Power Control for Multi-antenna Gaussian Channels with Delayed Feedback

Devdutt V. Marathe and Srikrishna Bhashyam
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai 600036, India
Phone: 91-44-22574439, Fax: 91-44-22570120
E-mail: skrishna@ee.iitm.ac.in

***Abstract*— Channel feedback is delayed in all practical wireless systems. Adaptive transmission strategies need to account for this feedback delay in order to be effective. We consider multiple-input single-output systems that perform power control based on a delayed SNR estimate. For perfect channel state information at the receiver (CSIR) and delayed channel state information at the transmitter (CSIT), we derive an expression for the power control function that minimizes outage probability at a constant rate. For imperfect CSIR and delayed CSIT, we derive an upper-bound on outage probability, and an expression for the power control function that minimizes this upper-bound.**

I. Introduction

The capacity and outage performance of multi-antenna Gaussian channels with perfect channel state information at the receiver (CSIR) and no channel state information at the transmitter (CSIT) have been derived in [1]. Perfect CSIT can significantly improve the outage performance [2]. However, in most systems, it is impractical to assume that the transmitter possesses error-free knowledge of the current channel. Multiple-antenna systems with various types of partial CSIT have been studied in [3], [4], [5]. The capacity of multiple-antenna systems with mean feedback or covariance feedback is discussed in [3]. Outage performance bounds for multiple-input single-output (MISO) systems with power control are analyzed for quantized SNR feedback in [4], while outage performance bounds with beam-forming are analyzed in [5] for quantized feedback of the channel realizations.

In this paper, we analyze the performance of multiple-input single-output (MISO) systems with power control under the assumption that the transmitter knows the channel SNR estimate after a feedback delay. This is a realistic assumption since the receiver usually feeds back some information (for example, SNR) once every frame. Using the Gaussian nature of the channel model, the effect of the feedback delay is captured using a correlation coefficient ρ and the performance is analyzed as a function of ρ. We choose to concentrate our attention on power control since it is more beneficial to use feedback for power control than for beamforming [4]. While beam-forming gives only a fixed gain, power control improves the decay rate of the outage probability as a function of average SNR. Moreover, power control requires only SNR feedback, while beam-forming requires feedback of individual channel coefficients.

First, we assume perfect CSIR and delayed CSIT, i.e., the transmitter knows the exact SNR at some earlier time, and find the power control function that minimizes outage probability. Using the theory of "Calculus of Variations" developed by Euler and Lagrange, we derive an implicit closed-form expression for the optimal power control function. Next, we relax the perfect CSIR assumption as well, and derive a lower-bound to the mutual information. Using this lower-bound, we upper-bound the outage probability and derive the optimal power control function. Finally, we consider maximizing the average mutual information instead of minimizing outage probability using power control.

II. Preliminaries

A. Modelling feedback delay and estimation error

We assume a flat-fading Gaussian channel. The MISO system with M transmit antennas and one receive antenna is, as usual [1], [4], modelled by the equation:

$$y = \sqrt{\frac{P}{M}} p(\gamma) \mathbf{h}^T \mathbf{x} + \eta, \tag{1}$$

$\mathbf{h}_{M\times 1}$, distributed as $\mathcal{CN}(\mathbf{0}, I_{M\times M})$, is the vector of the *normalized* channel gains, η (AWGN) is $\mathcal{CN}(0,1)$, P is the average SNR, $p(\gamma)$ is the power control function that relies on the transmitter's estimate of the normalized SNR (γ), and $\mathbf{E}_\Gamma(p(\gamma)) \leq 1$ is the average power constraint. We note here that the variance of each element in $\mathbf{h}$ and of η is 1, regardless of the SNR or the number of antennas. The term $\sqrt{\frac{P}{M}}$ takes care of the scaling due to SNR and multiple antennas.

Using the fact that the channel gains form a Gaussian process, we can model the relationship between the old channel and the current channel as

$$\mathbf{h} = \rho \mathbf{h}_{\text{old}} + \sqrt{1-\rho^2}\ \mathbf{v}, \tag{2}$$

where $\mathbf{h}_{\text{old}}$ and $\mathbf{v}$ are independent $\mathcal{CN}(\mathbf{0}, I_{M\times M})$. For example, in Jakes' model, $\rho = J_0(\omega_d \Delta t)$, where ω_d is the Doppler frequency and Δt is the feedback delay.

To model estimation error, we assume that the receiver uses the MMSE estimate of the channel gains, as in [4]. The MMSE estimate is derived using a preamble:

$$\hat{\mathbf{h}} = \frac{\sqrt{\frac{P_t}{M}}}{\frac{P_t}{M}+1} \left(\sqrt{\frac{P_t}{M}} \mathbf{h} + \mathbf{z} \right), \tag{3}$$

where P_t is the total power used for training and $\mathbf{z}$ is $\mathcal{CN}(\mathbf{0}, I_{M\times M})$.

The estimation error and feedback delay can be modelled together to relate the current channel $\mathbf{h}$ and the estimate of the old channel $\hat{\mathbf{h}}_{\text{old}}$ by the following equation:

$$\mathbf{h} = \rho\hat{\mathbf{h}}_{\text{old}} + \sqrt{1-\rho^2\frac{\frac{P_t}{M}}{\frac{P_t}{M}+1}}\,\mathbf{w}, \tag{4}$$

where $\mathbf{w}$ is $\mathcal{CN}(\mathbf{0}, I_{M\times M})$ and independent of $\hat{\mathbf{h}}_{\text{old}}$.

We can also model the relationship between the old estimate $\hat{\mathbf{h}}_{\text{old}}$ and the current estimate $\hat{\mathbf{h}}_{\text{new}}$ by the relationship:

$$\hat{\mathbf{h}}_{\text{new}} = \rho\hat{\mathbf{h}}_{\text{old}} + \sqrt{1-\rho^2}\sqrt{\frac{\frac{P_t}{M}}{\frac{P_t}{M}+1}}\mathbf{n}, \tag{5}$$

where $\hat{\mathbf{h}}_{\text{old}}$ is the MMSE estimate of the channel at some earlier time. We note here that $\hat{\mathbf{h}}_{\text{old}}$ and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, I_{M\times M})$ are independent.

B. Calculus of Variations

The theory of Calculus of Variations, developed by Euler and Lagrange, is a powerful tool used to solve problems that involve finding the functions that maximize or minimize a particular integral. Since there are many excellent books, for example [7], that deal with the theory of Calculus of Variations, we will only summarize the results that will be used later on in this paper.

Suppose we want to find the function $f(x)$ that will maximize or minimize the integral

$$\int_a^b F(f, \dot{f}, x)dx$$

with respect to the equality constraints

$$\phi_i(f, \dot{f}, x) = 0.$$

We define $\hat{F}(f, \dot{f}, x)$ to be

$$\hat{F}(f, \dot{f}, x) = F(f, \dot{f}, x) + \sum_{i=1}^{N}\lambda_i(x)\phi_i(f, \dot{f}, x), \tag{6}$$

and solve the *Euler-Lagrange equation*

$$\frac{\partial\hat{F}}{\partial f} - \frac{d}{dx}\frac{\partial\hat{F}}{\partial\dot{f}} = 0. \tag{7}$$

III. Power Control to minimize outage probability

Outage is defined as the event where the channel SNR is insufficient to allow transmission at a given rate. We will assume here that the transmitter knows the SNR corresponding to the receiver's estimate of the channel at some earlier time. For a particular value of the cross-correlation ρ, we seek to find the power control function that minimizes the outage probability. We consider two cases:

- Perfect CSIR and Delayed (but estimation-noise-free) CSIT
- Imperfect CSIR and Delayed and Noisy CSIT

A. Perfect CSIR and Delayed CSIT

Assuming perfect CSIR, the mutual information $I(X;Y|\mathbf{h}, \mathbf{h}_{\text{old}})$ can be written as

$$I(X;Y|\mathbf{h}, \mathbf{h}_{\text{old}}) = \log\left(1+\frac{P}{M}p(\gamma)\xi\right) \text{ nats/s/Hz}, \tag{8}$$

where $\xi = \|\mathbf{h}\|^2$ and $\gamma = \|\mathbf{h}_{\text{old}}\|^2$.

To pose the problem of finding the optimal power control function as a Calculus of Variations problem, we express the probability of outage as:

$$\begin{aligned}Pr(outage) &= Pr(I(X;Y) < R)\\ &= \int_0^\infty \underbrace{f_\Gamma(\gamma)Pr(outage|\gamma)}_{F(p,\dot{p},\gamma)}d\gamma.\end{aligned} \tag{9}$$

The average power constraint can be written as an equality constraint:

$$\underbrace{\mathbf{E}_\Gamma(p(\gamma)) - 1}_{\phi(p,\dot{p},\gamma)} = 0. \tag{10}$$

From (9) and (10), we can write $\hat{F}(p, \dot{p}, \gamma)$ (6) as:

$$\hat{F}(p, \dot{p}, \gamma) = f_\Gamma(\gamma)Pr(outage|\gamma) + \lambda(\gamma)(\mathbf{E}_\Gamma(p(\gamma)) - 1). \tag{11}$$

After substituting the various density functions, we solve the Euler-Lagrange equation (7) to get

$$\begin{aligned}\lambda(\gamma) =& f_\Gamma(\gamma)e^{-\mu\gamma}\frac{\beta}{p^2(\gamma)}\left(\frac{\beta}{\mu\gamma p(\gamma)}\right)^{\frac{M-1}{2}}\\ &\times e^{-\frac{\beta}{p(\gamma)}}I_{M-1}\left(2\sqrt{\mu\beta\frac{\gamma}{p(\gamma)}}\right),\end{aligned} \tag{12}$$

where $\mu = \frac{\rho^2}{1-\rho^2}$ and $\beta = \frac{e^R-1}{\frac{P}{M}}(\mu+1)$. $I_r(x)$ is the modified Bessel function of the first kind with order r.

Now we numerically find the $\lambda(\gamma)$ such that the resultant $p(\gamma)$ satisfies the power constraint and the non-negativity constraint. We have observed that the $\lambda(\gamma)$ that corresponds to the optimal $p(\gamma)$ is a scaled version of the density function $f_\Gamma(\gamma)$. With this observation we can cancel out $f_\Gamma(\gamma)$ on both sides of (12). This leads to faster convergence, especially in the regions where $f_\Gamma(\gamma)$ is small.

B. Extension to Imperfect CSIR and delayed CSIT

Using a method similar to that used in [4], we derive a lowerbound to the average mutual information. The final expression is given below. The details of this derivation are given in the Appendix.

$$\begin{aligned}I(X;Y|\hat{\mathbf{h}}_{\text{new}}) \geq& \frac{T-M}{T}\log\left(1+\frac{P_d p(\hat{\gamma})\hat{\xi}}{M\left(1+\sigma^2 P_d p(\hat{\gamma})\right)}\right)\\ =& \frac{T-M}{T}I_{\text{lb}}(X;Y|\hat{\mathbf{h}}_{\text{new}})\end{aligned} \tag{13}$$

where $\hat{\gamma} = \|\hat{\mathbf{h}}_{\text{old}}\|^2$, $\hat{\xi} = \|\hat{\mathbf{h}}_{\text{new}}\|^2$, $\sigma^2 = \frac{1}{\frac{P_t}{M}+1}$, P_t is the total training power, P_d is the average data power, and T is the frame length.

This bound can be further tightened by maximizing over (P_t,P_d), such that the total power constraint ($P_t + (T - M)P_d = PT$) is satisfied. Using the same approach as in [4], we find, for $T \gg M$, the optimal power splitting to be (P_t,P_d) = (PM,P).

Using this bound and results from the theory of Calculus of Variations, we derive expressions for the power control functions that (a) minimize an upperbound on outage probability, and (b) maximize a lowerbound on average mutual information.

The probability of outage can be upperbounded by:

$$Pr(outage) \leq Pr(I_{\text{lb}}(X;Y|\hat{\mathbf{h}}_{\text{new}}) < R') \tag{14}$$

where $R' = R\frac{T}{T-M}$.

As in the perfect CSIR case, we can derive the following expression for the Lagrangian $\lambda(\gamma)$:

$$\begin{aligned}\lambda(\gamma) =& f_{\hat{\Gamma}}(\hat{\gamma}) \exp\left(-\frac{\alpha(1+\sigma^2 P_d p(\hat{\gamma}))}{p(\hat{\gamma})} - \frac{\mu\hat{\gamma}}{\nu}\right)\frac{\alpha}{p^2(\hat{\gamma})} \\ &\times \left(\frac{\alpha(1+\sigma^2 P_d p(\hat{\gamma}))\nu}{p(\hat{\gamma})\mu\hat{\gamma}}\right)^{\frac{M-1}{2}} \\ &\times I_{M-1}\left(2\sqrt{\frac{\mu\hat{\gamma}\alpha(1+\sigma^2 P_d p(\hat{\gamma}))}{\nu p(\hat{\gamma})}}\right),\end{aligned} \tag{15}$$

where $\nu = \frac{\frac{P_t}{M}}{\frac{P_t}{M}+1} = 1 - \sigma^2$, $\alpha = \frac{(e^R-1)(\mu+1)}{\frac{P}{M}\nu}$. We observe that (15) is similar in form to (12).

C. Results and Observations

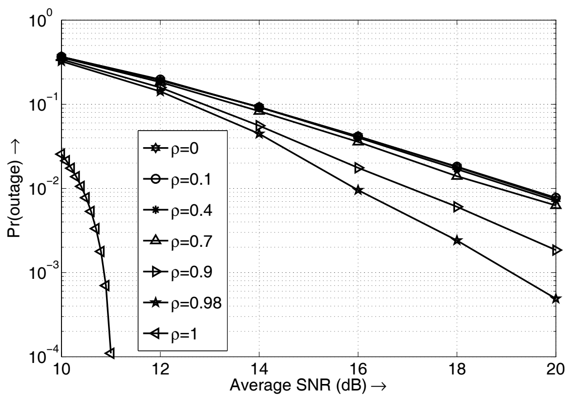

Fig. 1. Outage probability curves for various values of ρ for M=2 Tx antennas and R=2 nats/s/Hz with perfect CSIR

Fig. 1 shows the relationship between outage probability and average SNR for various values of ρ for a system with $M = 2$ transmit antennas and $R = 2$ nats/s/Hz. Fig. 2 shows the optimal power control functions for the same system at an average of SNR $P = 20$ dB with perfect CSIR. Fig. 3 shows the relationship between outage probability and rate at a fixed SNR of 20dB for M=2 antennas and perfect CSIR.

Fig. 4 compares the outage probability in the perfect CSIR case with the upperbound in the imperfect CSIR case (with

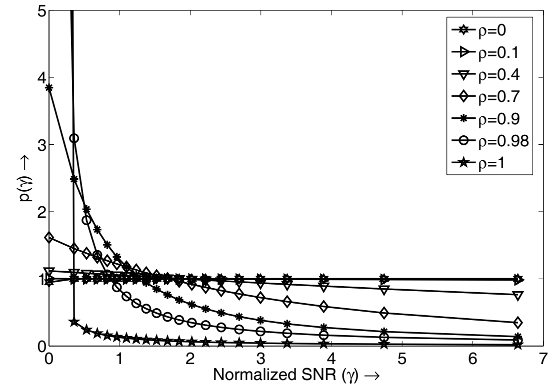

Fig. 2. Optimal power control functions for various values of ρ at average SNR of 20 dB, M=2 Tx Antennas and R=2 nats/s/Hz with perfect CSIR

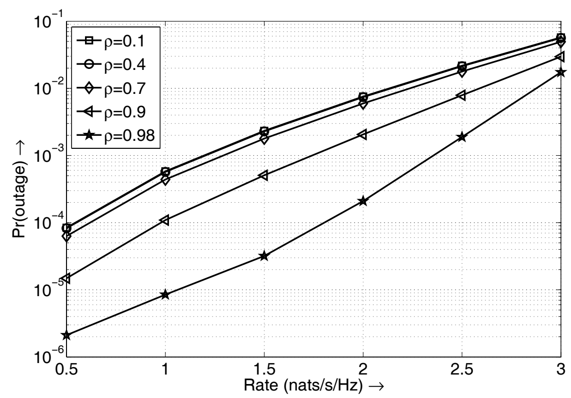

Fig. 3. Outage Probability vs Rate for M=2 antennas at average SNR of 20 dB

optimal power-splitting) for $M = 2$ transmit antennas, $R = 2$ nats/s/Hzm and $\rho = 0.9$.

We can make the following observations:

- Outage probability falls rapidly as ρ approaches 1, but is nearly constant as ρ increases from 0 to, say, 0.7. We conjecture that this is because the outage probability depends on $\mu = \frac{\rho^2}{1-\rho^2}$, which increases rapidly near $\rho = 1$.
- From Fig. 2, we can see that the optimal power control functions vary from no power control ($p(\gamma) = 1$) (which is optimal at $\rho = 0$) to the waterfilling curve (which is optimal for $\rho = 1$)
- From Fig. 3, we see that considerable gains can be achieved by better feedback. For example, at $Pr(outage) = 10^{-4}$, we can transmit at a rate of 1.8 nats/s/Hz with $\rho = 0.98$, instead of 0.5 nats/s/Hz with $\rho = 0.1$. The gain is lower for larger $Pr(outage)$.
- From Fig. 4, we see that the upperbound on $Pr(outage)$ with imperfect CSIR and optimal power-splitting is quite close to the value of $Pr(outage)$ with perfect CSIR. Moreover, the difference between the two curves reduces

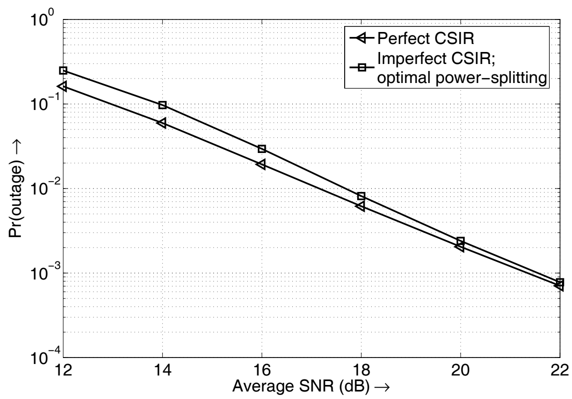

Fig. 4. Comparision between Pr(outage) for perfect and imperfect CSIR for $M = 2$ antennas at $R = 2$ nats/s/Hz and $\rho = 0.9$

with an increase in average SNR. It is reasonable to assume that these curves will be indistinguishable at asymptotically high SNRs.

In addition, we have observed that it is necessary that the power control function be matched to the exact value of ρ, especially as ρ nears 1. For example, using the optimal power control function for $\rho = 1$ when ρ is actually 0.98 results in an outage probability of nearly 0.5 in the above case ($M = 2, R = 2, SNR = 20dB$). However an interesting question to ask is what would happen if we consistently underestimated ρ. For example, if we assumed that our estimate and the actual SNR were statistically independent ($\rho = 0$) regardless of the actual value of ρ, we would achieve an outage probability of $Pr(outage) = 8 \times 10^{-3}$, which is the same as that for $\rho = 0$.

IV. Power Control to Maximize Average Mutual Information

The method of solving for the optimal power control functions that maximize the average mutual information for both perfect and imperfect CSIR is very similar to the approach used in the previous section. Therefore, we leave out the details here, and instead give only the results.

A. Perfect CSIR

We use the Calculus of Variations approach to obtain

$$\lambda(\gamma) = \mathbf{E}_{\Gamma,\Xi}\left(\frac{\frac{P}{M}\xi}{1 + \frac{P}{M}p(\gamma)\xi}\right), \qquad (16)$$

where $\gamma = ||\mathbf{h}_{\text{old}}||^2$ and $\xi = ||\mathbf{h}_{\text{new}}||^2$.

B. Imperfect CSIR

As before, we use the Calculus of Variations approach to obtain

$$\lambda(\gamma) = \mathbf{E}_{\Gamma,\Xi}\left(\frac{\frac{P_d}{M}\hat{\xi}}{(1 + \sigma^2 P_d p(\hat{\gamma}))\left(1 + P_d p(\hat{\gamma})(\sigma^2 + \frac{\hat{\xi}}{M})\right)}\right), \qquad (17)$$

where $\hat{\gamma} = ||\hat{\mathbf{h}}_{\text{old}}||^2$ and $\hat{\xi} = ||\hat{\mathbf{h}}_{\text{new}}||^2$.

C. Results and Observations

Fig. 5 shows the relationship between the average mutual information, maximized over all possible power control functions, and the average SNR, with perfect CSIR.

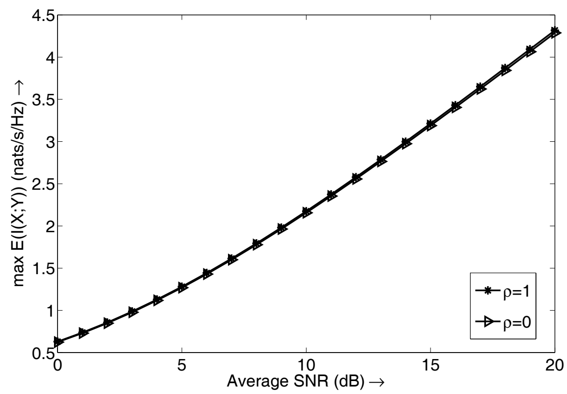

Fig. 5. Maximum average mutual information for $M = 2$ transmit antennas for $\rho = 0, 1$ with perfect CSIR

- We have observed that the maximum average mutual information is nearly independent of ρ. This was earlier observed for M=1 by Varaiya and Goldsmith in [8].
- We have found, via numerical simulations, that the optimal power control functions in this case nearly resemble the case of no power control ($p(\gamma) = 1 \ \forall \ \gamma$).
- We obtain similar results for the imperfect CSIR case, but these are of limited use since average mutual information is nearly independent of the correlation ρ.

V. Conclusions

In the preceding sections, we have derived implicit, closed-form equations for the optimal power control functions with respect to two metrics with the average SNR and the cross-correlation ρ as parameters. The actual values of the power control functions can be found by numerical methods. We have observed that power control is of limited use when the correlation between the actual SNR and the transmitter's estimate of SNR is not sufficiently high (of the order of 0.9, say). Another caveat is that the performance of the optimal power control functions depends heavily on the accuracy of the estimates of the parameters, especially the correlation ρ. We have also observed that power control does not significantly improve the average mutual information, and hence of not much use when performing rate control.

Appendix
Proof of Equation (13)

To prove (13), we first rewrite the system equation as:

$$y = \sqrt{\frac{P}{M}}p(\hat{\gamma})\hat{\mathbf{h}}_{\text{new}}^T\mathbf{x} + \hat{n}, \qquad (18)$$

where $\hat{n} = \sqrt{\frac{P}{M}} p(\hat{\gamma})(\mathbf{h} - \hat{\mathbf{h}}_{\text{new}})^T \mathbf{x} + n$. We observe that $\hat{n}$ possesses the following properties [4]:

$$\begin{aligned} \mathbf{E}(\hat{n}\mathbf{x}^\dagger | \hat{\mathbf{h}}_{\text{new}}) &= 0 \quad (19) \\ \mathbf{E}(|\hat{n}|^2 | \hat{\mathbf{h}}_{\text{new}}) &= 1 + \sigma^2 P_d \quad (20) \end{aligned}$$

To begin the derivation of (13), we write

$$\begin{aligned} I(X;Y|\hat{\mathbf{h}}_{\text{new}}) =& H(X|\hat{\mathbf{h}}_{\text{new}}) - H(X|Y,\hat{\mathbf{h}}_{\text{new}}) \\ =& H(X|\hat{\mathbf{h}}_{\text{new}}) - H(X - \mathbf{g}Y|Y,\hat{\mathbf{h}}_{\text{new}}) \\ \geq& H(X|\hat{\mathbf{h}}_{\text{new}}) - H(X - \mathbf{g}Y|\hat{\mathbf{h}}_{\text{new}}) \end{aligned} \quad (21)$$

The first term in the RHS above, $H(X|\hat{\mathbf{h}}_{\text{new}})$, can be upper-bounded using the Gaussian upperbound:

$$H(X|\hat{\mathbf{h}}_{\text{new}}) \leq \log(\det(\pi e Q)). \quad (22)$$

This bound is achieved if $\mathbf{x}$ is chosen to be Gaussian. Therefore, for Gaussian $\mathbf{x}$, (21) can be rewritten as

$$I(X;Y|\hat{\mathbf{h}}_{\text{new}}) \geq \frac{T-M}{T}\big(\log(\det(\pi e Q)) - \log(\det((\mathbf{x} - \mathbf{g}y)(\mathbf{x} - \mathbf{g}y)^\dagger))\big). \quad (23)$$

We must now choose $\mathbf{g}$ appropriately. Here we choose $\mathbf{g}$ to be the MMSE estimator of $\mathbf{x}$, given y:

$$\mathbf{g} = \frac{\frac{P_d}{M}}{\frac{P_d}{M}\|\hat{\mathbf{h}}_{\text{new}}\|^2 + \mathbf{E}(|\hat{n}|^2|\hat{\mathbf{h}}_{\text{new}})} \hat{\mathbf{h}}_{\text{new}}^\dagger \quad (24)$$

After appropriately simplifying (21), (similar to Appendix A in [4]), we get

$$\begin{aligned} I(X;Y|\hat{\mathbf{h}}_{\text{new}}) \geq& \frac{T-M}{T} \log\left(1 + \frac{P_d p(\hat{\gamma})\hat{\xi}}{M\left(1 + \sigma^2 P_d p(\hat{\gamma})\right)}\right) \\ =& \frac{T-M}{T} I_{\text{lb}}(X;Y|\hat{\mathbf{h}}_{\text{new}}). \end{aligned} \quad (25)$$

This bound can be further tightened by maximizing the argument of the $\log()$ term over all possible (P_t, P_d), subject to the total power constraint

$$P_t + (T-M)P_d = PT. \quad (26)$$

Using the method of Lagrangian multipliers, we can show that the optimal P_d must satisfy

$$\begin{aligned} (T-M)(T-2M)P_d^2 - & 2(T-M)(PT+M)P_d \\ & + PT(PT-M) = 0 \end{aligned} \quad (27)$$

This equation has already been derived, using a different method, in [6].

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," Euproean Trans. Telecomm., Vol. 10, No. 6, Nov. 1999. pp 585-596.

[2] E. Biglieri, G. Caire, and G. Taricco, "Limiting Performance of block-fading channels with multiple antennas," IEEE Transactions on Information Theory, vol. 47, no. 4, pp. 1273-1289, May 2001.

[3] E. Visotsky and U. Madhow, "Space-time precoding with imperfect feedback," in Proceedings of ISIT 2000, Sorrento, Italy, p. 312, June 2000.

[4] S. Bhashyam, A. Sabharwal and B. Aazhang, "Feedback gain in multiple antenna systems," IEEE Transactions on Communications, Vol. 50, No. 5, May 2002. pp 785-798.

[5] K. K. Mukkavilli, A. Sabharwal, E. Erkip and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems", IEEE Transactions on Information Theory, vol. 49, no. 10, pp. 2562-2579, October 2003.

[6] B. Hassibi and B.M. Hochwald, "How much training is needed in multiple-antenna wireless links?", IEEE Transactions on Information Theory, vol. 49, no. 4, pp 951-963, April 2003.

[7] L. E. Elsgolc, "Calculus of variations," International Series of Monographs in Pure and Applied Mathematics, Vol. 19, Pergamon Press.

[8] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," IEEE Transactions on Information Theory, Vol. 43, No. 6, Nov. 1997. pp 1986-1992.