On BER Analysis of the BLAST without Optimal Ordering over Rayleigh Fading Channel

S. Loyka
School of Information Technology and Engineering (SITE), University of Ottawa, 161 Louis Pasteur, Ottawa, Ontario, Canada, K1N 6N5
E-mail: sergey.loyka@ieee.org

F. Gagnon
Department of Electrical Engineering, Ecole de Technologie Superieure
1100, Notre-Dame St. West, Montreal
Quebec, H3C 1K3, Canada

Abstract— **The BLAST algorithm is simple and, hence, popular solution for a signal processing at the MIMO receiver. Its BER performance has been studied mainly using numerical (Monte-Carlo) techniques since exact analytical analysis presents serious difficulties. Close examination of the problem of BLAST BER performance analysis reveals that the major difficulty for analytical analysis is due to the optimal ordering procedure. Hence, we analyze the algorithm performance without optimal ordering. While this is a disadvantage of the analysis, there are certain advantages as well. Exact closed-form analytical analysis is possible in the general case of mxn system for i.i.d. Rayleigh channel, which provides deep insight and understanding that cannot be gained using Monte-Carlo approach alone. We present closed-form expressions for instantaneous and average BER at each detection step, which hold true for any modulation format and take simple form in some cases (i.e. BPSK, QPSK, BFSK etc.). Asymptotic form (for large SNR) of these expressions is especially simple. As expected, the first-step BER dominates the total BER and D-BLAST provides some improvement over V-BLAST in terms of the "worst" BER, which, however, disappears asymptotically. We stress that this conclusions hold true for any modulation. Comparison to the V-BLAST with optimal ordering allows to better understand the advantages provided by the optimal ordering procedure.**

Keywords-MIMO, BLAST, Performance Analysis, BER

I. Introduction

The BLAST algorithm is simple and, hence, popular solution for a signal processing at the MIMO receiver [1-3]. Its BER performance has been studied mainly numerically (Monte-Carlo techniques) since analytical analysis presents serious difficulties, especially when no bounds or approximations are used. While 2xn system (i.e. with 2 Tx and n Rx antennas) can be analytically analyzed in a closed form without any approximation or bounds [4,5], the extension of the analysis to general case of mxn system has not been found yet. Hence various bounds and approximations have been employed to attack the problem [6]. Consequently, the solutions found are limited in some ways.

Here we adopt a different approach. Close examination of the problem of BLAST BER performance analysis reveals that the major difficulty for closed-from exact analytical analysis is due to the optimal ordering procedure. Hence, we analyze the algorithm performance without optimal ordering. Clearly, this is a disadvantage of the analysis. However, there are certain advantages as well: (i) closed-form exact analytical analysis is possible in the general case of mxn system, (ii) this provides deep insight and understanding that cannot be gained using Monte-Carlo approach alone, (iii) there exists a hope that the techniques developed can be further extended to account for the optimal ordering, (iv) comparing the performance of the no-ordering algorithm to that with the optimal ordering allows one to better understand the advantages provided by the ordering and various differences in the performance, (v) extension to the D-BLAST is straightforward.

II. The V-BLAST Algorithm

The main idea of the BLAST architecture is to split the information bit stream into several sub-streams and transmit them in parallel using a set of Tx antennas (the number of Tx antennas equals the number of sub-streams) at the same time and frequency. At the Rx side, each Rx antennas "sees" all the transmitted signals, which are mixed due to the nature of the wireless propagation channel. Using appropriate signal processing at the Rx side, these signals can be unmixed so that the matrix wireless channel is transformed into a set of virtual parallel independent channels (provided that the multipath is rich enough).

The standard baseband system model is given by

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \boldsymbol{\xi} \tag{1}$$

where $\mathbf{s}$ and $\mathbf{r}$ are the Tx and Rx vectors correspondingly, $\mathbf{H}$ is the mxn channel matrix, i.e. the matrix of the complex channel gains between each Tx and each Rx antenna, n is the number of Rx antennas, m is the number of Tx antennas, $n \geq m$, and $\boldsymbol{\xi}$ is the additive white Gaussian noise (AWGN), which is assumed to be $\mathcal{CN}(0, \sigma_0^2 \mathbf{I})$, i.e. independent and identically distributed (i.i.d.) in each branch.

The job of the V-BLAST algorithm is to find $\mathbf{s}$ given $\mathbf{r}$ and $\mathbf{H}$ in a computationally-efficient way. The V-BLAST processing begins with the 1st Tx symbol and proceeds in sequence to the m-th symbol. When the optimal ordering procedure is employed, the Tx indexing is changed prior to the processing. The main steps of the algorithm are as follows [1,3]:

(1) *The interference cancellation step*: at the i-th processing step (i.e., when the signal from the i-th transmitter is detected) the interference from the first i-1 transmitters can be subtracted based on the estimations of the Tx symbols and the knowledge of the channel matrix $\mathbf{H}$,

0-7803-8521-7/04/$20.00 © 2004 IEEE

$$\mathbf{r}_i' = \mathbf{r} - \sum_{j=1}^{i-1} \mathbf{h}_j \hat{s}_j \quad (2)$$

where $\mathbf{h}_j$ is the j-th column of $\mathbf{H}$, and $\hat{s}_j$ are the detected symbols (which are assumed to be error-free).

(2) *The interference nulling step*: based on the knowledge of the channel matrix, the interference from yet-to-be-detected symbols can be nulled out using the Gramm-Schmidt orthogonalization process (applied to the column vectors of **H**) and orthogonal projection on the sub-spaced spaced by yet-to-be-detected symbols,

$$\mathbf{r}_i'' = \mathbf{P}_i \mathbf{r}_i' \quad (3)$$

where $\mathbf{P}_i$ is the projection matrix on the sub-space orthogonal to that spanned by $\{\mathbf{h}_{i+1}\mathbf{h}_{i+2}...\mathbf{h}_m\}$: $\mathbf{P}_i = \mathbf{I} - \mathbf{H}_i(\mathbf{H}_i^+\mathbf{H}_i)^{-1}\mathbf{H}_i^+$, where $\mathbf{H}_i = [\mathbf{h}_{i+1}\mathbf{h}_{i+2}...\mathbf{h}_m]$ [7].

(3) *The optimal ordering procedure*: the order of symbol processing is organized according to their after-processing SNRs in the decreasing order, i.e. the symbol with highest SNR is detected first.

III. Analysis of the V-BLAST Algorithm

The following basic assumptions are employed in the present paper:

(1) The channel is random, quasistatic (i.e. fixed for every frame of information bits but varying from frame to frame), frequency independent (i.e., negligible delay spread); the components of H are $\mathcal{CN}(0,\mathbf{I})$ (i.e., i.i.d. Rayleigh fading with unit average power gain).

(2) Equal-power constellation is used.

(3) The Tx signals, noise and channel gains are independent of each other.

(4) Perfect channel knowledge is available at the receiver.

(5) There is no performance degradation due to synchronization and timing errors.

As it was indicated above, the optimal ordering procedure will be omitted in the present paper. We follow the approach to V-BLAST analysis proposed in [4,5], where it was shown that the conditional (i.e. assuming no demodulation error at the first (i-1) steps) after-processing instantaneous (i.e. for given channel instant) signal power y_i at i-th processing step is

$$y_i \sim \chi^2_{2(n-m+i)} , \quad (4)$$

where ~ means equal in distribution, and different y_i are independent of each other. The i-th step has diversity order equal to (n-m+i), the smallest one being at the 1st step and the largest – at the last one. Note that the fact that the distribution is conditional (no error propagation) does not limit the analysis since, as we show below, the conditional distribution is sufficient to find the block error rate (BLER) and outage probability. The distribution of y_i follows also from the Bartlett decomposition of the complex Wishart matrix.

The best way to improve the output SNR is to use maximum ratio combining (MRC) after the interference nulling out step. However, the well-known expressions for the MRC output SNR cannot be applied directly since the orthogonal projection during the nulling-out step results in correlated branch noise. The after-projection noise correlation matrix is

$$\mathbf{C}_\xi = \sigma_0^2 \mathbf{P}_i , \quad (5)$$

The MRC combining weights $\boldsymbol{\alpha}$ are given in this case by the solution of the following generalized eigenvalue problem,

$$\left(\mathbf{C}_y - \gamma \mathbf{C}_\xi\right)\boldsymbol{\alpha} = 0 , \quad (6)$$

where $\mathbf{C}_y$ is the after-projection signal covariance matrix. Remarkably, as detailed analysis demonstrates (see Appendix A), the output SNR is still the same as that for i.i.d. branch noise,

$$\gamma_i = \sigma_0^{-2} \mathbf{y}_i^{+} \mathbf{y}_i , \quad (7)$$

where $\mathbf{y}_i = \mathbf{P}_i \mathbf{H}_{i-1} \mathbf{s}$ is the after-projection signal. We stress that this is a non-trivial result that holds true because of the special structure of the projection matrix $\mathbf{P}$.

As the last equation demonstrates, the output SNR is proportional to the output signal power. Hence, the statistical analysis in terms of the output signal power is the same as that in terms of SNR. Using the distribution of the signal power at each detection step given above, the conditional outage probability can be found immediately,

$$F_i(x) = \Pr\left[\gamma_i / \gamma_0 < x\right] = 1 - e^{-x} \sum\nolimits_{k=0}^{n-m+i-1} x^k / k! , \quad (8)$$

where γ_0 is the average per-branch SNR, and the corresponding pdf is

$$f_i(x) = dF_i(x)/dx = x^{n-m+i-1} e^{-x} / (n-m+i-1)! , \quad (9)$$

Using this, the conditional (i.e. given no error at first (i-1) steps) average (over the channel statistics) BER at i-th step can be expressed in the standard form,

$$\overline{P_{e,i}} = \gamma_0^{-1} \int_0^\infty P_e(\gamma) f_i(\gamma/\gamma_0) d\gamma , \quad (10)$$

where $P_e(\gamma)$ is the instantaneous BER (i.e. BER for given SNR γ), which is determined by the modulation format. Noting that f_i is χ^2 density, the average BER can be expressed in closed form for many modulation formats. The instantaneous conditional BER at step i is simply $P_e(\gamma_i)$, and, using the Bayes formula, the instantaneous unconditional BER (i.e. including the error propagation from first (i-1) steps) is

$$\widetilde{P}_{e,i} = \sum\nolimits_{j=1}^{i} P_e(\gamma_j) \prod\nolimits_{k=1}^{j-1} (1 - P_e(\gamma_k)) , \quad (11)$$

where the product gives the probability of no error at first (j-1) steps, and the entire expression gives the probability of at least one error at first i steps assuming that, due to the error propagation, an error at any step form 1 to (i-1) will result at an error at step i. We stress that this expression accounts for the effect of error propagation (however, in an exaggerated form, i.e. "100%" error propagation, with the true one being somewhat less than that). Defining a block error rate (BLER) as a probability of having at least one error at the demodulated Tx vector, one can express it as

$$\widetilde{P}_b = \widetilde{P}_{e,m} = \sum_{j=1}^{m} P_e(\gamma_j) \prod_{k=1}^{j-1} (1 - P_e(\gamma_k)) . \quad (12)$$

We emphasize that this is a rigorous expression, which accounts for the true error propagation. The total instantaneous BER (i.e. when all the sub-streams are merged together after demodulation) is

$$\widetilde{P}_t = m^{-1} \sum_{i=1}^{m} \widetilde{P}_{e,i} , \quad (13)$$

Using the fact that γ_i and γ_k are independent for $i \neq k$, the (unconditional) average BER at i-th step is given by

$$\overline{\widetilde{P}}_{e,i} = \sum_{j=1}^{i} \overline{P}_{e,j} \prod_{k=1}^{j-1} (1 - \overline{P}_{e,k}) , \quad (14)$$

and the average BLER is

$$\overline{\widetilde{P}}_b = \overline{\widetilde{P}}_{e,m} = \sum_{j=1}^{m} \overline{P}_{e,j} \prod_{k=1}^{j-1} (1 - \overline{P}_{e,k}) \quad (15)$$

Let us now consider its asymptotic behavior for large average SNR, $\gamma_0 >> 1$. In this case, $\overline{P}_{e,i} << 1$ and the product term represents a second-order effect which can be neglected,

$$\overline{\widetilde{P}}_{e,i} \approx \sum_{j=1}^{i} \overline{P}_{e,j} \approx \overline{P}_{e,1}, \;\; \overline{\widetilde{P}}_b \approx \sum_{i=1}^{m} \overline{P}_{e,i} \approx \overline{P}_{e,1} , \quad (16)$$

where the last equality is due to the fact that the diversity order increases with the step number i, the smallest one being at step 1, which results asymptotically in $\overline{P}_{e,1} >> \overline{P}_{e,2} >> ... >> \overline{P}_{e,m}$. Clearly, the 1st step BER has the dominant effect, which agrees well with intuitive expectation based on the diversity order argument. The analysis above, however, gives more detailed and precise picture. Note also that the 1st approximation in (16) is more accurate than the 2nd one.

IV. Analysis of the D-BLAST Algorithm

It is straightforward to extend this analysis to the D-BLAST algorithm. We assume that the sub-stream launched by i-th antenna is always detected at i-th step regardless of the Tx the antenna is connected to. Then, the antenna cycling results in instantaneous unconditional BER of each sub-stream being the same and equal to the mean unconditional step BER of the corresponding V-BLAST system,

$$\widetilde{P}_{e,s} = m^{-1} \sum_{i=1}^{m} \widetilde{P}_{e,i} = m^{-1} \sum_{i=1}^{m} (m+1-i) \widehat{P}_{e,i} , \quad (17)$$

where $\widehat{P}_{e,j} = P_e(\gamma_j) \prod_{k=1}^{j-1} (1 - P_e(\gamma_k))$. Note that $\widetilde{P}_{e,s}$ is identical the total instantaneous BER since all the sub-stream BERs are the same. It is also the same as the V-BLAST total BER (13). Hence, the D-BLAST does not provide any improvement over the V-BLAST in terms of the total instantaneous BER. The only difference is that all the D-BLAST sub-streams have the same BER, while some V-BLAST sub-streams are "worse" than the average, which constitutes a bottle-neck of the system performance (unless some special measures are taken to equate them). Clearly, the "worst" sub-stream performance of the D-BLAST is better than that of the V-BLAST,

$$\widetilde{P}_{e,s} = m^{-1} \sum_{i=1}^{m} (m+1-i) \widehat{P}_{e,i} < \widetilde{P}_{e,m} = \sum_{i=1}^{m} \widehat{P}_{e,i} , \quad (18)$$

Taking the expectation of (16) over the channel statistics, the average step BER can be found,

$$\overline{\widetilde{P}}_{e,s} = m^{-1} \sum_{i=1}^{m} (m+1-i) \overline{\widehat{P}}_{e,i} , \quad (19)$$

where $\overline{\widehat{P}}_{e,j} = \overline{P}_{e,j} \prod_{k=1}^{j-1} (1 - \overline{P}_{e,k})$. The same improvement over the V-BLAST in terms of the "worst" BER can be observed. It should be noted that similar results hold true when the optimal ordering is used as well (since (17) holds true in that case as well). In the high SNR regime,

$$\overline{\widetilde{P}}_{e,s} \approx m^{-1} \sum_{i=1}^{m} (m+1-i) \overline{P}_{e,i} \approx \overline{P}_{e,1} . \quad (20)$$

Hence, asymptotically (for large average SNR), the advantage of the D-BLAST disappears since both algorithms have the same "worst" BER equal to $\overline{P}_{e,1}$, i.e. the 1st step BER is dominant for both algorithms. We emphasize that the results above hold true for any modulation format as no specific details of $P_e(\gamma)$ have been used (except for the asymptotic behavior which is the same for any modulation up to a constant). Application of these generic results to a particular modulation format is straightforward and can be done analytically in closed-form in some cases. Finally, we would like to note that most of the results above hold true for fading channels other than Rayleigh one.

V. Monte-Carlo Simulations

In order to validate the analytical results above, extensive Monte-Carlo simulations have been undertaken. Specifically, we used the Rayleigh i.i.d. fading channel and BPSK modulation demodulated coherently. First, the instantaneous BER expressions have been validated. No statistically-significant difference between analytical and MC results have been found for conditional BER (i.e. without error propagation). However, for unconditional BER (true one), it was found that the "100%" error propagation assumption results in overestimated BER. Hence, the actual error propagation rate is less than 100%. As expected, the BLER is predicted accurately. Some of the results are shown in Fig. 1. Clearly, 1st step BER dominates the BLER for high SNR (>5 dB). Note also that the error propagation has significant effect on the 2nd step BER, as comparison to the "no error propagation" BER demonstrates. However, noting that the "100%" error propagation would result in the 2nd step BER being equal to the BLER, we conclude that the actual error propagation is smaller. In terms of the average SNR, the difference is about 6 dB.

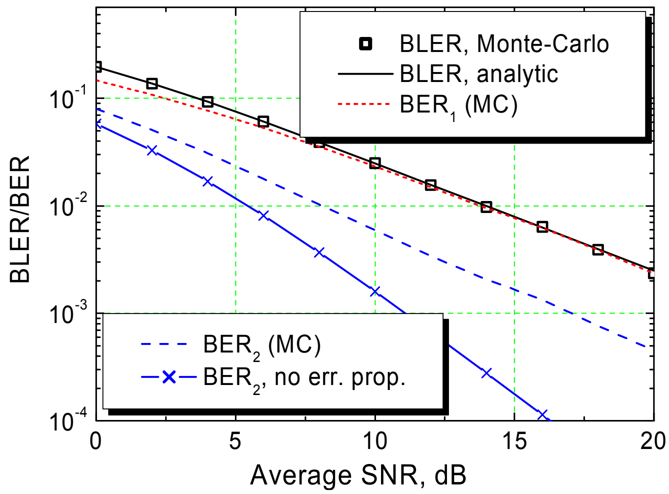

Fig. 1. BLER/BER of 2x2 V-BLAST

References

[1] G.J. Foschini et al, Analysis and Performance of Some Basic Space-Time Architectures, IEEE Journal Selected Areas Comm., v. 21, N. 3, pp. 281-320, April 2003.

[2] G.J Foschini, 'Layered space-time architecture for wireless communication in a fading environment when using multiple antennas', Bell Lab. Tech. J., vol. 1, N. 2, pp. 41-59, 1996.

[3] G.J Foschini et al, Simplified Processing for High Spectral Efficiency Wireless Communication Employing Multi-Element Arrays, IEEE Journal on Selected Areas in Communications, v. 17, N. 11, pp. 1841-1852, Nov. 1999.

[4] S. Loyka, V-BLAST Outage Probability: Analytical Analysis, IEEE Vehicular Technology Conference, 24–28 September 2002, Vancouver, Canada.

[5] S. Loyka, F. Gagnon, Performance Analysis of the V-BLAST Algorithm: an Analytical Approach, IEEE Trans. Wireless Comm., July 2004.

[6] S. Loyka, F. Gagnon, Analytical Framework for Outage and BER Analysis of the V-BLAST Algorithm, International Zurich Seminar on Communications, Feb. 18-20 2004, ETH Zurich, Switzerland.

[7] T.W. Anderson, An Introduction to Multivariate Analysis (3rd Ed.), Wiley, 2003.

[8] H.L. Van Trees, Optimum Array Processing, Wiley, 2002.

[9] W.C. Jakes Jr., Microwave Mobile Communication. New York: Wiley, 1974.

[10] J.G. Proakis, Digital Communications, McGraw Hill, Boston, 2001.

Appendix A

Consider the after-projection noise vector,

$$\xi_i = \mathbf{P}_i\xi,\ \ \mathbf{P}_i = \mathbf{I} - \mathbf{H}_i^+(\mathbf{H}_i\mathbf{H}_i^+)^{-1}\mathbf{H}_i \tag{A1}$$

Since the correlation matrix of the original noise vector is

$$\left\langle \xi\xi^+ \right\rangle = \sigma_0^2\mathbf{I}, \tag{A2}$$

where the expectation is over the noise, the correlation matrix of the projected noise is

$$\mathbf{C}_{\xi_i} = \left\langle \xi_i\xi_i^+ \right\rangle = \sigma_0^2\mathbf{P}_i\mathbf{P}_i^+ = \sigma_0^2\mathbf{P}_i, \tag{A3}$$

where we used the following property of the projection matrix: $\mathbf{P}_i = \mathbf{P}_i^+ = \mathbf{P}_i\mathbf{P}_i^+$. Clearly, the noise is correlated. Let us consider the MRC for the correlated noise case. To simplify the notations, we further drop index *i*. The output signal of the combiner is

$$r_{out} = \alpha^+\mathbf{y} + \alpha^+\xi \tag{A4}$$

where α^+ is the weight vector. The output SNR (to be maximized) is

$$\gamma = \frac{P_y}{P_\xi} = \frac{\alpha^+\mathbf{C}_y\alpha}{\alpha^+\mathbf{C}_\xi\alpha} \tag{A5}$$

where P_y and P_ξ are the signal and noise powers at the combiner output, and $\mathbf{C}_y = \mathbf{y}\mathbf{y}^+$ is the signal instantaneous correlation matrix. The weight vector α^+ that maximizes (A5) can be found by taking the derivative $\partial\gamma/\partial\alpha^+$ and setting it to zero [7],

$$\frac{\partial\gamma}{\partial\alpha^+} = 0 \to \mathbf{C}_y\alpha = \gamma\mathbf{C}_\xi\alpha \tag{A6}$$

This is the classical eigenvalue problem, which could be reduced to the following if the noise correlation matrix were nonsingular ($\det\mathbf{C}_\xi \neq 0$),

$$\mathbf{C}_\xi^{-1}\mathbf{C}_y\alpha = \gamma\alpha \tag{A7}$$

Clearly, the optimum weight vector is the eigenvector of $\mathbf{C}_\xi^{-1}\mathbf{C}_y$ corresponding to the largest eigenvalue equal to γ . Unfortunately, (A7) cannot be used since $\mathbf{C}_\xi$ is singular, $\det\mathbf{C}_\xi = 0$ (this can be seen from (A3) by noting that the projection matrix has (*m-i*) zeros as its eigenvalues, with the rest of eigenvalues being equal to 1). Hence, we have to solve the generalized eigenvalue problem,

$$\left(\mathbf{C}_y - \gamma\mathbf{C}_\xi\right)\alpha = 0 \tag{A8}$$

To this end, we use unitary transformation $\mathbf{U}$ such that

$$\mathbf{UP} = \begin{bmatrix} \mathbf{I}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{A9}$$

where $\mathbf{I}_i$ is (*ixi*) identity matrix. This can always be done by introducing new orthonormal basis vectors $\{\mathbf{e}_1\mathbf{e}_2...\mathbf{e}_m\}$ in such a way that $Span\{\mathbf{e}_1\mathbf{e}_2...\mathbf{e}_i\} = Span\{\mathbf{h}_1\mathbf{h}_2...\mathbf{h}_i\}$ (in fact, we split the entire space into the signal sub-space $Span\{\mathbf{h}_1\mathbf{h}_2...\mathbf{h}_i\}$ and its orthogonal complement $Span\{\mathbf{h}_{i+1}\mathbf{h}_{i+2}...\mathbf{h}_m\}$, which is cancelled by the orthogonal projection). Applying the unitary transformation $\mathbf{U}$ to (A8), one obtains

$$\left(\mathbf{C}_y{}' - \gamma\mathbf{C}_\xi{}'\right)\boldsymbol{\alpha}' = 0 \quad \text{(A10)}$$

where $\mathbf{C}_y{}' = \mathbf{U}\mathbf{C}_y\mathbf{U}^+, \mathbf{C}_\xi{}' = \mathbf{U}\mathbf{C}_\xi\mathbf{U}^+, \boldsymbol{\alpha}' = \mathbf{U}\boldsymbol{\alpha}$. Further, we note that

$$\mathbf{y}' = \mathbf{U}\mathbf{y} = [y_1 y_2 ... y_i 0...0]^T \rightarrow \mathbf{C}_y{}' = \begin{bmatrix} \tilde{\mathbf{C}}_\mathbf{y} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \text{(A11)}$$

where $\tilde{\mathbf{C}}_\mathbf{y} = \tilde{\mathbf{y}}\tilde{\mathbf{y}}^+, \tilde{\mathbf{y}} = [y_1 y_2 ... y_i]^T$, i.e. we drop the orthogonal sub-space components, which are equal to zero. Similarly,

$$\mathbf{C}_\xi{}' = \sigma_0^2 \begin{bmatrix} \mathbf{I}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \tilde{\boldsymbol{\alpha}} = [\alpha_1 \alpha_2 ... \alpha_i]^T \quad \text{(A12)}$$

Clearly, the optimization of $\boldsymbol{\alpha}^+$ can be done in the signal sub-space without loss of generality. Noting that the noise is uncorrelated in that sub-space, the classical MRC solution applies. Alternatively, (A10) can be presented as

$$\tilde{\mathbf{y}}\tilde{\mathbf{y}}^+\tilde{\boldsymbol{\alpha}} = \gamma\sigma_0^2\tilde{\boldsymbol{\alpha}} \quad \text{(A13)}$$

which results in the same MRC solution,

$$\tilde{\boldsymbol{\alpha}} = a\tilde{\mathbf{y}} \quad \text{(A14)}$$

where a is a scalar constant, which does not affect the output SNR and, hence, can be chosen arbitrarily. We chose $a = 1$. Adding the zero components of the orthogonal sub-space to (A14) and multiplying it by $\mathbf{U}^+$, one obtains

$$\boldsymbol{\alpha} = \mathbf{y} \quad \text{(A15)}$$

i.e. despite of the correlated noise, the classical MRC solution still applies (it becomes clear when one sets $a = 1/\sigma_0^2$ rather than 1)! The output noise power is

$$P_\xi = \boldsymbol{\alpha}^+\mathbf{C}_\xi\boldsymbol{\alpha} = \tilde{\boldsymbol{\alpha}}^+\tilde{\mathbf{C}}_\xi\tilde{\boldsymbol{\alpha}} = \sigma_0^2\tilde{\mathbf{y}}^+\tilde{\mathbf{y}} = \sigma_0^2\mathbf{y}^+\mathbf{y} \quad \text{(A16)}$$

Finally, the output SNR can be expressed as

$$\gamma = \frac{\mathbf{y}^+\mathbf{y}}{\sigma_0^2} \quad \text{(A17)}$$

Remarkably, the output SNR is not affected by the noise correlation and is the same as if the noise were i.i.d. in each branch (after projection). We attribute it to the special structure of the projection matrix (A1). (A17) clearly indicates that the analysis in terms of the total signal power $\mathbf{y}^+\mathbf{y}$ is equivalent (up to a constant) to the analysis in terms of the SNR γ if the MRC is used after the orthogonal projection.