max}(k+\mathcal{D}) = \max_{1 \leq i \leq N_t} \hat{p}_i(k+\mathcal{D}) \quad (10)$$

Examining the statistics of the power predictor will naturally lead us to the derivation of the fading PDF of the post-processing SNR, given in II-D. Note that the average value of error $\epsilon_p(k+\mathcal{D}) = p_i(k+\mathcal{D}) - \hat{p}_i(k+\mathcal{D})$, while predicting based on power, is non-zero as was the case with simple channel amplitude prediction; it is known to be biased [9], [10]. It can be seen that $\mathbb{E}[\epsilon_p(k+\mathcal{D})] = N_r(\sigma_h^2 - \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r})$. Thus the power prediction is itself biased. The MSE for the biased predictor is $\sigma_{\epsilon_p}^2 = \mathbb{E}[|p_i(k+\mathcal{D}) - \hat{p}_i(k+\mathcal{D})|^2]$. With use of the average channel power gain $\mathbb{E}[p_i(k+\mathcal{D})] = N_r \sigma_h^2$ and the average predicted power gain $\mathbb{E}[\hat{p}_i(k+\mathcal{D})] = N_r \sigma_{\hat{h}}^2 = N_r \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$, along with $\mathbb{E}[h_{ij}(k+\mathcal{D}) \hat{h}_{ij}^*(k+\mathcal{D})] = \hat{\rho}_{h\hat{h}}$, and identity (A-5), the value of the MSE for the biased power predictor can now be determined, after deriving the following:

$$\mathbb{E}[p_i(k+\mathcal{D})^2] = N_r(N_r+1)\sigma_h^4 \quad (11)$$

$$\mathbb{E}[\hat{p}_i(k+\mathcal{D})^2] = N_r(N_r+1)\sigma_{\hat{h}}^4 \quad (12)$$

$$\mathbb{E}[p_i(k+\mathcal{D})\hat{p}_i(k+\mathcal{D})] = N_r^2 \sigma_h^2 \sigma_{\hat{h}}^2 + N_r |\hat{\rho}_{h\hat{h}}|^4 \quad (13)$$

The above equations are also useful in determining the normalised correlation coefficient ρ_p between the predicted and actual SNR, which is described in II-D. For simplicity, we shall assume $\sigma_h^2 = 1$ for the rest of the calculations. Then the instantaneous post-processing SNR $\gamma_i(k+\mathcal{D})$ for any transmit antenna is:

$$\gamma_i(k+\mathcal{D}) = \frac{E_s}{N_0} p_i(k+\mathcal{D}) \quad (14)$$

and the corresponding predicted SNR gain is,

$$\hat{\gamma}_i(k+\mathcal{D}) = \frac{E_s}{N_0} \hat{p}_i(k+\mathcal{D}) \quad (15)$$

with their means as:

$$\bar{\gamma}_i = \mathbb{E}[\gamma_i(k+\mathcal{D})] = \frac{E_s}{N_0} N_r \sigma_h^2 = \frac{E_s}{N_0} N_r \quad (16)$$

$$\bar{\hat{\gamma}}_i = \mathbb{E}[\hat{\gamma}_i(k+\mathcal{D})] = \frac{E_s}{N_0} N_r \sigma_{\hat{h}}^2 = \frac{E_s}{N_0} N_r \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r} \quad (17)$$

Both are gamma distributed with their probability density function (PDF): $\gamma_i \sim G(N_r, \bar{\gamma})$ and $\hat{\gamma}_i \sim G(N_r, \bar{\hat{\gamma}})$ where N_r is the shape factor, $\bar{\gamma} = \frac{E_s}{N_0}$ and $\bar{\hat{\gamma}} = \frac{E_s}{N_0} \sigma_{\hat{h}}^2$ the scale factors of the gamma distributions, also equal to the average SNR per symbol for the true and predicted channel respectively. For an error rate analysis it is required to determine the fading

distribution of the true maximum channel SNR $\gamma_{max}(k + \mathcal{D})$, which is the maximum of the SNRs $\gamma_i(k + \mathcal{D})$. Similarly the predicted maximum SNR is $\hat{\gamma}_{max}(k + \mathcal{D})$.

D. Fading distribution of a TASP/MRC with prediction

The TASP/MRC symbol error probability, outage probability and fading statistics, can be derived by first obtaining an expression for the PDF $f_{\gamma_{max}}(\gamma)$, which is the distribution of the random variable γ_{max} . We will omit time indices because of the assumption of a stationary random process. The PDF of $f_{\gamma_{max}}(\gamma)$ is:

$$f_{\gamma_{max}}(\gamma) = \int_0^\infty f(\gamma|\hat{\gamma})f_{\hat{\gamma}_{max}}(\hat{\gamma})d\hat{\gamma} \quad (18)$$

$$= \int_0^\infty \frac{f(\gamma, \hat{\gamma})}{f_{\hat{\gamma}}(\hat{\gamma})} f_{\hat{\gamma}_{max}}(\hat{\gamma})d\hat{\gamma} \quad (19)$$

$f(\gamma, \hat{\gamma})$ being the joint distribution of γ and $\hat{\gamma}$, $f(\gamma|\hat{\gamma})$ the PDF of γ conditioned on $\hat{\gamma}$, and $f_{\hat{\gamma}}(\hat{\gamma})$ the PDF of the predicted power given by (we have dropped index i since the distribution is the same for all antennas):

$$f_{\hat{\gamma}}(\hat{\gamma}) = \frac{\hat{\gamma}^{N_r-1}}{\bar{\gamma}^{N_r} (N_r - 1)!} \exp\left(-\frac{\hat{\gamma}}{\bar{\gamma}}\right) \quad (20)$$

$f_{\hat{\gamma}_{max}}(\hat{\gamma})$ is the PDF of $\hat{\gamma}_{max}$. To derive $f_{\hat{\gamma}_{max}}(\hat{\gamma})$, we determine the cumulative distribution function (CDF) of $\hat{\gamma}$:

$$F_{\hat{\gamma}}(\hat{\gamma}) = 1 - \exp(-\hat{\gamma}/\bar{\gamma}) \sum_{m=0}^{N_r-1} (1/m!)(\hat{\gamma}/\bar{\gamma})^m \quad (21)$$

Making use of order statistics [27], the PDF $f_{max}(\hat{\gamma})$ can be calculated as follows:

$$f_{\hat{\gamma}_{max}}(\hat{\gamma}) = N_t f(\hat{\gamma}) [F(\hat{\gamma})]^{N_t-1} \quad (22)$$

Both γ and $\hat{\gamma}$ are mutually correlated, determined through a bi-variate gamma distribution $G' \sim (N_r, \bar{\gamma}, \hat{\gamma}, \rho_p)$ as follows:

$$f_{\gamma, \hat{\gamma}}(\gamma, \hat{\gamma}) = \left\{ \frac{f(\gamma)f(\hat{\gamma})(N_r-1)!}{1-\rho_p} \left(\frac{\rho_p \gamma \hat{\gamma}}{\bar{\gamma} \hat{\gamma}} \right)^{-(N_r-1)/2} I_{N_r-1} \left(\frac{2\sqrt{\rho_p}}{(1-\rho_p)} \sqrt{\frac{\gamma \hat{\gamma}}{\bar{\gamma} \hat{\gamma}}} \right) \exp\left(-\frac{\rho_p}{1-\rho_p} \left(\frac{\gamma}{\bar{\gamma}} + \frac{\hat{\gamma}}{\bar{\gamma}} \right) \right) \right\} \quad (23)$$

$I_{N_r-1}(\cdot)$ is the modified Bessel function of the first kind with order $(N_r - 1)$ and ρ_p is the power correlation coefficient of the true and predicted fading SNRs. To proceed further, we express $\hat{\gamma} = r\bar{\gamma}$ (as seen from (16) and (17)) where $0 \leq r \leq 1$. Utilising $f_{\hat{\gamma}}(\hat{\gamma})$, the PDF of $\hat{\gamma}$, we can write $f(\gamma|\hat{\gamma})$ as:

$$f_{(\gamma|\hat{\gamma})}(\gamma|\hat{\gamma}) = \left\{ \left(\frac{r\gamma}{\rho_p \hat{\gamma}} \right)^{(N_r-1)/2} I_{N_r-1} \left(\frac{2\sqrt{\rho_p \gamma \hat{\gamma}}}{(1-\rho_p) \sqrt{r\bar{\gamma}}} \right) \exp\left(-\frac{\rho_p \hat{\gamma} + r\gamma}{r(1-\rho_p)\bar{\gamma}}\right) \cdot \frac{1}{(1-\rho_p)\bar{\gamma}} \right\} \quad (24)$$

Furthermore, to determine ρ_p we know that,

$$\rho_p = Cov(\gamma, \hat{\gamma}) [Var(\gamma) Var(\hat{\gamma})]^{-0.5} \quad (25)$$

Using (11-13) and (16-17), the value of ρ_p can now be determined. The power correlation coefficient can be shown to equal the square of the channel correlation coefficient so that $\rho_p = \rho_{hh}^2$ (note that this is because the coefficient vector \mathbf{w}_{opt} is also the underlying factor used in the channel power prediction calculation). Note that in this case also $r = \rho_p$. Examining (24) more closely, we see that as ρ_p approaches zero (meaning very poor prediction), the term in the exponential becomes independent of $\hat{\gamma}$. By re-factorising, $f_{(\gamma|\hat{\gamma})}(\gamma|\hat{\gamma})$ is found to be independent of $\hat{\gamma}$ and equates to $f_{\gamma}(\gamma)$. This means that the true instantaneous SNR γ will be independent of the predicted SNR $\hat{\gamma}$, thereby rendering TASP ineffective with the system behaving like a simple MRC system with a single transmit antenna $(1, 1; N_r)$.

By use of multi-nominal and binomial theorems, the power of the polynomial in equation (22) can be expanded for use with (24). Finally, using integration results from [28] (See A-3), the final closed form PDF expression is derived as:

$$f_{\gamma_{max}}(\gamma) = \frac{N_t! \gamma^{N_r-1}}{(N_r - 1)!} \cdot \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \exp\left(-\frac{(i+1)\gamma}{[i(1-\rho_p) + 1]\bar{\gamma}}\right) \cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) j! \frac{[(1-\rho_p)\bar{\gamma}]^j}{[[i(1-\rho_p) + 1]\bar{\gamma}]^{j+N_r}} \cdot L_j^{(N_r-1)}\left(-\frac{\rho_p \gamma}{[i(1-\rho_p) + 1](1-\rho_p)\bar{\gamma}}\right) \quad (26)$$

where $L_j^{N_r-1}(\cdot)$ is the Laguerre polynomial [29] (see A-2) and $\eta_{N_r}(i, j)$ [28] are the coefficients of z^j , $j = 0, 1, \dots, i(N_r - 1)$, in the expansion of $(\sum_{j=0}^{N_r-1} z^j/j!)^i$, obtained as per eqns. (A-1) and (A-2) in the appendix. For the TASP case, the fading PDF is derived in a similar manner. However the important difference is that in the TASP case, as per (4) and (5) we will use a perfectly estimated or noiseless channel coefficient for antenna selection. This can be thought of as a highly smoothed version of the channel estimate, used for generating instantaneous antenna index. Note that for TASP, the PDF of the maximum of SNRs would result in the same form as (26), the only difference being in the power correlation coefficient ρ_p , which has to be replaced now with $\rho_p = J_0^2(2\pi f_d \tau)$ [19]. Later we will see that power correlation ρ_p in TASP can be made to exceed ρ_d for any channel feedback delay, thus outperforming a TASP system; even one that uses a clear channel estimate for index generation. The advantage of the TASP scheme will be observed in the following sections.

III. ERROR RATE ANALYSIS

The average SER at an average SNR $\bar{\gamma}$ per receive path is found by averaging the probability of symbol error in AWGN over the fading distribution of the TASP/P schemes, in slow fading or in quasi-static scenarios. Exact, or approximate (but accurate), values of SER for coherent demodulation can be found from:

$$\bar{P}_s(\bar{\gamma}) = \int_0^\infty P_s(\gamma) f_{\gamma_{max}}(\gamma) d\gamma \quad (27)$$

where $P_s(\gamma) \approx \alpha Q(\sqrt{\beta\gamma})$. α and β are determined by specific constellations and $Q(\cdot)$ is the Q -function [30]. With $\alpha = 1$, $\beta = 2$, the exact BER for BPSK is found, while an approximate SER for rectangular M -QAM (quadrature amplitude modulation) is found with $\alpha = 4(\sqrt{M} - 1)/\sqrt{M}$ and $\beta = 3/(M - 1)$. To evaluate (27) we first expand the Laguerre polynomial using (A-2). Next collecting the terms containing exponential and the powers of γ in (26) along with the Q -function, we have the following inner integral to solve:

$$J_{N_r}(i, k, \beta) = \int_0^\infty \alpha Q\sqrt{\beta\gamma} \exp\left(-\frac{i+1}{i(1-\rho_p)+1}\right) \left(\frac{\gamma}{\bar{\gamma}}\right)^{k+N_r+1} d\gamma$$

$$= \alpha (k + N_r - 1)! \left[\frac{(i(1-\rho_p)+1)(1-\mu_i)}{2(i+1)} \right]^{k+N_r}$$

$$\cdot \sum_{l=0}^{k+N_r-1} \frac{(k+N_r-1+l)!}{l!(k+N_r-1)!} \left(\frac{1+\mu_i}{2}\right)^l \quad (28)$$

$$\mu_i = \sqrt{\frac{\beta(i(1-\rho_p)+1)\bar{\gamma}}{2(i+1) + \beta(i(1-\rho_p)+1)\bar{\gamma}}} \quad (29)$$

which can be solved using the result of the integral [3] (5A.5). Finally eqn. (27) is presented as:

$$\bar{P}_s(\bar{\gamma}) = \alpha \frac{N_t!}{(N_r-1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t-1-i)!}$$

$$\cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) (j + N_r - 1)! \quad (30)$$

$$\cdot \sum_{k=0}^j \frac{\rho_p^k (1-\rho_p)^{j-k}}{[(i(1-\rho_p)+1)]^j} \left[\frac{(1-\mu_i)}{2(i+1)} \right]^{k+N_r}$$

$$\cdot \sum_{l=0}^{k+N_r-1} \frac{(k+N_r-1+l)!}{l!(k+N_r-1)!} \left(\frac{1+\mu_i}{2}\right)^l$$

As a check, we may consider $(N_t, 1; 1)$ with $\beta = 2$ for BPSK modulation, and $\rho_p = 1$, implying full prediction or the immediate (no delay) case, then $\eta_1(i, j) = 1$, and the BER in (30) becomes:

$$\bar{P}_s(\bar{\gamma}) = N_t! \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t-1-i)!} \left[\frac{(1-\sqrt{\bar{\gamma}/(\bar{\gamma}+1)})}{2(i+1)} \right]$$

which verifies against the simple case of [4], equation (26). Similarly when $\rho_p = 0$, $\bar{P}_s(\bar{\gamma}) = \left[\frac{1-\sqrt{\bar{\gamma}/(\bar{\gamma}+1)}}{2} \right]$ which is the trivial case of BER of BPSK in a Rayleigh fading SISO link.

Numerical example: To illustrate the benefits of TASP, let us consider BER performance of TASP (2, 1; 2) compared to simple TASD. For all numerical and Monte Carlo simulation setups, we shall set $T_s = 1e-06$, $E_p/N_0 = 30$ dB, $f_d = 100$ Hz, $L_b = 1/(100f_dT_s)$, $L = 5$ (5 tap predictor). The correlation coefficient is affected by many system parameters, so that several dimensions of system performance trade-off are possible. The choice of these parameters, and how they impact system performance, will be explored in Section VI.

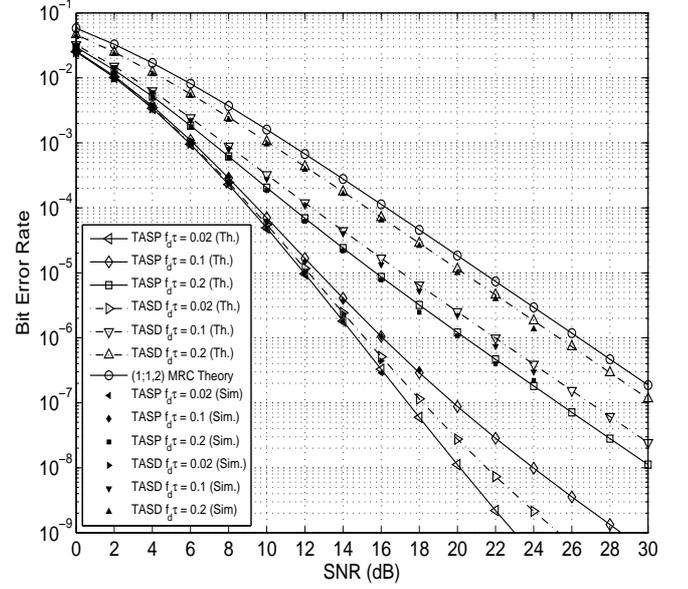


Fig. 3: BER performance comparison of BPSK for TAS/MRC (2, 1; 2) with and without prediction at different delays, showing theoretical and simulation results.

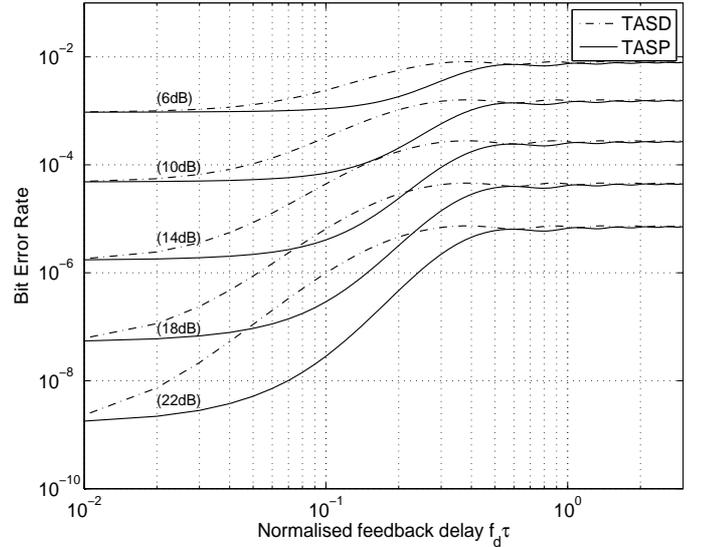


Fig. 4: Feedback delay tolerance of BER in BPSK of (2, 1; 2) scheme for different SNRs of 6, 10, 14, 18, 22 dB with and without prediction.

Fig. 3 compares the BER performance of BPSK, in the TAS/MRC system, with and without prediction, for small to severe normalised feedback delays of 0.02, 0.1 and 0.2. In order to verify eqn. (30) we also plot simulation results. For small $\bar{\gamma}$ and small delay, the performance of both schemes is nearly equal. For a $1e-3$ BER, compared to TASD, TASP offers a gain of 2dB at 0.1 delay, and 3dB at 0.2 delay. Similarly for a $1e-5$ BER, the gains are about 4dB for 0.1 and 0.2 feedback delay; while for lower BERs the gains are found to be 7dB and 5dB respectively. At large delays, as expected, the TASD system behaves like an open loop (1, 1; 2) MRC system: as seen in the graph, TASD tends towards the (1, 1; 2) MRC BER curve, while the TASP system also tends

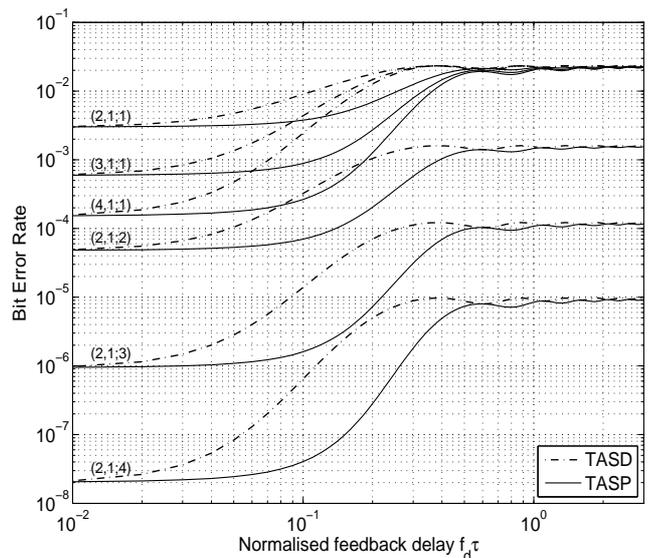
to the (1, 1; 2) MRC BER at a slower rate. These observations also indicate that there is an increase in degradation rate w.r.t delay for the TASP/D case.

The degradation effect is seen more clearly in Fig. 4 which plots BER against normalised feedback delay for five different average receive SNRs in a (2, 1; 2) configuration. In comparison to TASP for a delay of 0.02, the performance of the TASP system is almost as good as the system without any feedback delay. For small SNRs and small delays, the performance of TASP and TASP are quite similar, with additive noise dominating performance. TASP degradation begins at around 0.03 and higher, while TASP holds BER steady until about 0.1 delay. At higher SNRs of around 22 dB, TASP experiences degradation as early as 0.01 delay, while TASP can tolerate up to about 0.04. For small SNRs, in both TASP/P, degradation is more pronounced at larger feedback delays than at smaller ones. This effect can be explained by looking at the power correlation coefficient ρ_p in the BER equations. For TASP, $\rho_p = J_0^2(2\pi f_d \tau)$ [19], which falls off rapidly for increasing values of τ . By comparison, the power correlation coefficient for the TASP case is equal to $\mathbf{r}^H \mathbf{R}^{-1} \mathbf{r}$, which is a quadratic function, with shallower slope at smaller values of delay compared to the coefficient in TASP. The comparison of the correlation coefficients for both cases will be explored in Section VI.

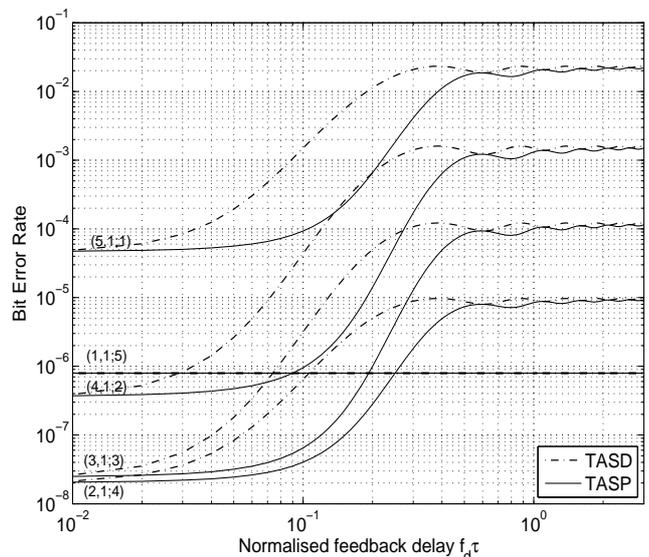
Fig. 5 plots the degradation of BER for different setups of TAS evaluated at a receive SNR $\bar{\gamma}$ of 10 dB, chosen so as to yield a range of BER values. Even small delays starting from 0.02 in the TASP arrangement cause significant deterioration in BER. By contrast the predictive scheme sustains performance almost unchanged out to around delays of 0.1. Relating this to a real system, at a carrier frequency of 900 MHz at a walking speed of 1 m/s, a time delay of not more than 6 ms can be tolerated by TASP. TASP, by contrast, can withstand up to 33 ms. Similarly for vehicles moving at 27.7 m/s (100 km/h), TASP tolerates up to 0.72 ms while TASP can tolerate about 3.6 ms of delay.

Fig. 5(a) looks at the influence of increasing N_t keeping N_r fixed and vice-versa at different delays. For increasing N_r , the BER decreases considerably at smaller delays, while at extreme delays, it gets capped to a value corresponding to its open loop MRC configuration. The BER improvement due to increasing N_t is less pronounced for higher delays in the TASP system, while the TASP system can help to harness this gain. It is also noteworthy that, for the TASP case, systems with lower BER requirement are more sensitive to feedback delay. The effects of adding extra diversity are also discussed in the following sections on outage probability and average fading gains.

Fig. 5(b) shows the BER at $\bar{\gamma} = 10$ dB for different systems of equal $N_r + N_t$. Such a plot can aid in choosing the total number of antennas used in a communications system for maximum performance w.r.t BER, at any given delay. For example, with prediction, in the region of $0.01 < f_d \tau < 0.25$, for $\bar{\gamma} = 10$ dB, the (2, 1; 4) setup outperforms all other schemes, with the same number of antennas, in terms of BER. For $f_d \tau \geq 0.25$, the non-selective (1, 1; 5) MRC setup performs best. However, when prediction is not employed,



(a)



(b)

Fig. 5: Feedback delay tolerance of BER in BPSK at $\bar{\gamma} = 10$ dB for different TAS/MRC configurations with and without prediction. (a) Influence of increasing N_r and N_t . (b) Different setups with $N_r + N_t = 6$.

(2, 1; 4) outperforms the other non-predictive setups in the region of $0.01 < f_d \tau < 0.11$.

IV. OUTAGE PROBABILITY

The outage probability P_{out} at any SNR $\bar{\gamma}$ is defined as the probability that the instantaneous capacity $C_{\bar{\gamma}}$ (bits/s/Hz) is less than a given capacity rate R and is given as $P_{out} = Pr\{C_{\bar{\gamma}} < R\}$ [31], which, in a block fading channel essentially is:

$$\begin{aligned} P_{out}(\bar{\gamma}, R) &= Pr\{\log_2(1 + \gamma_{max}) < R\} \\ &= Pr\{\gamma_{max} < (2^R - 1)\} \end{aligned} \quad (31)$$

Setting $z = (2^R - 1)$, outage probability is given as:

$$\begin{aligned}
 P_{out}(z) &= \int_0^z f_{\gamma_{max}}(\gamma) d\gamma \\
 &= \frac{N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) \\
 &\cdot (N_r + j - 1)! \sum_{k=0}^j \binom{j}{k} \frac{\rho_p^k (1 - \rho_p)^{j-k}}{[(i(1 - \rho_p) + 1)]^j} \\
 &\cdot \left\{ \frac{1}{(i+1)^{N_r+k}} - \exp\left(-\frac{(i+1)(z/\bar{\gamma})}{[i(1 - \rho_p) + 1]}\right) \right. \\
 &\cdot \left. \sum_{l=0}^{k+N_r-1} \frac{[1/(N_r + k - 1 - l)!](z/\bar{\gamma})^{N_r+k-l-1}}{(i+1)^{l+1}[i(1 - \rho_p) + 1]^{N_r+k-l-1}} \right\}
 \end{aligned} \tag{32}$$

where the integral in (A-4) has been utilized. As a check, consider the $(N_t, 1; 1)$ case: here $\eta_1(i, j) = 1$. Hence,

$$\begin{aligned}
 P_{out}(z) &= N_t \sum_{i=0}^{N_t-1} \frac{(-1)^i}{(i+1)} \binom{N_t-1}{i} \\
 &\cdot \left[1 - \exp\left(-\frac{(i+1)z}{[i(1 - \rho_p) + 1]\bar{\gamma}}\right) \right]
 \end{aligned} \tag{33}$$

Using $\frac{N_t}{i+1} \binom{N_t-1}{i} = \binom{N_t}{i+1}$ and $\sum_0^{N_t} (-1)^i \binom{N_t}{i} = 0$, when $\rho_p = 1$ for full prediction, (33) reduces to $P_{out}(z) = (1 - \exp(-z/\bar{\gamma}))^{N_t}$ which, for high SNR, has a diversity order of N_t . Similarly, when $\rho_p = 0$, $P_{out}(z)$ becomes $1 - \exp(-z/\bar{\gamma})$ which is the simple SISO case with diversity reducing to unity – verifying eqn. (32). Similar analysis can be done for higher order configurations. Fig. (6) shows the outage probability plots for different setups of TASP/D, at $\bar{\gamma} = 10$ dB with $R = 1$ bit/s/Hz. These reveal how the systems behave for different degrees of fading correlation and number of antennas. Simulation results are also shown to verify the equations.

Fig. 6(a) shows outage plots for $(N_t, 1; 2)$, and the dependency of increasing N_t having 2 receive antennas for TASP/D. For TASP, increasing N_t decreases P_{out} considerably for smaller delays, however for $N_t = 4$, degradation begins earlier than 0.01 delay. At larger delays, the effect of increasing N_t offers little improvement as seen around $f_d\tau = 0.1$. The TASP setup however tolerates delays up to around $f_d\tau = 0.05$, and behaves in a similar fashion to TASP for greater delays; while at extreme delays, as expected, P_{out} tends to a limit corresponding to the open loop MRC configuration of $(1, 1; 2)$ for both cases. It is also seen that for high order schemes, the degradation rate is higher for higher delays; whereby $(4, 1; 2)$ degrades faster than a $(3, 1; 2)$ scheme. Fig. 6(b) is plotted for increasing N_r as $(2, 1; N_r)$ and behaves in a similar fashion except being more delay tolerant compared to $(N_t, 1; 2)$. For increasing values of N_r , P_{out} decreases considerably at smaller delays. Similar to the previous case, for extreme delays, each setup tends to its corresponding $(1, 1; N_r)$ MRC configuration. The setups here have the same number of antennas as in Fig. 6(a), with MRC being the dominant factor in providing increased array gains. They also having the advantage of decaying at a slower rate as delay increases. Thus a $(2, 1; 4)$ setup may be better than a $(4, 1; 2)$ setup.

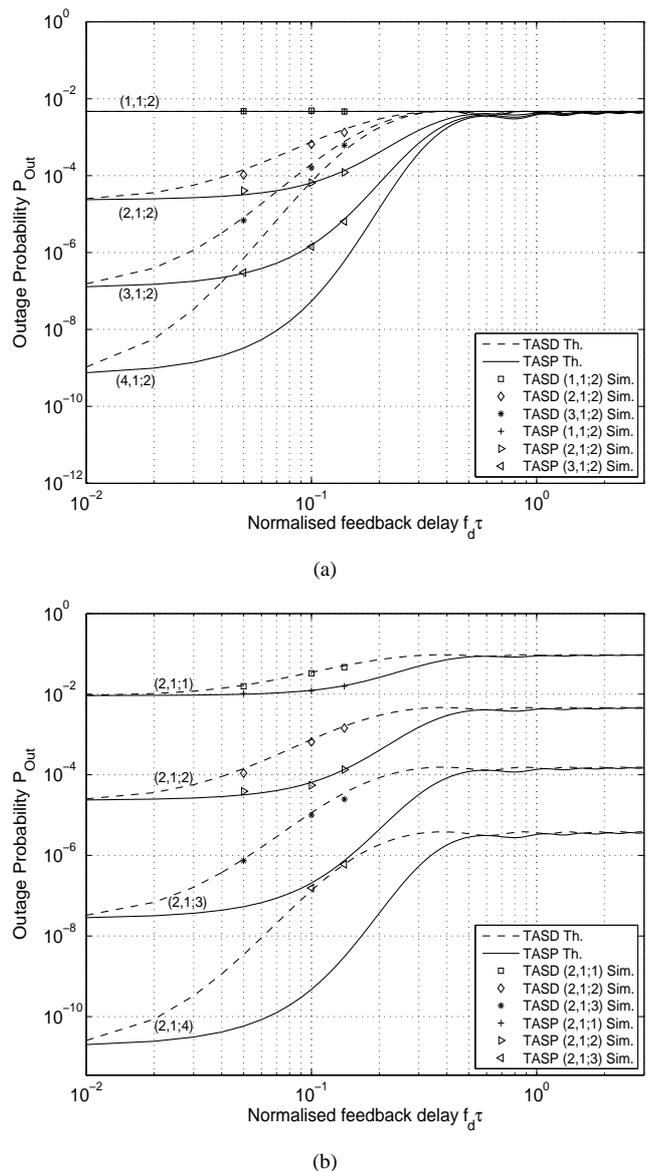


Fig. 6: Outage probability comparisons at $\bar{\gamma} = 10$ dB SNR between TASP/D schemes with fixed N_r and increasing N_t and vice-versa, at $R = 1$ bit/s/Hz, showing theoretical and simulation results.

V. FADING STATISTICS OF THE OUTPUT SNR

We evaluate and compare the different TAS configurations based on moments of the signal amplitude γ_{max} . Closed form expressions for average SNR gain, signal variance and amount of fading (AF) are derived. Such measures could find use in comparing system performance of the existing TASP/MRC scheme with schemes that use different fading models, and possibly across other MIMO architectures, where closed form solutions of common performance metrics such as BER and outage probability are not readily available. The average SNR gain can be used to quantify and compare across different TAS/MRC schemes. Specifically here we compare the gains of different TAS/MRC schemes in the presence and absence of the predictor.

The average SNR gain of the TASP $(N_t, 1; N_r)$ system is

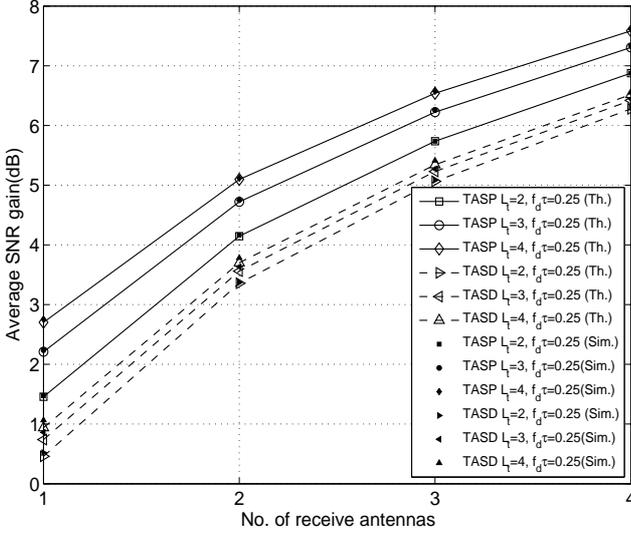


Fig. 7: Comparison of normalised average SNR gain $\bar{\gamma}_{TAS}/\bar{\gamma}$ as a function of number of receive and transmit antennas at specific feedback delays of $f_d\tau = 0.25$ with and without prediction. Theoretical results are verified by simulations.

denoted as $\bar{\gamma}_{TAS} = \mathbb{E}[\gamma_{max}]$ and is given by:

$$\mathbb{E}[\gamma_{max}] = \int_0^\infty \gamma f_{\gamma_{max}}(\gamma) d\gamma \quad (34)$$

For a $(N_t, 1; N_r)$ scheme, using (26),

$$\begin{aligned} \mathbb{E}[\gamma_{max}] &= \frac{N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ &\cdot \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) j! \frac{[(1 - \rho_p)\bar{\gamma}]^j}{[i(1 - \rho_p) + 1]\bar{\gamma}^{j+N_r}} \\ &\cdot \int_0^\infty \gamma^{N_r} \exp\left(-\frac{(i+1)\gamma}{[i(1 - \rho_p) + 1]\bar{\gamma}}\right) \\ &\cdot L_j^{(N_r-1)}\left(-\frac{[\rho_p/(1 - \rho_p)]\gamma}{[i(1 - \rho_p) + 1]\bar{\gamma}}\right) d\gamma \end{aligned} \quad (35)$$

then expanding the Laguerre polynomial, and integrating with the help of (A-4), we have:

$$\begin{aligned} \mathbb{E}[\gamma_{max}] &= \frac{\bar{\gamma}N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ &\cdot \sum_{j=0}^{i(N_r-1)} \frac{\eta_{N_r}(i, j) j! (j + N_r - 1)!}{[i(1 - \rho_y) + 1]^{(j-1)}} \\ &\cdot \sum_{k=0}^j \frac{\rho_y^k (1 - \rho_y)^{j-k} (k + N_r)}{(j - k)! k! [(i + 1)]^{N_r+k+1}} \end{aligned} \quad (36)$$

Fig. 7 plots the gains of the predictor and delayed scheme for a fixed delay of 0.25, verified by simulation results which are also plotted. Note that for the delayed scheme, the diversity benefits from adding an additional transmit antenna diminish as delay gets larger. Prediction is therefore crucial in maintaining transmit diversity. When prediction is not used, it is beneficial to invest in receive diversity. For larger N_r , the gain increase from adding an extra transmit antenna is less. Fig.

8 compares post-processing SNR gains achieved by different TASD and TASP configurations for a wide range of delays. As before both TASP and TASD show a decrease in gains as delay increases. Depending upon the amount of gain required at a given delay, the predictor can attempt to keep gain constant. This has a direct bearing on the amount of power put into the pilot symbols. In all graphs, increasing N_r keeping N_t fixed increases the gain, but not linearly. The gain in going from one receive antenna to two is greater than from going from two to three: in general increasing N_r gives diminishing returns of post-processing SNR gain. Looking across the graphs, at given delay, any gain improvement obtained by increasing N_t (keeping N_r constant) is less than from increasing N_r (keeping N_t constant). Depending on the required gain, a system designer can choose the most appropriate configuration based upon available resources. For example, as seen from Fig. 8, at a delay of 0.1, it would be wiser to invest in a $(2, 1; 4)$ configuration with prediction than to use a $(4, 1; 3)$ without prediction under the given conditions, since the $(2, 1; 4)$ is better protected through MRC than $(4, 1; 3)$.

Although this performance criterion is useful in comparing different TAS schemes, it does not capture all the diversity benefits. The fading fluctuations are also dependent on the second order statistics, namely the signal amplitude variance, or the second moment of the output SNR. The second moment of the random variable γ_{max} can be similarly obtained as:

$$\begin{aligned} \mathbb{E}[\gamma_{max}^2] &= \frac{\bar{\gamma}^2 N_t!}{(N_r - 1)!} \sum_{i=0}^{N_t-1} \frac{(-1)^i}{i!(N_t - 1 - i)!} \\ &\cdot \sum_{j=0}^{i(N_r-1)} \frac{\eta_{N_r}(i, j) j! (j + N_r - 1)!}{[i(1 - \rho_p) + 1]^{j-2}} \\ &\cdot \sum_{k=0}^j \frac{\rho_p^k (1 - \rho_p)^{j-k} (k + N_r)(k + N_r + 1)}{(j - k)! k! [(i + 1)]^{N_r+k+2}} \end{aligned} \quad (37)$$

The amount of fading (AF), associated with the fading PDF of strength γ_{max} , is derived as:

$$AF = \frac{\text{var}[\gamma_{max}]}{\mathbb{E}[\gamma_{max}]^2} = \frac{\mathbb{E}[\gamma_{max}^2] - \mathbb{E}[\gamma_{max}]^2}{\mathbb{E}[\gamma_{max}]^2} \quad (38)$$

This performance measure was introduced by Charash [32] as a measure of fading severity and is independent of the average SNR per symbol $\bar{\gamma}$. Thus, two systems with the same AF may be considered to fade similarly. Win and Winters [33] used the square root of the ratio of the variance of the combined output SNR to its squared mean, to assess the effectiveness of a hybrid diversity combining scheme in the presence of Rayleigh fading. This performance measure is also equal to the square-root of the AF . For example, in comparing the fading severity of $(2, 1; 4)$ and $(3, 1; 3)$ under no delay, AF evaluated as per eqn. (38), is approximately -8 dB for both, and one may expect them to have a similar BER performance. As a check, Fig. 5(b) verifies this, and both schemes have nearly equal BER for small delays at $\bar{\gamma} = 10$ dB. However, AF is not always a good indicator of BER performance (although may be better than average SNR). Higher order statistics of the

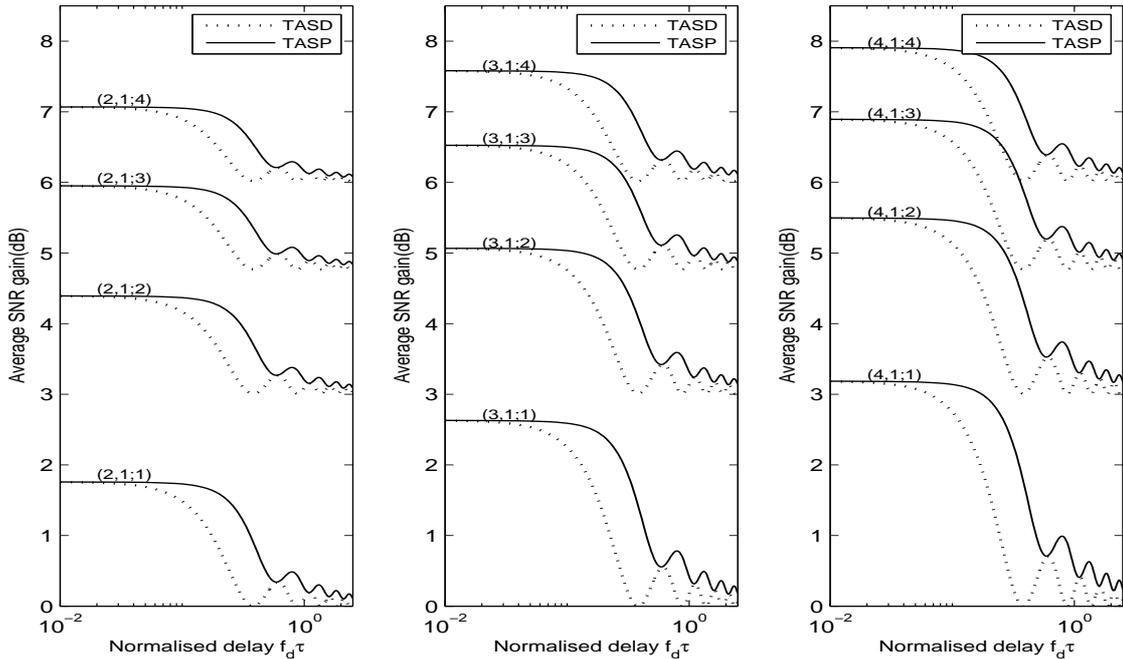


Fig. 8: Normalised average SNR gains $\bar{\gamma}_{TAS/\bar{\gamma}}$ of different arrangements with and without prediction for a wide range of delays.

fading SNR can also be derived, and used for further system analysis and design.

VI. INVESTIGATION OF SYSTEM PARAMETERS AND OPERATING POINTS

A. Influence of system parameters on channel correlation

The correlation coefficient which governs BER of the TASP/MRC system, generally depends on a number of parameters such as FIR (finite impulse response) predictor length, pilot SNR, symbol rate, channel sampling frequency (location of pilot symbol), training length, Doppler frequency and feedback delay. At higher bit-rates, to limit the number of filter coefficients, it is beneficial to reduce pilot insertion frequency. This sub-sampling frequency (SSF) is usually kept at multiples of the Doppler frequency ($SSF \geq 2f_d$), satisfying the sampling theorem. Having a relatively high over-sampling rate may mean clearer channel estimates but also poorer long range prediction with a fixed filter order, and vice-versa. We shall look at the influence of these parameters one by one.

1) *Influence of E_p/N_0 and SSF:* As mentioned, our simulations and numerical evaluations use a symbol period of $T = 1\mu s$, with Doppler of 100 Hz. To find the optimal sampling or adaptation rate in the presence of noise, we must determine the optimal delay spacing at any given E_p/N_0 and fixed filter length. We fix the filter length to a nominal value of $L = 5$ in our simulations, since having a small filter length reduces the computational burden of receiver end processing; it also reduces errors in estimation of the coefficients themselves. Next we plot channel correlation w.r.t. delay spacing for different E_p/N_0 values and find the delay spacing at which maximum correlation occurs, as shown in Fig. 9(a). For small prediction ranges, we could use a low sampling rate, since sufficient channel correlation exists over longer periods. For

larger prediction ranges, the adaptation rate varies with the SNR as shown. For lower SNR it is beneficial to keep a larger block length, and vice-versa. Note however, no prior noise smoothing has been done and it is possible that any out-of-band noise gets amplified, decreasing correlation. Thus higher order filters can efficiently perform better prediction by suppressing noise. Also increasing the filter tap length reduces the optimal sampling frequency, since more predictor coefficients can provide better long term prediction.

2) *Influence of E_p/N_0 and predictor filter length:* At a given $\bar{\gamma}$, normalised feedback delay and filter delay spacing, we can determine the dependency of the correlation and filter length with increasing E_p/N_0 . For smaller delays, shorter training and prediction-filter lengths are able to assure sufficient correlation. However at increased delays, greater training length offers more correlation gain than increasing filter length, since clearer channel estimates offer greater prediction performance. This is shown in Fig. 9(b) where two sets of curves for normalised delays of 0.02 and 0.1 are plotted with different E_p/N_0 and increasing filter lengths. We can see that, for a larger delay, the effect of changing pilot power is more pronounced. Maintaining the same parameters, and improving the E_p/N_0 , naturally increases correlation again because of better channel estimates. For small filter lengths like $L = 5$, the absolute change in correlation value upon increasing pilot power is greater at larger delays than at small delays. This trend is also seen for higher order filters, but reducing slowly with increasing L . These results and observations corroborate findings in [6], [9]. Depending on the filter order and the amount of power of the pilot, the predictor can always be made to offer greater correlation. These system parameters influence the correlation for the TASP system which increases or decreases at a fixed delay.

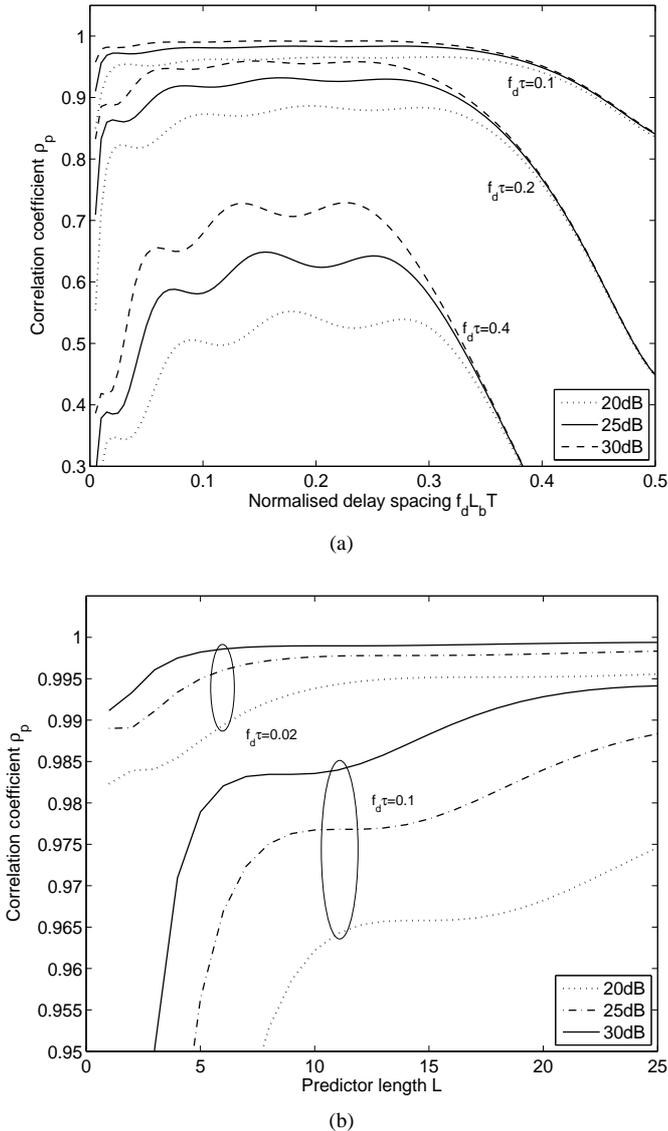


Fig. 9: (a) Influence of filter tap spacing and E_p/N_0 at normalised feedback delays of 0.1, 0.2 and 0.4 on power correlation with a five tap filter (b) Influence of predictor length and E_p/N_0 on power correlation at fixed delays of 0.02 and 0.1 and $L_b = 100$.

However for better prediction at greater delays, the block length may have to be increased, which may in turn account for channel variability and estimation errors during demodulation at the receiver.

3) *Choice of block length (L_b):* As noted above the choice of block length or pilot spacing will also be determined by the capabilities of the noise smoother and interpolation that happen during channel estimation at the receiver. For Rayleigh flat fading, we need to estimate a single tap for each of the $(N_t \times N_r)$ channels. In general, for a continuous flat fading channel, the estimation error will be contributed by the Doppler variance and noise [34]. For quasi-static or block fading, if we neglect the noise introduced by Doppler variance, then the MSE or the variance of the estimated channel will be $N_0/(L_t E_s)$ where L_t is the number of training symbols

for each transmit antenna (where each is transmitted round-robin), over the entire frame. As a rule of thumb, pilot symbols are inserted every $1/(100 f_d)$ s to ensure that the channel changes negligibly over this period. As seen from Fig. 9a when the prediction distance ($f_d \tau$) is relatively small, for example, $f_d \tau = 0.1$, this condition can be easily met, without significantly lowering the correlation coefficient or predictor performance. Fig. 10 shows the comparison of correlation coefficients for the delayed and predicted cases with $L_b=100$.

B. Operating point of TASP and trade-off point with TASD

Depending on threshold output BER, BER_T for a given configuration, and knowing the normalised delay that must be tolerated, the graphs presented in the previous subsection may be used to first determine the threshold correlation, and from that, the required filter order and E_p/N_0 . This is worst-case design: when delay is less, BER will naturally improve. It may happen that TASP performs well enough to achieve the target BER at that delay, in which case prediction is not necessary. Otherwise, TASP can be chosen to improve performance. It is interesting to determine the point at which TASP becomes advantageous over TASD, through setting $\rho_p = \rho_d$ (i.e when the correlation of both schemes are equal). A few such intersecting points can be seen by following the TASP curve in Fig. 10.

To explore further, Fig. 11 plots the equal-correlation points for different delays against pilot SNR with varying filter order. Thus, knowing the feedback delay of a given system, it is possible to determine whether TASP of a given filter order outperforms TASD. If one wants to improve system performance beyond this point (i.e. in terms of increasing correlation or BER) then TASP should be used. Thus it can be seen that even when TASP has a clear, but outdated, estimate available, TASP can perform better by predicting into the future from noisy channel estimates.

The setups of Fig. 12 further highlight design tradeoffs, for $(3, 1; 2)$ and $(2, 1; 4)$ arrangements respectively: they exhibit crossover points between the TASP and TASD curves for each arrangement. It is at these points that the performance of both TASP and TASD are equal. For example at a delay of 0.08 and $E_p/N_0 = 15$ dB (top solid curves), the BER of TASP and TASD equate. This can be seen in both $(3, 1; 2)$ and $(2, 1; 4)$, which provide BERs of 5×10^{-5} and 3×10^{-7} respectively. From Fig. 11 we can see that both schemes offer a correlation of almost 0.9 at the given delay, and that TASP requires a 5 tap filter, found from the intersection of $E_p/N_0 = 15$ dB and $f_d \tau = 0.08$ which lies between the $L = 4$ and $L = 5$ curves. Note that, if better BER is required, further improvement is possible with TASP through increasing L or E_p/N_0 . However from Fig. 12, we note that at extremes of very small ($< .02$) and very large normalised delays (≥ 0.4), the bundle of TASP BER curves approach asymptotes. This implies that in those extreme regions, the effect of increases in E_p/N_0 will be minimal.

VII. CONCLUSION

This paper has explored the issue of transmit antenna selection based on an index fed back from an MRC receiver.

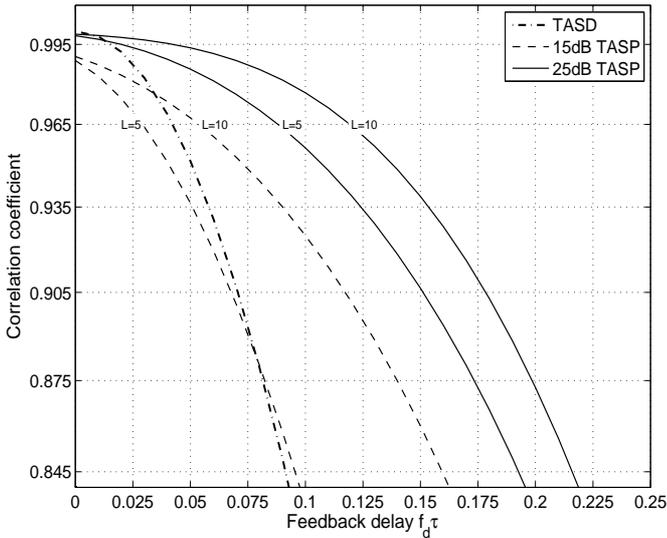


Fig. 10: Correlation coefficient comparison for TASP, and TASP for E_p/N_0 of 15 dB and 25 dB with two different filter lengths of 5 and 10.

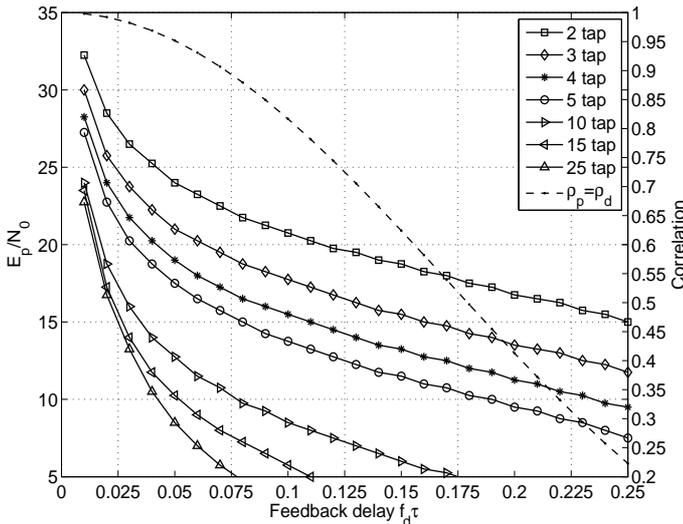


Fig. 11: Equal correlation plot for TASP and TASP, where at all points $\rho_p = \rho_d$ with the corresponding E_p/N_0 (for TASP case) against feedback delay for various filter orders. The points on the dashed line corresponds to the correlation value at any feedback delay, read off from the RH axis.

The system has been shown to be sensitive to normalised switching delay (thus reducing channel correlation), which may arise from processing time in the reverse communications path and mobility between the transmitter and the receiver. In an effort to mitigate against the effects of normalised delay, a power predictor based on Wiener filtering of past outdated CSI, is introduced at the receiver, to predict a best transmit antenna for future transmission slots.

The effects and benefits of channel prediction on several important performance metrics were studied and closed form expressions derived for each of them, all verified by simulation. Prediction was found useful in maintaining transmit diversity and improving both BER and outage performance. The inter-relation between systems parameters such as pre-

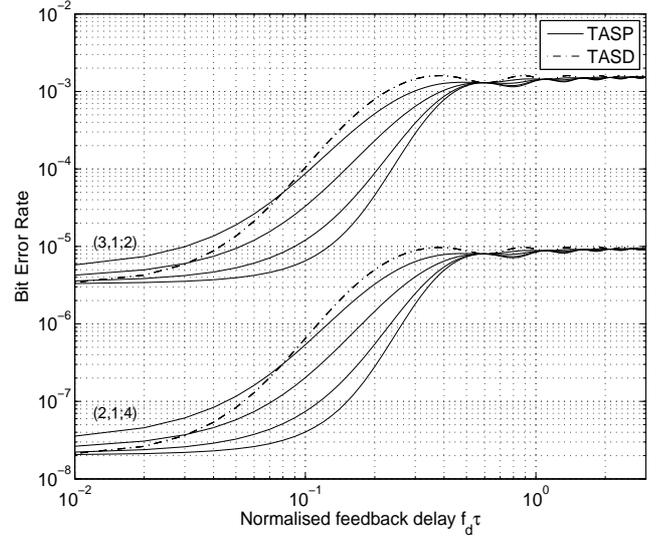


Fig. 12: Effect of increasing E_p/N_0 for TASP at a filter length of $L = 5$ on BER performance at $\bar{\gamma} = 10$ dB for two different configurations. The solid black curves represent TASP at different $E_p/N_0 = 15, 20, 25, 30$ dB from top to bottom, in two example systems.

dictor length, pilot SNR and block length on long range prediction has been explored, and their influence on the channel correlation coefficient noted (where pilot power can improve predictor performance more than increasing the filter order). The system has been shown capable of alleviating much of the performance loss associated with outdated transmitter selection knowledge, even for delays which would have caused non-predictive TAS to be ineffective. At any given delay, a minimum threshold pilot SNR and filter order required for the predictive system to just overtake the non-predictive case was presented graphically. The predictive approach is found beneficial in combating diversity loss in systems, which would otherwise substantially degrade for normalised delays greater than 0.02. For most configurations, it was shown that the use of only a 5 tap filter, per transmit-receive path, was effective in sustaining BER out to about 0.1 normalised delay.

APPENDIX

A. The coefficients $\eta_{N_r}(i, j)$ that appear in eqn. (26) were obtained from the following expansion [35]:

$$\left[\sum_{j=0}^{N_r-1} z^j / j! \right]^i = \sum_{j=0}^{i(N_r-1)} \eta_{N_r}(i, j) z^j \quad (\text{A-1})$$

with the coefficients:

$$\eta_{N_r}(i, 0) = 1, \eta_{N_r}(i, 1) = i$$

$$\eta_{N_r}(i, i(N_r - 1)) = 1 / ((N_r - 1)!)^i$$

$$\eta_{N_r}(i, j) = (1/j) \sum_{l=1}^{J_0} (l(i+1) - j) / l! \eta_{N_r}(i, j-l)$$

$$J_0 = \min(j, N_r - 1), 2 \leq j \leq i(N_r - 1) - 1$$

B. The Laguerre polynomial is given by [29]:

$$L_n^b(x) = \sum_{k=0}^n (-1)^k \binom{n+b}{n-k} \frac{x^k}{k!} \quad (\text{A-2})$$

where $\binom{n}{k}$ is a binomial coefficient.

$$\begin{aligned} C. \int_0^\infty x^{n+\frac{1}{2}} e^{-\alpha x} J_v(2\beta\sqrt{x}) dx \\ = n! \beta^v e^{-\frac{\beta^2}{\alpha}} \alpha^{-(n-v-1)} L_n^v\left(\frac{\beta^2}{\alpha}\right) \end{aligned} \quad (\text{A-3})$$

where $J_v(\cdot)$ is the Bessel function of the first kind of order v .

D. The lower incomplete gamma function is given by:

$$\int_0^u x^{v-1} e^{-\mu x} dx = \frac{(v-1)!}{\mu^v} \left(1 - e^{-\mu u} \sum_{l=0}^{v-1} \frac{(\mu u)^l}{l!}\right) \quad (\text{A-4})$$

when v is an integer.

E. If $x_{i=1,2,3,4}$ are complex normal random variables with zero mean, with $\mathbb{E}[x_i x_j^*] = C_{ij}$ then [36]:

$$\mathbb{E}[x_1 x_2 x_3 x_4] = C_{12} C_{34} + C_{13} C_{24} + C_{14} C_{23} \quad (\text{A-5})$$

REFERENCES

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge Univ. Press, 2003.
- [2] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.
- [3] M. Simon and M. Alouini, *Digital communication over fading channels*. Wiley-Interscience, 2005.
- [4] Z. Chen, J. Yuan, and B. Vucetic, "Analysis of Transmit Antenna Selection/Maximal-Ratio Combining in Rayleigh Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 4, pp. 1312–1321, July 2005.
- [5] A. Wittneben, "Analysis and comparison of optimal predictive transmitter selection and combining diversity for DECT," in *Proceedings. Global Telecommunications Conference, 1995. GLOBECOM '95., IEEE*, Nov. 1995, pp. 1527–1531 vol.2.
- [6] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Communications Letters*, vol. 2, no. 9, pp. 254–256, Sep. 1998.
- [7] S. Hu, T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Transmitter antenna diversity and adaptive signaling using long range prediction for fast fading DS/CDMA mobile radio channels," in *Proceedings. Wireless Communications and Networking Conference, 1999. WCNC. 1999 IEEE*, 1999, pp. 824–828 vol.2.
- [8] S. Guncavdi and A. Duel-Hallen, "Performance analysis of space-time transmitter diversity techniques for WCDMA using long range prediction," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 40–45, Jan. 2005.
- [9] T. Ekman, "Prediction of mobile radio channels, modeling and design," Ph.D. dissertation, Uppsala University, Sweden, 2002.
- [10] T. Ekman, M. Sternad, and A. Ahlén, "Unbiased power prediction of Rayleigh fading channels," in *Proceedings. Vehicular Technology Conference, VTC 2002-Fall.*, 2002, pp. 280–284 vol.1.
- [11] G. Øien, H. Holm, and K. Hole, "Impact of channel prediction on adaptive coded modulation performance in Rayleigh fading," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 3, pp. 758–769, May 2004.
- [12] S. Zhou and G. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?" *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1285–1294, July 2004.
- [13] K. Baddour, C. Squires, and T. Willink, "Mobile Channel Prediction with Application to Transmitter Antenna Selection for Alamouti Systems," in *Proceedings. Vehicular Technology Conference, 2006. VTC-2006 Fall 2006.*, Sept. 2006, pp. 1–6.
- [14] S. Thoen, L. Van der Perre, B. Gyselinckx, and M. Engels, "Performance analysis of combined transmit-SC/receive-MRC," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 5–8, Jan. 2001.
- [15] J. Tang and X. Zhang, "Error probability analysis of TAS/MRC-based scheme for wireless networks [point-to-point link example]," in *Proceedings. Wireless Communications and Networking Conference, 2005 IEEE*, March 2005, pp. 877–882 vol. 2.
- [16] X. Cai and G. Giannakis, "Performance analysis of combined transmit selection diversity and receive generalized selection combining in Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1980–1983, Nov. 2004.
- [17] S. Prakash and I. McLoughlin, "Predictive Transmit Antenna Selection with Maximal Ratio Combining," in *Proceedings. Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, Nov.30 2009-Dec. 4 2009, pp. 1–6.
- [18] J. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio]," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, Nov. 1991.
- [19] W. Jakes, *Microwave mobile communications*. Wiley-IEEE Press, 1994.
- [20] G. Stüber, *Principles of mobile communication*. Kluwer Academic Pub, 2001.
- [21] A. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.
- [22] M. Alouini and A. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Wireless Personal Communications*, vol. 13, no. 1, pp. 119–143, May 2000.
- [23] X. Tang, M.-S. Alouini, and A. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Transactions on Communications*, vol. 47, no. 12, pp. 1856–1864, Dec. 1999.
- [24] D. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 844–855, Jun. 1999.
- [25] S. Zhou and G. Giannakis, "Adaptive modulation for multi-antenna transmissions with channel mean feedback," *Communications, 2003. ICC '03. IEEE International Conference on*, vol. 4, pp. 2281–2285 vol.4, May 2003.
- [26] X. Cai and G. Giannakis, "Adaptive PSAM accounting for channel estimation and prediction errors," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 246–256, Jan. 2005.
- [27] H. David and H. Nagaraja, *Order statistics*. Wiley-Interscience, 2003.
- [28] I. Gradshteyn and I. Ryzhik, *Table of integrals, Series, and Products*, vol. 22.
- [29] M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, 1972.
- [30] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [31] I. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–596, Nov./Dec 1999.
- [32] U. Charash, "Reception Through Nakagami Fading Multipath Channels with Random Delays," *IEEE Transactions on Communications*, vol. 27, no. 4, pp. 657–670, Apr. 1979.
- [33] M. Win and J. Winters, "Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Transactions on Communications*, vol. 47, no. 12, pp. 1773–1776, Dec. 1999.
- [34] Q. Sun, D. Cox, H. Huang, and A. Lozano, "Estimation of continuous flat fading MIMO channels," in *Proceedings. Wireless Communications and Networking Conference, 2002. WCNC2002. 2002 IEEE*, Mar. 2002, pp. 189–193 vol.1.
- [35] G. Fedele, "N-branch diversity reception of many DPSK signals in slow and nonselective Nakagami fading," *European Transactions on Telecommunications*, vol. 7, no. 2, Apr. 1996.
- [36] A. Papoulis and S. Pillai, *Probability, random variables and stochastic processes*. McGraw-Hill Higher Education, 2002.