Designing PSAM Schemes: How Optimal are SISO Pilot Parameters for Spatially Correlated SIMO?

Xiangyun Zhou, Tharaka A. Lamahewa, Parastoo Sadeghi and Salman Durrani
College of Engineering and Computer Science, The Australian National University, Canberra 0200 ACT, Australia
Email:{xiangyun.zhou, tharaka.lamahewa, parastoo.sadeghi, salman.durrani}@anu.edu.au

***Abstract*—We study the design parameters of pilot-symbol-assisted modulation (PSAM) schemes for spatially correlated single-input multiple-output (SIMO) systems in time-varying Gauss-Markov flat-fading channels. We use an information capacity lower bound as our figure of merit. We investigate the optimum design parameters, including the ratio of power allocated to the pilots and the fraction of time occupied by the pilots, for SIMO systems with different antenna sizes and with spatial channel correlation. Our main finding is that by optimally designing the training parameters for single-input single-output (SISO) systems, the same parameters can be used to achieve near optimum capacity in both spatially independent and correlated SIMO systems for the same fading rate and signal-to-noise ratio (SNR). In addition, we show that spatially independent channels give the lowest capacity at sufficiently low SNR. These findings provide insights into the design of practical PSAM systems.**

I. Introduction

Channel estimation is crucial for reliable high data rate transmission in wireless communications with coherent detection. Pilot-symbol-assisted modulation (PSAM) has been used in many practical communication systems, e.g. in Global System for Mobile Communications (GSM) [1], to assist estimation of unknown channel parameters. In PSAM schemes, training symbols are inserted into data blocks periodically to acquire the channel state information (CSI) [2]. However, the insertion of pilots also reduces the information capacity as less transmission resource is allocated to data. Therefore, trade-off analysis in PSAM parameter design is required on the resource allocation to pilots and data. Furthermore, the parameters of the wireless channel, such as the number of the channel inputs/outputs, the fading rate and the spatial correlations, add another level of complexity into the design problem.

Optimal PSAM designs for time-varying fading channels with low-pass Doppler spectra were studied in [3, 4], where Wiener filtering was used for non-causal channel estimation. In [3], the authors studied a lower bound on channel capacity and concluded that optimal sampling frequency of the fading process equals the Nyquist rate. On the other hand, the studies on channel capacity with ideal interleaving via Monte Carlo simulations in [4] showed that pilot symbols should be sent more frequently than the Nyquist rate.

More recently, studies on optimal PSAM design in single-input single-output (SISO) systems adopted a Gauss-Markov channel model which is an alternative model for time-varying fading channels [5–7]. With fixed ratio of pilot insertion, the authors in [5] found that allocating only one pilot per transmission block minimized the channel estimation error at the last data symbol in the block. From an information theoretic viewpoint, the authors in [6] investigated the power distribution among data symbols and showed that data symbols closer to the pilot symbols should have more power than those further away from the pilots. In [7], the authors assumed uniform power distribution among data symbols, and jointly optimized a channel capacity lower bound to find the optimum pilot power allocation ratio and pilot spacing.

For multiple-input multiple-output (MIMO) systems, the authors in [8] studied a lower bound on the information capacity in PSAM schemes for block fading channels, and derived the optimal pilot power allocation and optimal number of pilot symbols per transmission block. For slow bandlimited fading channels with maximum likelihood estimation, the authors in [9] found that the optimal pilot spacing is nearly independent of the number of receive antennas. However, studies in [9] assumed spatially independent channels with equal power allocation to pilot and data symbols.

The impact of channel spatial correlations on the capacity has been studied mainly in non-PSAM schemes. With the knowledge of the channel spatial covariance at the transmitter, correlations among transmit antennas increase the channel capacity when perfect CSI is present at the receiver [10]. On the other hand, the authors in [11] showed that correlations among the transmit/receive antennas always reduce capacity, assuming perfect CSI is available only at the receiver.

In this paper, we consider the optimal PSAM design from an information theoretic viewpoint for SIMO systems in time-varying Gauss-Markov channels. We investigate the following questions: Does channel spatial correlation at the receiver always reduce information capacity? Are optimal parameters for SISO systems also optimal for SIMO systems with spatially correlated channels? The main contributions of this paper are:

- In Section IV, we show that spatially independent SIMO channels result in the highest channel estimation error, and hence the lowest capacity at sufficiently low SNR.
- In Section V, we show that the optimum PSAM design parameters for SISO systems are very close to optimal for spatially independent SIMO systems for Gauss-Markov channels with the same fading rate and operating SNR.
- In Section V, we show that the optimal design for spatially independent channels are also near optimal for correlated channels, which is an extension of [9]. Based on the above results, we conclude that by optimally

978-1-4244-2644-7/08/$25.00 ©2008 IEEE

designing the training parameters for SISO systems, the same parameters can be used to achieve near optimum capacity in both spatially independent and correlated SIMO systems.

Throughout the paper, the following notations will be used: Boldface upper and lower cases denote matrices and column vectors, respectively. The matrix $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. $[\cdot]^*$ denotes the complex conjugate operation, and $[\cdot]^\dagger$ denotes the conjugate transpose operation. The notation $E\{\cdot\}$ denotes the mathematical expectation. $\mathrm{tr}\{\cdot\}$, $|\cdot|$ and $\mathrm{rank}\{\cdot\}$ denote the matrix trace, determinant and rank, respectively.

II. System Model

We consider a SIMO system with N_r receive antennas in time-varying flat-fading channels. After matched filtering, the received symbols at time index ℓ are given by

$$\boldsymbol{y}_\ell = \boldsymbol{h}_\ell x_\ell + \boldsymbol{n}_\ell, \tag{1}$$

where x_ℓ is the transmitted symbol, $\boldsymbol{y}_\ell$ is the $N_r \times 1$ received symbol vector, $\boldsymbol{n}_\ell$ is the $N_r \times 1$ noise vector with covariance matrix $\boldsymbol{R}_{\boldsymbol{n}} = E\{\boldsymbol{n}_\ell \boldsymbol{n}_\ell{}^\dagger\}$. The noise at each receive antenna is independent, identically distributed (i.i.d.) and zero-mean circularly symmetric complex Gaussian (ZMCSCG), each with variance $\sigma_{\boldsymbol{n}}^2$, i.e. $\boldsymbol{R}_{\boldsymbol{n}} = \sigma_{\boldsymbol{n}}^2 \boldsymbol{I}_{N_r}$. $\boldsymbol{h}_\ell$ is the $N_r \times 1$ channel vector with ZMCSCG entries. The spatial correlation of the channels is characterized by $\boldsymbol{R}_{\boldsymbol{h}} = E\{\boldsymbol{h}_\ell \boldsymbol{h}_\ell^\dagger\}$. In the case where the channels are spatially i.i.d., $\boldsymbol{R}_{\boldsymbol{h}} = \sigma_{\boldsymbol{h}}^2 \boldsymbol{I}_{N_r}$, where $\sigma_{\boldsymbol{h}}^2$ denotes the variance of each entry of $\boldsymbol{h}_\ell$.

The temporal correlation of the channels is modelled as a Gauss-Markov process:

$$\boldsymbol{h}_\ell = \alpha \boldsymbol{h}_{\ell-1} + \boldsymbol{w}_\ell, \tag{2}$$

where $\boldsymbol{w}_\ell$ is a ZMCSCG process noise with covariance matrix $\boldsymbol{R}_{\boldsymbol{w}} = E\{\boldsymbol{w}_\ell \boldsymbol{w}_\ell{}^\dagger\} = (1-\alpha^2)\sigma_{\boldsymbol{h}}^2 \boldsymbol{I}_{N_r}$. α is the temporal correlation coefficient given by $\alpha = J_0(2\pi f_D T_s)$, where J_0 is the zero-order Bessel function of the first kind, f_D is the Doppler frequency shift, and T_s is the transmitted symbol period. Therefore, $f_D T_s$ is the normalized fading rate. We assume that α is known and is constant over a large number of transmitted symbols.

A. Pilot Transmission Scheme

In PSAM schemes, the channel estimation is performed during pilot transmission. During data transmission the channels can be predicted based on the temporal correlation. From the previous studies on the optimal design of pilot insertion [5, 8], we know the optimal strategy for SIMO systems is to allocate one pilot per transmission block. Therefore, we assume each transmission block of T symbols consists of one pilot followed by $T-1$ data symbols. We denote the pilot spacing by $\eta = 1/T$. The average power per symbol is denoted by $\mathcal{E}$, and the power of pilot and data symbols are denoted by $\mathcal{E}_p$ and $\mathcal{E}_d$, respectively. We assume a fraction of γ of the total power budget is allocated to pilots. Hence, we have

$$\mathcal{E}_p = \gamma \mathcal{E} T = \frac{\gamma}{\eta}\mathcal{E} \quad \text{and} \quad \mathcal{E}_d = \frac{(1-\gamma)\mathcal{E}T}{T-1} = \frac{1-\gamma}{1-\eta}\mathcal{E}. \tag{3}$$

We denote the average received symbol SNR by $\rho = \sigma_{\boldsymbol{h}}^2 \mathcal{E}/\sigma_{\boldsymbol{n}}^2$. Using (3), the pilot and data symbol SNRs are given as

$$\rho_p = \frac{\sigma_{\boldsymbol{h}}^2 \mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2} = \frac{\gamma}{\eta}\rho \quad \text{and} \quad \rho_d = \frac{\sigma_{\boldsymbol{h}}^2 \mathcal{E}_d}{\sigma_{\boldsymbol{n}}^2} = \frac{1-\gamma}{1-\eta}\rho. \tag{4}$$

III. Channel Estimation

Due to the channel temporal correlation, we use the Kalman filter as an iterative linear minimum mean square error (LMMSE) estimator based on state space models. In particular, (1) and (2) are the observation equation and the state update equation, respectively. For channels with Gaussian statistics, the LMMSE estimator is the MMSE estimator.

During pilot transmission, i.e. $\ell = 1, T+1, ...$, the Kalman filter gain is given by [12]

$$\boldsymbol{K}_\ell = (\alpha^2 \boldsymbol{M}_{\ell-1} + \boldsymbol{R}_{\boldsymbol{w}}) x_\ell^* \left((\alpha^2 \boldsymbol{M}_{\ell-1} + \boldsymbol{R}_{\boldsymbol{w}})\mathcal{E}_p + \sigma_{\boldsymbol{n}}^2 \boldsymbol{I}_{N_r}\right)^{-1}, \tag{5}$$

and the channel estimate update equation is given by

$$\hat{\boldsymbol{h}}_\ell = \alpha \hat{\boldsymbol{h}}_{\ell-1} + \boldsymbol{K}_\ell (\boldsymbol{y}_\ell - \alpha \hat{\boldsymbol{h}}_{\ell-1} x_\ell). \tag{6}$$

It can be observed that $\hat{\boldsymbol{h}}_\ell$ is ZMCSCG. We denote the estimation error by $\tilde{\boldsymbol{h}}_\ell = \boldsymbol{h}_\ell - \hat{\boldsymbol{h}}_\ell$, and its covariance matrix by $\boldsymbol{M}_\ell = E\{\tilde{\boldsymbol{h}}_\ell \tilde{\boldsymbol{h}}_\ell^\dagger\}$. The update equation of $\boldsymbol{M}_\ell$ is given by

$$\boldsymbol{M}_\ell = (1 - \boldsymbol{K}_\ell x_\ell)(\alpha^2 \boldsymbol{M}_{\ell-1} + \boldsymbol{R}_{\boldsymbol{w}}). \tag{7}$$

Without loss of generality, we initialize $\hat{\boldsymbol{h}}_0 = 0$ and $\boldsymbol{M}_0 = \boldsymbol{R}_{\boldsymbol{h}}$. The orthogonality property of LMMSE estimator states that $\hat{\boldsymbol{h}}_\ell$ and $\tilde{\boldsymbol{h}}_\ell$ are uncorrelated, which implies the covariance of $\hat{\boldsymbol{h}}_\ell$ is given by $\boldsymbol{R}_{\hat{\boldsymbol{h}}_\ell} = \boldsymbol{R}_{\boldsymbol{h}} - \boldsymbol{M}_\ell$.

We are interested in the steady-state behaviour of the Kalman filter. At the periodic steady-state, we have $\boldsymbol{M}_\ell = \boldsymbol{M}_{\ell-T}$ $\forall \ell$. In order to find a closed-form expression for the steady-state error covariance matrix, we first focus on the pilot transmission mode and express $\boldsymbol{M}_\ell$ in terms of $\boldsymbol{M}_{\ell-T}$.

From (2), the channel states between two consecutive pilot transmissions are related as

$$\boldsymbol{h}_\ell = \alpha^T \boldsymbol{h}_{\ell-T} + \boldsymbol{v}_\ell, \qquad \text{where} \qquad \boldsymbol{v}_\ell = \sum_{k=0}^{T-1} \alpha^k \boldsymbol{w}_{\ell-k}.$$

Following the Kalman filter update equations in (5) and (7), the covariance of channel estimation error between two consecutive pilot transmissions can be written as

$$\begin{aligned} \boldsymbol{M}_\ell \;=\; & \left((\alpha^{2T} \boldsymbol{M}_{\ell-T} + (1-\alpha^{2T})\boldsymbol{R}_{\boldsymbol{h}}) \frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2} + \boldsymbol{I}_{N_r} \right)^{-1} \times \\ & \left(\alpha^{2T} \boldsymbol{M}_{\ell-T} + (1-\alpha^{2T})\boldsymbol{R}_{\boldsymbol{h}} \right). \end{aligned} \tag{8}$$

When the Kalman filter reaches the steady-state, we denote the steady-state error covariance matrix at pilot transmission by $\boldsymbol{M}_{ss,1} \doteq \boldsymbol{M}_\ell = \boldsymbol{M}_{\ell-T}$. Rearranging the above equation, we obtain the following quadratic matrix equation

$$\boldsymbol{M}_{ss,1}^2 + \frac{1-\alpha^{2T}}{\alpha^{2T}\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}} \left(\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2} \boldsymbol{R}_{\boldsymbol{h}} + \boldsymbol{I}_{N_r} \right) \boldsymbol{M}_{ss,1} - \frac{1-\alpha^{2T}}{\alpha^{2T}\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}} \boldsymbol{R}_{\boldsymbol{h}} = 0.$$

It can be shown using result in [13] that the above quadratic matrix equation satisfies the conditions for an explicit solution

in the same form as in the scalar case. Therefore, the solution is given as

$$\boldsymbol{M}_{ss,1} = -\frac{1}{2}\kappa\boldsymbol{Q} + \frac{1}{2}\left(\kappa^2\boldsymbol{Q}^2 + 4\kappa\boldsymbol{R}_{\boldsymbol{h}}\right)^{1/2}, \tag{9}$$

where $(\boldsymbol{A})^{1/2}$ denotes the matrix square root of $\boldsymbol{A}$, and

$$\boldsymbol{Q} = \frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}\boldsymbol{R}_{\boldsymbol{h}} + \boldsymbol{I}_{N_r}, \qquad \text{and} \qquad \kappa = \frac{1-\alpha^{2T}}{\alpha^{2T}\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}}.$$

During data transmission, the channel prediction is given by

$$\hat{\boldsymbol{h}}_\ell = \alpha\hat{\boldsymbol{h}}_{\ell-1}, \tag{10}$$

and the error covariance update equation is given by

$$\boldsymbol{M}_\ell = \alpha^2\boldsymbol{M}_{\ell-1} + \boldsymbol{R}_{\boldsymbol{w}}. \tag{11}$$

The steady-state error covariance matrix during data transmissions can be calculated iteratively using (11) as

$$\boldsymbol{M}_{ss,\ell} = \boldsymbol{R}_{\boldsymbol{h}} + \alpha^{2(\ell-1)}(\boldsymbol{M}_{ss,1} - \boldsymbol{R}_{\boldsymbol{h}}), \tag{12}$$

where $\ell = 2, 3, ..., T$.

IV. Channel Capacity

In this section, we study the channel capacity when the Kalman filter has reached the steady-state.

A. A Capacity Lower Bound

Without loss of generality, we normalize the variance of the channel gains, i.e. $\sigma_{\boldsymbol{h}}^2 = 1$. For systems with imperfect CSI at the receiver, the exact capacity expression is still unavailable. Alternatively, we consider a lower bound for the instantaneous capacity, which has been used for information-theoretic studies [14], given by

$$C_{\mathrm{LB},\ell} = E_{\hat{\boldsymbol{h}}_\ell}\left\{\log_2\left|\boldsymbol{I}_{N_r} + \rho_d(\rho_d\boldsymbol{M}_{ss,\ell} + \boldsymbol{I}_{N_r})^{-1}\hat{\boldsymbol{h}}_\ell\hat{\boldsymbol{h}}_\ell^\dagger\right|\right\}. \tag{13}$$

For the case where entries of $\boldsymbol{h}_\ell$ are i.i.d., $\boldsymbol{M}_{ss,\ell}$ is a diagonal matrix with the same diagonal entry denoted by $\sigma_{e,\ell}^2$, which is proven later in *Lemma 1*, and the entries of $\hat{\boldsymbol{h}}_\ell$ are i.i.d. with variance $1 - \sigma_{e,\ell}^2$. Let $\zeta_\ell = \hat{\boldsymbol{h}}_\ell^\dagger\hat{\boldsymbol{h}}_\ell$. Note that ζ_ℓ is a Gamma distributed random variable with parameters $(N_r, 1-\sigma_{e,\ell}^2)$. From (10), one can show that $\zeta_\ell = \alpha^{2(\ell-1)}\zeta_1$. Using (4) and the matrix determinant lemma $|\boldsymbol{I}+\boldsymbol{AB}| = |\boldsymbol{I}+\boldsymbol{BA}|$, the instantaneous capacity lower bound in (13) can be rewritten as

$$C_{\mathrm{LB},\ell} = E_{\zeta_1}\left\{\log_2\left(1 + \frac{\frac{1-\gamma}{1-\eta}\rho\alpha^{2(\ell-1)}\zeta_1}{\frac{1-\gamma}{1-\eta}\rho\sigma_{e,\ell}^2 + 1}\right)\right\}. \tag{14}$$

When entries of $\boldsymbol{h}_\ell$ are correlated, $\hat{\boldsymbol{h}}_\ell$ has correlated entries as well. In this case, Monte Carlo simulation will be used to carry out the numerical analysis in Section V.

For both spatially i.i.d. and correlated channels, the capacity lower bound per transmission block is given by

$$C_{\mathrm{LB}} = \frac{1}{T}\sum_{\ell=2}^{T} C_{\mathrm{LB},\ell}. \tag{15}$$

B. The Effect of Spatial Correlation on the Capacity

The authors in [11] showed that spatial correlation always reduces capacity, assuming perfect CSI at the receiver. We would like to ask whether it is still true when imperfect CSI is available at the receiver. To answer this question, we first look at the effect of channel spatial correlation on the channel estimation MMSE. Our finding is summarized in the following lemma with the proof given in Appendix A.

Lemma 1: Consider the system model given by (1) and (2) with the assumption that the channel covariance matrix $\boldsymbol{R}_{\boldsymbol{h}}$ is full rank. Under the Kalman filter setup for channel estimation, spatially i.i.d. channels result in the maximum channel estimation MMSE, and the covariance matrix of the estimation error is diagonal with the same entries on its main diagonal.

At very high operating SNR, the channel estimation error is negligible. Therefore, we expect the effect of channel spatial correlation on the capacity to be the same as in the perfect CSI case.

Here we focus on the effect of spatial correlation in sufficiently low data symbol SNR (ρ_d) regime. In this regime, the instantaneous capacity lower bound in (13) can be approximated as

$$\begin{aligned} C_{\mathrm{LB},\ell} &\approx E_{\hat{\boldsymbol{h}}_\ell}\left\{\log_2\left|\boldsymbol{I}_{N_r} + \rho_d\hat{\boldsymbol{h}}_\ell\hat{\boldsymbol{h}}_\ell^\dagger\right|\right\}, \\ &= \frac{1}{\ln 2}E_{\hat{\boldsymbol{h}}_\ell}\,\mathrm{tr}\left\{\ln\left(\boldsymbol{I}_{N_r} + \rho_d\hat{\boldsymbol{h}}_\ell\hat{\boldsymbol{h}}_\ell^\dagger\right)\right\}, \quad (16) \\ &\approx \frac{1}{\ln 2}E_{\hat{\boldsymbol{h}}_\ell}\,\mathrm{tr}\left\{\rho_d\hat{\boldsymbol{h}}_\ell\hat{\boldsymbol{h}}_\ell^\dagger\right\}, \quad (17) \\ &= \frac{\rho_d}{\ln 2}(N_r - \mathrm{tr}\{\boldsymbol{M}_{ss,\ell}\}), \quad (18) \end{aligned}$$

where (16) is obtained using $\ln|\cdot| = \mathrm{tr}\{\ln(\cdot)\}$, and (17) is obtained using Taylor's series expansion of $\ln(\cdot)$. It can be seen that (18) is a decreasing function of $\mathrm{tr}\{\boldsymbol{M}_{ss,\ell}\}$. From *Lemma 1*, we know that spatially i.i.d. channels result in maximum $\mathrm{tr}\{\boldsymbol{M}_{ss,\ell}\}$. *This implies that at sufficiently low SNR, spatial correlation between channels are desirable for higher capacity.* This finding signifies the effect of channel estimation error on the capacity in contrast to the perfect CSI assumption in [11].

V. Numerical Results

In this section, we perform numerical analysis on the optimal values of the PSAM design parameters, i.e. the pilot power ratio γ and the pilot spacing η (or equivalently the block length T), which maximize the capacity lower bound in (15).

A. Spatially i.i.d. Channels

Fig. 1 shows a 3D plot of the capacity lower bound in (15) for a 1×4 SIMO system. The average SNR budget is $\rho = 10$dB, and normalized fading rate is $f_D T_s = 0.11$ which gives the Gauss-Markov parameter $\alpha = 0.884$ in (2), i.e. moderately fast time-varying channels. From Fig. 1, the optimal parameter values are $\gamma_{\mathrm{opt}} = 0.37$ and $\eta_{\mathrm{opt}} = 0.25$ (or $T_{\mathrm{opt}} = 4$). We observe that C_{LB} is more sensitive to the pilot spacing η (or the block length T) than to the pilot power ratio γ. Particularly,

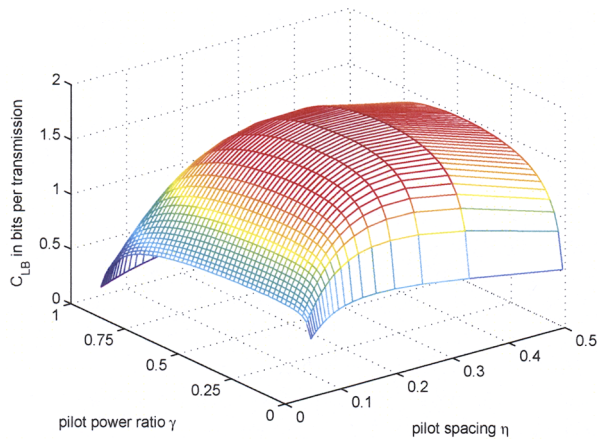

Fig. 1. The capacity lower bound in (15) at SNR budget $\rho = 10$dB, and normalized fading rate $f_D T_s = 0.11$ (i.e. $\alpha = 0.884$), for a 1×4 system.

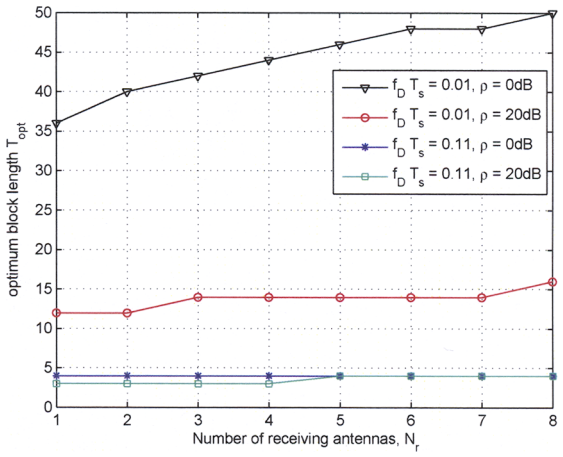

Fig. 3. The optimum block length for different numbers of receive antennas, at both high and low SNR budgets and normalized fading rates.

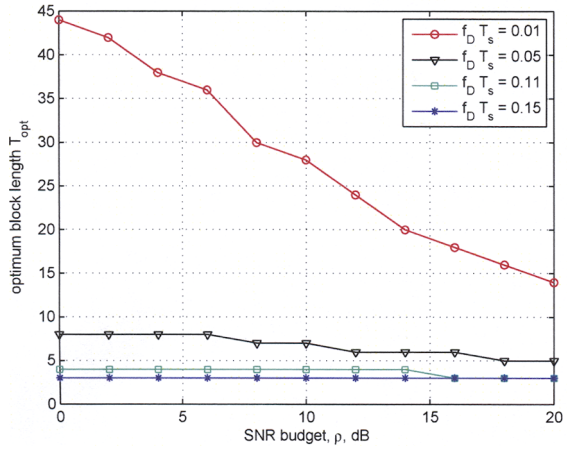

Fig. 2. The optimum block length T_{opt} for a wide range of SNR budget and normalized fading rates, for a 1×4 system. T_{opt} is found by a numerical search in (15) where different T and γ values are checked.

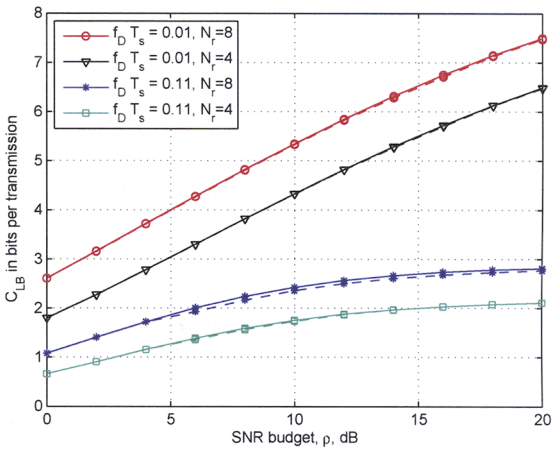

Fig. 4. The capacity lower bound in (15) for a wide range of SNR budget for different normalized fading rates and numbers of receive antennas. The solid lines are capacity lower bound using optimum parameters for SIMO systems, and the dashed lines are the capacity lower bound achieved using optimum parameters SISO systems for the same antenna size and fading rate.

C_{LB} is almost constant for γ ranging from 0.2 to 0.55, with less than 5% degradation from the optimal value. We also studied the plots for $N_r = 1, 2, 8$ (not shown in this paper), and the trends are very similar to the one shown in Fig. 1. Our observations indicate that the optimal design of the pilot spacing is more important than that of the pilot power ratio. Hence, we focus on the optimal pilot spacing in the following capacity analysis.

Fig. 2 shows the optimal block length $T_{\text{opt}} = 1/\eta_{\text{opt}}$ for a wide range of SNR and normalized fading rates, for a 1×4 SIMO system. We also produced the plots for $N_r = 1, 2, 8$ (not shown in this paper), and the same trends are found in all cases. Firstly, the optimum block length decreases as the fading rate increases for a fixed SNR. This is expected as more frequent training is needed when channel varies faster. Secondly, for slow fading channels, e.g. $f_D T_s = 0.01$, the optimum block length decreases dramatically as the SNR increases, while for fast fading channels, e.g. $f_D T_s = 0.11, 0.15$, the optimum block length is almost constant as SNR increases.

Fig. 3 shows the optimum block length for different numbers of receive antennas. We see that the optimal block length generally increases with N_r. Since more antennas produce higher diversity, which improves the tolerance on the channel estimation error, the system can allow a larger pilot spacing. Furthermore, the increases in the optimum block length is more sensitive to N_r in the slow fading channel ($f_D T_s = 0.01$) than in the faster fading channel ($f_D T_s = 0.11$).

Fig. 4 shows the capacity lower bound for a wide range of SNR budget for different normalized fading rates and number of receive antennas. We plot both the capacity lower bound achieved using optimum parameters for each SIMO case (solid line) and that achieved using optimal parameter for the SISO case (dashed line). We see that the difference between the two is negligible, except for the case where $f_D T_s = 0.11$ and $N_r = 8$, in which the maximum difference is approximately 3% at $\rho = 8$dB. *Therefore, by optimizing parameters for SISO*

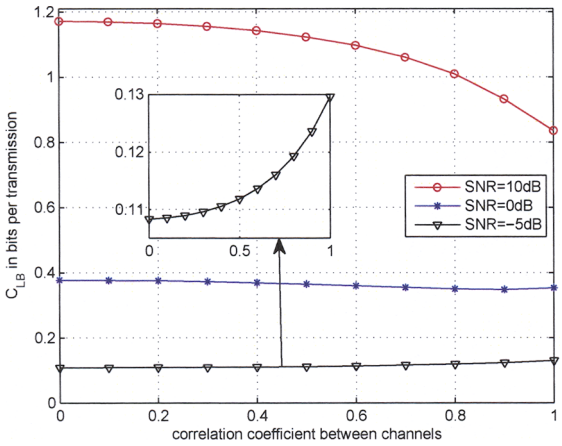

Fig. 5. The capacity lower bound in (15) using optimum parameters vs. spatial correlation coefficient for normalized fading rate $f_D T_s = 0.11$ and different SNR, for a 1×2 SIMO system.

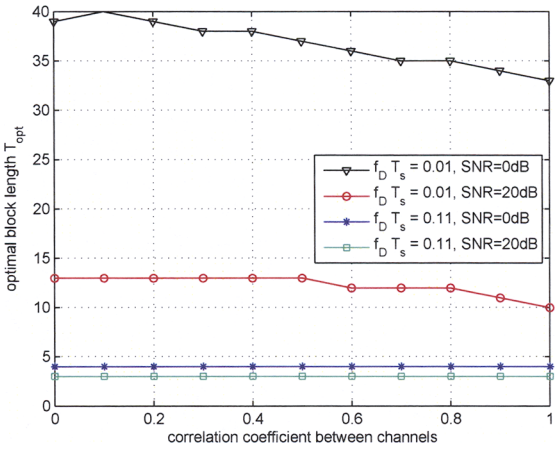

Fig. 6. The optimum block length vs. spatial correlation coefficient for a 1×2 system, at both high and low SNR budgets and normalized fading rates.

channel, one can achieve near optimum capacity for spatially i.i.d. SIMO channels of practical antenna sizes ($N_r \leq 8$).

B. Spatially Correlated Channels

We study the effect of channel spatial correlation on the optimal design parameters. Without loss of generality, we place the receive antennas on a uniform circular array and use the standard Jake's model to calculate the spatial correlation under isotropic scattering environment [15]. From the analysis in Section IV, we expect that spatial correlation may increase the capacity through reduction in channel estimation error at sufficiently low SNR, while we argued that correlation reduces the capacity at high SNR.

Fig. 5 shows the capacity lower bound in (15) using optimum parameters for a 1×2 SIMO system varying from spatially i.i.d. channels to identical channels (fully correlated). At moderately high SNR, e.g. $\rho = 10$dB, spatially i.i.d. channels result in the maximum capacity lower bound, and the capacity lower bound decreases as the two channels become more correlated. However, a different trend is found at low SNR. At $\rho = 0$dB, spatially i.i.d. channel still results in the maximum capacity lower bound, but the minimum occurs at correlation coefficient of 0.9 and not at identical channels. Furthermore, at $\rho = -5$dB, the capacity lower bound has a 20% increase from spatially i.i.d. channels to identical channels. These observations confirm our earlier analysis in Section IV. In addition, these numerical results confirm that spatially i.i.d. channels give the maximum information capacity at practical operating SNRs.

Fig. 6 shows the optimum block length verses the spatial correlation coefficient between channels for a 1×2 SIMO system. We also plotted for optimum pilot power ratio (not shown in this paper) and observed the same trend. In general, we see that the optimum parameters remain roughly constant when channel correlation coefficient is less than 0.5, and have some gradual changes for slow fading channels when correlation coefficient increases above 0.5. Therefore, we can say

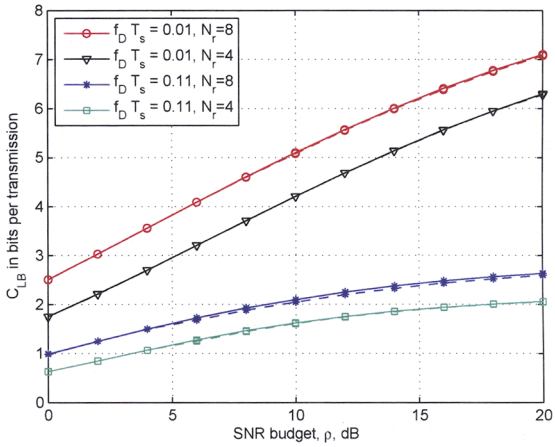

Fig. 7. The capacity lower bound in (15) for a wide range of SNR, and different fading rates and numbers of antennas. The solid lines are the capacity lower bounds using optimum parameters for the correlated SIMO systems, and the dashed lines are the capacity lower bounds using optimum parameters for SISO systems.

that the parameters which optimize the capacity for spatially i.i.d. channels may also achieve near optimum capacity for correlated channels.

Fig. 7 shows the capacity lower bound in (15) for correlated channels, fixing the inter-antenna distance to be a quarter of a wavelength. We plot both the capacity lower bound achieved using the optimum parameters for each correlated SIMO case (solid line) and that achieved using the optimal parameter for the SISO case (dashed line). We see that the difference between the two is negligible, except for the case where $f_D T_s = 0.11$ and $N_r = 8$, in which the maximum difference is approximately 3% at $\rho = 6$dB. *Together with the result obtained for spatially i.i.d. channels, we conclude that by optimizing parameters for SISO channel, the same parameters achieve near optimum capacity for both spatially i.i.d. and correlated SIMO system of practical antenna size ($N_r \leq 8$).*

VI. Conclusion

We studied the design parameters of PSAM schemes from an information-theoretic viewpoint in time-varying flat-fading SIMO channels with and without spatial correlation. The design parameters to maximize the channel capacity were the ratio of power allocated to pilot symbols and the fraction of time allocated to pilot symbols. We studied a capacity lower bound and showed that it is more sensitive to changes in the pilot spacing than to the pilot power ratio. The optimum pilot power ratio and pilot spacing remain relatively constant as the channel spatial correlation increases. We showed that at sufficiently low SNR, spatial correlation results in an increase in the capacity compared with spatially i.i.d. channels. Most importantly, our findings showed that by optimally designing the training parameters for SISO systems, the same parameters can be used to achieve near optimum capacity in both spatially i.i.d. and correlated SIMO systems for the same SNR and fading rate. We are currently extending this work to MIMO systems.

Appendix A
Proof of Lemma 1

We use the mathematical induction approach by firstly looking at the initial channel estimation, i.e. $\boldsymbol{M}_1$. From (5), (7) and the given initializations of the Kalman filter, one can show that

$$\boldsymbol{M}_1=\boldsymbol{R_h}-\boldsymbol{R_h}\mathcal{E}_p(\boldsymbol{R_h}\mathcal{E}_p+\sigma_{\boldsymbol{n}}^2\boldsymbol{I}_{N_r})^{-1}\boldsymbol{R_h}=(\boldsymbol{R_h}^{-1}+\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}\boldsymbol{I}_{N_r})^{-1},$$

where the second equality is obtained using the matrix inversion lemma. Therefore,

$$\mathrm{tr}\{\boldsymbol{M}_1\}=\sum_{i=1}^{N_r}(g_i^{-1}+\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1},$$

where g_i, $i=1,...,N_r$ are the eigenvalues of $\boldsymbol{R_h}$. The values of g_i at which the maxima of $\mathrm{tr}\{\boldsymbol{M}_1\}$ occurs under the constraint $\sum_{i=1}^{N_r} g_i = \mathrm{tr}\{\boldsymbol{R_h}\} = \sigma_{\boldsymbol{h}}^2 N_r$ form a Lagrange multiplier problem, and the solution is given by $g_{i,max}=\sigma_{\boldsymbol{h}}^2\ \forall i$. This implies that the channels are i.i.d with $\boldsymbol{R_h}=\sigma_{\boldsymbol{h}}^2\boldsymbol{I}_{N_r}$. In this case, $\boldsymbol{M}_1$ is a diagonal matrix with diagonal entries taking the same values. Now we have proven the claim in *Lemma 1* for $\ell=1$. The next step is to prove this holds for all ℓ, given it holds for $\ell-1$. The proof for all data transmission time slots is trivial. Here, we only show the proof for all pilot transmission time slots, i.e. $\ell=1,T+1,2T+1,....$

Consider pilot transmission based on (8), we let $\bar{\boldsymbol{M}}_\ell = \alpha^{2T}\boldsymbol{M}_{\ell-T}+(1-\alpha^{2T})\boldsymbol{R_h}$. We assume the claim in *Lemma 1* is true for $\ell-T$ or effectively for $\boldsymbol{M}_{\ell-T}$. Then, clearly $\bar{\boldsymbol{M}}_\ell$ is diagonal and $\sum_{i=1}^{N_r}\bar{m}_i$ achieves its maximum with the same value for $\bar{m}_i$, where $\bar{m}_i$, $i=1,...,N_r$ are the eigenvalues of $\bar{\boldsymbol{M}}_\ell$. To complete the proof, we only need to show that the properties of $\bar{m}_i$ imply the claim in *Lemma 1* is also true for $\boldsymbol{M}_\ell$.

Using the matrix inversion lemma, (8) can be rewritten as

$$\boldsymbol{M}_\ell = (\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1}\boldsymbol{I}_{N_r}-(\bar{\boldsymbol{M}}_\ell\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}+\boldsymbol{I}_{N_r})^{-1}(\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1}.$$

Therefore,

$$\mathrm{tr}\{\boldsymbol{M}_\ell\}=N_r(\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1}-(\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1}\sum_{i=1}^{N_r}(\bar{m}_i\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}+1)^{-1}.$$

Let us impose an arbitrary constraint $\sum_{i=1}^{N_r}\bar{m}_i=\psi>0$. The values of $\bar{m}_i$ at which the maxima of $\mathrm{tr}\{\boldsymbol{M}_\ell\}$ occurs under this constraint form a Lagrange multiplier problem, and the solution is given by $\bar{m}_{i,max}=\frac{\psi}{N_r}\ \forall i$. The constrained maximum value of $\mathrm{tr}\{\boldsymbol{M}_\ell\}$ is given by

$$\mathrm{tr}\{\boldsymbol{M}_\ell\}=N_r(\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1}-(\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2})^{-1}N_r(\frac{\psi}{N_r}\frac{\mathcal{E}_p}{\sigma_{\boldsymbol{n}}^2}+1)^{-1}.$$

It is clear that $\mathrm{tr}\{\boldsymbol{M}_\ell\}$ is maximized when ψ takes its maximum value. Together with the i.i.d. channels assumption, it is easy to show that $\boldsymbol{M}_\ell$ is diagonal with the same diagonal entries. And we have completed the proof.

Acknowledgements

This work was supported by the Australian Research Council Discovery Grant DP0773898.

References

[1] S. M. Redl, M. K. Weber, and M. W. Oliphan, *An introduction to GSM*, 1st ed. Boston: Artech House, 1995.

[2] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[3] S. Ohno and G. B. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 712–720, Oct. 2002.

[4] G. F. J. Baltersee and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.

[5] M. Dong, L. Tong, and B. M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1403–1418, May 2004.

[6] A. Bdeir, I. Abou-Faycal, and M. Médard, "Power allocation schemes for pilot symbol assisted modulation over Rayleigh fading channels with no feedback," in *Proc. IEEE ICC*, vol. 2, June 2004, pp. 737–741.

[7] P. Sadeghi, Y. Liu, R. A. Kennedy, and P. B. Rapajic, "Pilot symbol transmission for time-varying fading channels: An information-theoretic optimization," in *Proc. ICSPCS*, Dec. 2007.

[8] B. Hassibi and M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[9] V. Pohl, P. H. Nguyen, V. Jungnickel, and C. Helmolt, "Continuous flat-fading MIMO channels: Achievable rate and optimal length of the training and data phases," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1889–1900, July 2005.

[10] E. A. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Trans. Signal Processing*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[11] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[12] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Dover Publications, Inc., 1979.

[13] N. J. Higham and H. Kim, "Solving a quadratic matrix equation by newton's method with exact line searches," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 2, pp. 303–316, 2001.

[14] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[15] T. A. Lamahewa, R. A. Kennedy, T. D. Abhayapala, and T. Betlehem, "MIMO channel correlation in general scattering environments," in *Proc. AusCTW*, Feb. 2006, pp. 93–98.