DIMENSIONALITY LOSS IN MIMO COMMUNICATION SYSTEMS

Sergey Loyka[1], Ammar Kouki[2]

[1]School of Information Technology and Engineering (SITE)
University of Ottawa, 161 Louis Pasteur, Ottawa, Ontario,
Canada, K1N 6N5
Email: sergey.loyka@ieee.org

[2]Department of Electrical Engineering
Ecole de Technologie Superieure
1100, Notre-Dame St. West
Montreal (Quebec), H3C 1K3, Canada

Abstract

Traditional viewpoint is that the effect of correlation is to decrease MIMO capacity. In this paper, the effect of correlation is studied from a different perspective: a concept of effective dimensionality (ED) of a MIMO system is introduced and investigated using the correlation matrix approach. It is shown that the channel correlation results in ED decrease. Simple formulas, which give an explicit dependence of ED on the channel correlation, are given for practically-important cases. A comparison of the ED concept with the recently-introduced concept of effective degrees of freedom is presented.

Key words

MIMO Channels, Capacity, Correlation

1. Introduction

Multiple-Input Multiple-Output (MIMO) communication architecture has recently emerged as a new paradigm for efficient wireless communications in rich multipath environments [1,2]. Using multi-element antenna arrays (MEA) at both transmitter and receiver, which effectively exploits the third (spatial) dimension in addition to the time and frequency dimensions, this architecture achieves channel capacities far beyond those of traditional techniques. In uncorrelated Rayleigh channels the MIMO capacity scales linearly as the number of antennas [1,2]. However, there are several limitations to the performance of this architecture in real-world conditions [1,3-6]. One of the major limitations is the correlation of individual sub-channels, i.e. links between one transmitter and one receiver antennas, of the matrix channel, which may result in severe degradation of MIMO performance [3-6]. One way to characterise this phenomenon is to consider the MIMO system dimensionality. When we have $n\times n$ MIMO system (i.e. n transmit and n receive antennas) and n parallel independent sub-channels, we say that the system dimensionality is n. The MIMO channel capacity achieved under these conditions (and also under some reasonable assumptions) is maximum. If some, or all, of the sub-channels are correlated, the channel capacity decreases. Spatial dimensions, which are used through n transmit and n receive antennas, provide smaller advantage in this case. Hence, we say that the system *effective dimensionality* (ED) decreases.

We give a definition of the MIMO system effective dimensionality through a comparison of the MIMO capacity for correlated and uncorrelated channels, and investigate it numerically and analytically using the correlation matrix approach. We show that strong correlation between some subchannels results in ED reduction. We also give a comparison of the effective dimensionality with the number of effective degrees of freedom (EDOF) introduced in [3].

By introducing the ED concept, we try to isolate and study the effect of correlation. In the ideal MIMO space, all the "dimensions" are orthogonal and the capacity achieves its maximum. For an actual MIMO space, some or all "dimensions" are non-orthogonal (correlated) and the capacity decreases. From a viewpoint of system performance, it is equivalent to going into a space of fewer dimensions. In fact, ED tells us how many of the actual number of transmit/receive branches are effectively used. In a sense, the ED concept is similar to the power efficiency concept, when lossless (ideal) and lossy (actual) systems are compared.

2. Effective Dimensionality of a MIMO System

Under some reasonable assumptions, the channel capacity of $n\times n$ MIMO system is [1]:

$$C=\log_2 \det\left(\mathbf{I}+\rho\mathbf{H}\cdot\mathbf{H}^{+}/n\right) \text{ bits/s/Hz} \tag{1}$$

where n is the number of transmit/receive antennas, ρ is the signal-to-noise ratio (SNR), $\mathbf{I}$ is n×n identity matrix, $\mathbf{H}$ is the normalized channel matrix, and "$^{+}$" means transpose conjugate. The following normalization of $\mathbf{H}$ is adopted in this paper [5]: $tr\left(\mathbf{HH}^{+}\right)=n$, where tr means trace. Then ρ/n is per-subchannel SNR. Another normalization can also be used but ρ/n will have a different meaning in this case. For simplicity, we further consider the case of equal received powers in every receive branch. In this case $\sum_k |h_{ik}|^2=1$ and (1) simplifies to

$$C(\mathbf{R},n)=\log_2 \det\left(\mathbf{I}+\rho\mathbf{R}/n\right), \tag{2}$$

where **R** is the normalized channel correlation matrix, $|r_{ij}| \le 1$, whose components are $r_{ij} = \sum_k h_{ik} h_{jk}{}^*$, where "*" denotes complex conjugate. In fact, r_{ij} is the correlation coefficient of i-th and j-th receive branches. Strictly speaking, this definition is valid for fixed channels only. However, an expectation operator can be employed for stochastic channels. (2) emphasizes that the MIMO channel capacity $C(\mathbf{R},n)$ is a function of the correlation matrix **R** and of the number n of antennas. We define the MIMO effective dimensionality n_e from the following equation:

$$C(\mathbf{R},n) = C(\mathbf{I},n_e), \tag{3}$$

where $C(\mathbf{I},n_e)$ is the MIMO capacity of uncorrelated channel and is given by [1]:

$$C(\mathbf{I},n_e) = n_e \cdot \log_2 (1 + \rho / n_e) \tag{4}$$

Thus, the effective dimensionality is the number of dimensions of a MIMO system operating over an uncorrelated channel, which has the same channel capacity as the actual system operating over the actual correlated channel (the signal-to-noise ratio ρ being the same for both cases). The ED shows how efficiently we use the actual receive and transmit branches and is, hence, a system performance parameter. In general, (3) is a transcendental equation and cannot be solved analytically for n_e. Numerical methods can be applied to solve it. No convergence problems are anticipated since both sides of (3) are monotonous functions of n. For some specific cases, analytical techniques can be used which allow us to gain some insight and to obtain simple analytical solutions for practically-important cases.

A similar concept has been introduced in [3] as the number of effective degrees of freedom of a MIMO system:

$$EDOF = \frac{d}{d\delta} C(2^{\delta}\rho)\Big|_{\delta=0} \tag{5}$$

In some cases, the ED and the EDOF give very close values while in some other cases their values are very different. We give below a comparative analysis of the ED and the EDOF for some important cases and discuss the implication of their differences for practical system performance analysis. Similar parameters have also been considered in [7].

3. The Correlation Matrix Approach

Let us now consider the case where the correlation matrix **R** has a block diagonal structure:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-k} \end{bmatrix} \tag{6}$$

where $\mathbf{I}_{n-k}$ is the $(n-k)\times(n-k)$ identity matrix, **0** is the zero matrix, and $\mathbf{R}_k$ is a $k \times k$ correlation sub-matrix with non-zero components. In this model, only k receive branches are correlated; the rest (n-k) branches are independent.

In order to demonstrate how the ED concept works, we adopt here the exponential model[1] for $\mathbf{R}_k$, which can approximate some realistic scenarios [5]:

$$[\mathbf{R}_k]_{ij} = \begin{cases} r^{j-i}, & i \le j \\ \left(r^{i-j}\right)^*, & i > j \end{cases}, \quad |r| \le 1 \tag{7}$$

where r is the (complex) correlation coefficient of adjacent receive branches. Using (6) and (7), $C(\mathbf{R},n)$ can be presented in the following form:

$$C(\mathbf{R},n) = \log_2 \left(\left(1 + \frac{\rho}{n}\right)^n \det \overline{\mathbf{R}}_k \right), \tag{8}$$

where $\overline{\mathbf{R}}_k$ is the following $k \times k$ matrix:

$$\left[\overline{\mathbf{R}}_k\right]_{ij} = \begin{cases} \beta r^{j-i}, & i \le j \\ 1, & i = j \\ \beta\left(r^{i-j}\right)^*, & i > j \end{cases} \tag{9}$$

and

$$\beta = \frac{\rho}{n}\left(1 + \frac{\rho}{n}\right)^{-1} \tag{10}$$

A closed-form expression for the $det\,\overline{\mathbf{R}}_k$ is derived in Appendix A,

$$D_k = \frac{1}{2}\left[\left(\frac{\alpha+\xi}{2}\right)^n + \left(\frac{\alpha-\xi}{2}\right)^n\right] + \frac{2-\alpha}{2\xi}\left[\left(\frac{\alpha+\xi}{2}\right)^n - \left(\frac{\alpha-\xi}{2}\right)^n\right] \tag{11}$$

where $\alpha = 1 - |r|^2 (2\beta - 1)$, $\gamma = r(\beta - 1)$., $\xi = \sqrt{\alpha^2 - 4|\gamma|^2}$.

For some practically-important cases, a simplified expression may be obtained. For a system having a large signal-to-noise ratio ($r/n >> 1$) and a large number of antennas (n>>1), after some transformations which do not change the determinant, we obtain (see Appendix A for details):

$$\det\left[\overline{\mathbf{R}}_k\right] \approx \left(1 - \beta |r|^2\right)^{k-1} \tag{12}$$

Approximate solution of (3) then takes the simple form:

$$n_e \approx n - \sigma(k-1) \tag{13}$$

where

---

[1] other models of **R** (using, for example, electromagnetic simulation) can also be used for the present analysis

$$\sigma = 1 - \frac{\log_2\left(1+\frac{\rho}{n}\left(1-|r|^2\right)\right)}{\log_2\left(1+\frac{\rho}{n}\right)} \tag{14}$$

For $|r|=1$, one obtains $\sigma=1$ and, consequently, $n_e \approx n-k+1$. Thus, the reduction in system effective dimensionality Δn due to the strong correlation of k receive branches is $\Delta n = n - n_e \approx k-1$. This is a physically reasonable conclusion because we cannot transmit information independently over these k branches but have to use them as only one branch. For r=0, $\sigma=0$ and $\Delta n = 0$, and the system has full dimensionality as it should be.

The interpretation of this effect in terms of eigenvalue analysis is straightforward. For large correlation, $|r| \to 1$, the eigenvalues of $\mathbf{R}_k$ can be approximated as (see Appendix A):

$$\lambda_1 \approx ... \approx \lambda_{k-1} \approx \frac{\left(1-|r|\right)^2}{1+|r|^2} << 1, \lambda_k \approx \frac{2k-(k-1)\left(1-|r|\right)^2}{2k-(k-1)\left(1+|r|^2\right)} \tag{15}$$

Recall that the eigenvalues are the virtual channel gains. Hence, we have one full-gain channel and (k-1) channels with reduced gains (due to correlation). When $|r|=1$, $\lambda_k = k$ and $\lambda_1 = ... = \lambda_{k-1} = 0$, as expected, i.e. there is only one channel out of k, and the optimum strategy is to do the beamforming and to transmit only one bit stream over k channels. k channels are effectively collapsed into one. In the case of zero correlation, $|r|=0$, one obtains $\lambda_1 = ... = \lambda_k = 1$, as expected, i.e. there are k channels with full gains. It is interesting to note that, despite of the fact that (15) has been derived under the assumption of large correlation, $|r| \to 1$, it holds true for $|r|=0$ as well.

Fig. 1 shows Δn as a function of $|r|$ for different values of k computed using (12) and (13) and by numerical solution of (3) using (6) and (7). As one may see, (12) and (13) provide quite a good approximation. An interesting question, however, is how strong should the correlation be for $\Delta n \approx k-1$. A detailed analysis (as well as an examination of Fig. 1) gives the following rough estimation:

$$|r| \ge 1 - n/(2\rho) \tag{16}$$

For the scenario of Fig. 1, one obtains: $|r| \ge 0.995$. Thus, for almost all practical cases (when $\rho / n >> 1$) Δn will be smaller than $k-1$.

Let us now compare the ED and the EDOF concepts for a practically-important case of $\rho/n >> 1$ and n>>1. A detail analysis shows that for $|r|=1$ and $|r|=0$ the ED and the EDOF give approximately the same prediction. But in between these two extreme cases their values are different. One may wonder: why? Both parameters have very similar physical meaning and both characterize the system performance from the same viewpoint. However, in the ED concept the real system performance (channel capacity) is compared to the ideal system performance, which operates over uncorrelated parallel subchannels, for the same total transmitted power that is distributed between the effective dimensions only, not between the actual number of transmitters. On the other hand, EDOF is determined through variation in SNR, i.e. in the transmitted power,

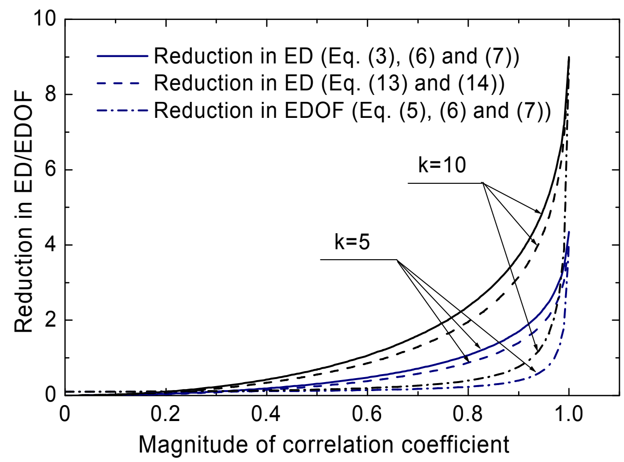

Figure 1. Reduction in ED/EDOF versus the magnitude of correlation coefficient, n=10, ρ=30 dB.

for fixed channel correlation. The total transmitted power is always distributed between the actual number of transmitters. Both concepts can be used for estimating MIMO system performance but from different viewpoints. In the EDOF concept, the effect of SNR is emphasized and, hence, it is more relevant when one wants to know how the actual system performance varies with transmitted power for fixed correlation. On the contrary, the effect of channel correlation is emphasized in the ED concept and, consequently, it is more relevant when one wants to know how the system performance varies with channel correlation for fixed power. In general, a definition of the number of dimensions of a MIMO system depends on a problem considered.

It should be noted that the concept of effective dimensionality introduced in this paper are somewhat similar to the concept of effective diversity order, which was successfully used in the analysis of diversity combining techniques over correlated channels [10].

4. Comparison to Other Correlation Matrix Models

Let us now examine the effective dimensionality loss in correlated channels for various correlation matrix models. This allows one to find how sensitive the results above are with respect to correlation model variations.

Specifically, we examine the uniform model, the three-diagonal model and the squared exponential model. For the uniform model,

$$[\mathbf{R}_k]_{ij} = \begin{cases} 1, & i = j \\ r, & i \neq j \end{cases}, \quad \mathrm{Im}\{r\} = 0\,, \tag{17}$$

assuming that r is real, the eigenvalues can be found explicitly in a closed form,

$$\lambda_1 = \ldots = \lambda_{k-1} = 1 - r\,,\ \lambda_k = 1 + (k-1)r \tag{18}$$

and the capacity can be expressed in a closed-form as well [14]. For the tri-diagonal model,

$$[\mathbf{R}_k]_{ij} = \begin{cases} 1, & i = j \\ r, & i = j-1 \\ r^*, & i = j+1 \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

the eigenvalues can be expressed in the following closed form [13],

$$\lambda_i = 1 - 2|r|\cos\frac{\pi i}{k+1}\,, \tag{20}$$

Note that the model is physical only when

$$|r| < \frac{1}{2}\left(\cos\frac{\pi}{k+1}\right)^{-1}, \tag{21}$$

i.e. for large k, $|r| < 1/2$. Otherwise, as it can be seen from (20), some eigenvalues are negative, which is not possible for a physical correlation matrix that is always positive definite. For the squared-exponential model, which can be obtained based on some physical arguments [11,12],

$$[\mathbf{R}_k]_{ij} = \begin{cases} r^{(j-i)^2}, & i \le j \\ \left(r^{(i-j)^2}\right)^*, & i > j \end{cases}, \quad |r| \le 1 \tag{22}$$

the eigenvalues are not known in a closed form and numerical techniques should be used.

Comparing all 4 models, one may note that the uniform model presents a worst-case since all the branches are equally correlated (in real world, correlation decreases for widely-separated antennas, which is not accounted for in the uniform model). The exponential model is more physical since it accounts for the correlation decrease for widely-separated antennas. The squared exponential model, which is based on some physical arguments [11,12], also accounts for the correlation decrease for distant antennas. In fact, it predicts faster decrease in correlation then the exponential model does. Finally, the three-diagonal model accounts for the correlation between neighboring antennas only assuming no correlation between distant antennas, which may be considered as the best-case scenario. The reason for this model being non-physical at $|r| > 1/2$ is that it is not physically possible to have strong correlation between adjacent antennas and no correlation for the others. Based on this reasoning, one may expect that the capacities of the three-diagonal, squared exponential, exponential and uniform models correspondingly will satisfy to the following inequality: $C_{3d} \le C_{se} \le C_e \le C_u$. However, as a detailed numerical analysis demonstrates (see Fig.2), it is not that simple. The reason behind this is that the eigenvalues and the correlation matrix entries are related in a very complicated way. For example, almost all the entries of the squared-exponential model (except for the diagonal ones) are smaller than those of the exponential model, but some of the eigenvalues of the latter are nonetheless larger than those of the former. A detailed analysis shows that the uniform model capacity is always smaller than that of the exponential model and that the squared exponential model capacity may be smaller or larger than that of the exponential model, depending on correlation and SNR values. It is interesting to note that all the models (except for the double-diagonal one, which is "pathological" at and above r=0.57, see (21)) predict almost the same behavior of the capacity versus correlation. Hence, the result seems to be of a general nature because it is quite model-independent.

Fig. 3 compares loss in ED for all 4 models. The results are again quite model-independent that suggest they may be of a general nature.

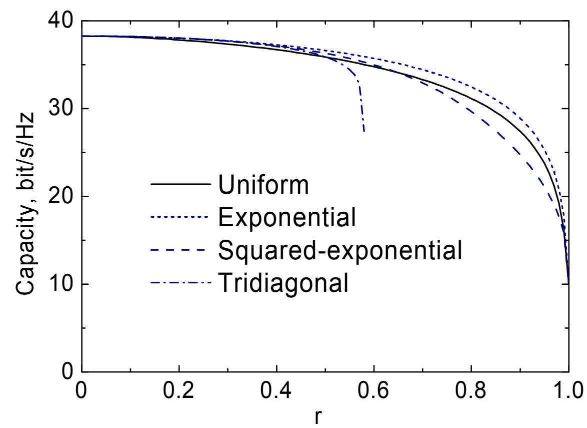

Figure 2. Capacity versus correlation coefficient for various models, n=5, ρ=30 dB.

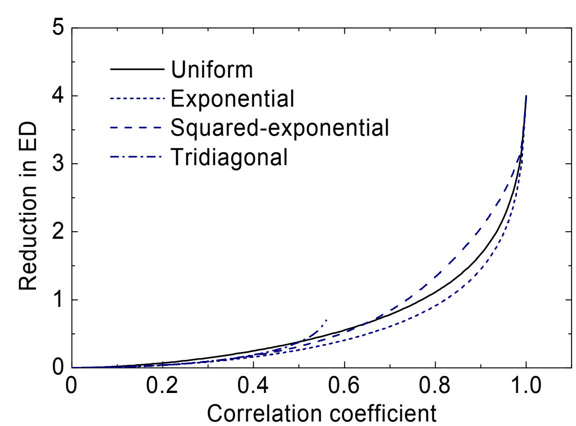

Figure 3. Loss in ED versus correlation coefficient for various models, n=5, ρ=30 dB.

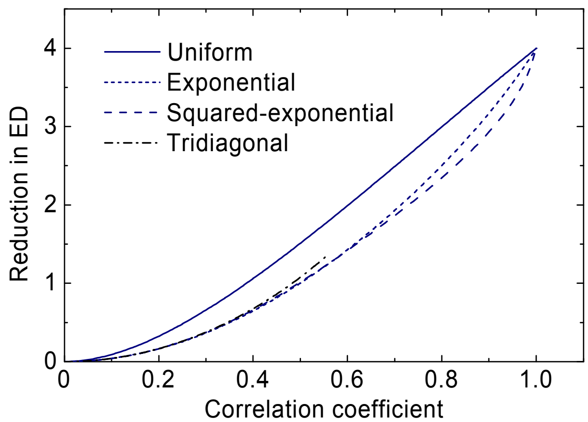

Figure 4. Loss in ED versus correlation coefficient for various models, n=5, ρ=10 dB.

As Fig. 4 suggests, there is some variation due to the SNR. In particular, the difference between the models slightly increases as the SNR decreases. The squared exponential model does provide better performance and the uniform model provides the lower bound on the performance (i.e. the worst-case scenario) at $\rho = 10dB$. The generic tendency remains however the same. Overall, as the SNR increases all the curves approach closer and closer the unit step function at r=1.

5. Conclusions

The effective dimensionality characterizes the MIMO system performance in a realistic environment, i.e., how efficiently we use multiple transmit and receive branches in a correlated channel. We have investigated it using the correlation matrix approach and have shown that for strongly-correlated k receive branches the reduction in ED is approximately k-1. Roughly speaking, ED is a factor in front of log(1+SNR) after correlation has been taken into account. Its purpose is to compare the actual system performance to that of the ideal system, whose channel capacity is maximum, and to characterize in this way the effect of correlation. For example, when k receive branches are completely correlated (r=1), this factor is (n-k+1) instead of n. The comparison of EDOF and ED shows that the former gives a more optimistic prediction of the system performance in a correlated channel.

References

[1] Foschini, G.J., Gans M.J.: ‘On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas’, Wireless Personal Communications, vol. 6, No. 3, pp. 311-335, March 1998

[2] I.E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," AT&T Bell Lab. Internal Tech. Memo., June 1995 (European Trans. Telecom., v.10, N.6, Dec.1999)

[3] Shiu, D.S., Foschini, G.J., Gans, M.J., Kahn, J.M., 'Fading Correlation and Its Effect on the Capacity of Multielement Antenna Systems,' IEEE Trans. on Communications, v. 48, N. 3, Mar. 2000, pp. 502-513.

[4] S. Loyka, G. Tsoulos, Estimating MIMO System Performance Using the Correlation Matrix Approach, IEEE Communication Letters, v. 6, N. 1, pp. 19-21, Jan. 2002.

[5] S.L. Loyka, Channel Capacity of MIMO Architecture Using the Exponential Correlation Matrix, IEEE Communication Letters, v.5, N. 9, pp. 369 –371, Sep 2001.

[6] S. Loyka, A. Kouki, New Compound Upper Bound on MIMO Channel Capacity, IEEE Communications Letters, v.6, N. 3, pp.96-98, Mar. 2002.

[7] Rayleigh, G.G., Gioffi, J.M.: "Spatio-Temporal Coding for Wireless Communications," IEEE Trans. Commun., v.46, N.3, pp. 357-366, 1998.

[8] S. Loyka, A. Kouki, Correlation and MIMO Communication Architecture (Invited), 8th International Symposium on Microwave and Optical Technology, Montreal, Canada, June 19-23, 2001

[9] M.A. Korn, T.M. Korn, Mathematical Handbook, McGraw Hill, New York, 1968.

[10] M. Schwartz, W.R. Bennet, S. Stein, Communication Systems and Techniques, IEEE Press, New York, 1996.

[11] R.B. Ertel at al, Overview of Spatial Channel Models for Antenna Array Communication Systems, IEEE Personal Communications Magazine, Feb. 1998, pp.10-22.

[12] T.S. Rappaport, Wireless Communications: Principles and Practice, Prentice Hall, New Jersey, 2002.

[13] J.N. Pierce, S. Stein, Multiple Diversity with Nonindependent Fading, IRE Proceed., Jan. 1960, v. 48, pp. 89-104.

[14] S.L. Loyka, J.R. Mosig, Channel Capacity of N-Antenna BLAST Architecture, Electronics Letters, vol. 36, No.7, pp. 660-661, Mar. 2000.

6. Apeendix A

We use the following equivalent transformation of the determinant:

$$\det \overline{\mathbf{R}}_k = D_k = \begin{vmatrix} 1 & \beta r^* & \beta r^{*2} & \dots & \beta r^{*k-1} \\ \beta r & 1 & \beta r^* & \dots & \beta r^{*k-2} \\ \beta r^2 & \beta r & 1 & \dots & \beta r^{*k-3} \\ \dots & \dots & \dots & \dots & \dots \\ \beta r^{k-1} & \beta r^{k-2} & \beta r^{k-3} & \dots & 1 \end{vmatrix} \overset{(a)}{=}$$

$$
=\begin{vmatrix} 1-\beta|r|^2 & r^*(\beta-1) & 0 & \dots & 0 \\ \beta r\left(1-|r|^2\right) & 1-\beta|r|^2 & r^*(\beta-1) & \dots & 0 \\ \beta r^2\left(1-|r|^2\right) & \beta r\left(1-|r|^2\right) & 1-\beta|r|^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \beta r^{k-1} & \beta r^{k-2} & \beta r^{k-3} & \dots & 1 \end{vmatrix} \overset{(b)}{=}
$$

$$
=\begin{vmatrix} 1-|r|^2(2\beta-1) & r^*(\beta-1) & 0 & \dots & 0 \\ r(\beta-1) & 1-|r|^2(2\beta-1) & r^*(\beta-1) & \dots & 0 \\ 0 & r(\beta-1) & 1-|r|^2(2\beta-1) & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} \tag{A1}
$$

Where we (a) first multiply each row (except for the first one) by r^* and substract it from the previous row, and (b) multiply each colum by r and substract it from the previous one (except for the first column). The last determinant can be presented in the following recursive form:

$$
D_k = \alpha D_{k-1} - |\gamma|^2 D_{k-2} \tag{A2}
$$

where $\alpha = 1-|r|^2(2\beta-1)$, $\gamma = r(\beta-1)$.. This is a difference equation [10] and its solution can be presented as:

$$
D_k = \frac{1}{2}\left[\left(\frac{\alpha+\xi}{2}\right)^n + \left(\frac{\alpha-\xi}{2}\right)^n\right] + \frac{2-\alpha}{2\xi}\left[\left(\frac{\alpha+\xi}{2}\right)^n - \left(\frac{\alpha-\xi}{2}\right)^n\right] \tag{A3}
$$

where $\xi = \sqrt{\alpha^2 - 4|\gamma|^2}$.

In high SNR mode, $\rho/n >> 1 \to \beta \approx 1$, this can be approximated as,

$$
D_k \approx \alpha^{k-1} = \left(1-|r|^2\right)^{k-1} \tag{A4}
$$

This approximation is accurate provided that r is not too close to 1.

More accurate approximation for D_k can be derived as follows [5]. We start with the second determinant in (A1) and note that, in high SNR regime, $\rho/n >> 1 \to \beta \approx 1$, the upper diagonal elements (above the main diagonal) are very small and, hence, can be neglected. Then,

$$
D_k \approx \begin{vmatrix} 1-\beta|r|^2 & 0 & 0 & \dots & 0 \\ \beta r\left(1-|r|^2\right) & 1-\beta|r|^2 & 0 & \dots & 0 \\ \beta r^2\left(1-|r|^2\right) & \beta r\left(1-|r|^2\right) & 1-\beta|r|^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \beta r^{k-1} & \beta r^{k-2} & \beta r^{k-3} & \dots & 1 \end{vmatrix} = \tag{A5}
$$

$$
=\left(1-\beta|r|^2\right)^{k-1}
$$

Using a similar approach, we can now estimate the eigenvalues of $\mathbf{R}_k$. We start with the eigenvalue equation,

$$
|\mathbf{R}_k - \lambda\mathbf{I}| = 0 \tag{A6}
$$

and, using the transformations similar to (A1), transform the determinant to the following form:

$$
|\mathbf{R}_k - \lambda\mathbf{I}| = \begin{vmatrix} \alpha & \gamma & 0 & \dots & 0 \\ \gamma^* & \alpha & \gamma & \dots & 0 \\ 0 & \gamma^* & \alpha & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1-\lambda \end{vmatrix} = 0 \tag{A7}
$$

where $\alpha = \left(1-|r|\right)^2 - \lambda\left(1+|r|^2\right)$, $\gamma = r^*\lambda$. For $|r| \to 1$, an approximate solution of (A7) can be obtained as follows. Using (A3) and keeping the first-order terms in ξ, which is very small when $|r| \to 1$, one obtains

$$
\left(\frac{\alpha}{2}\right)^{k-1}\left[\frac{\alpha}{2} + k\left(1-\lambda-\frac{\alpha}{2}\right)\right] = 0 \tag{A8}
$$

(k-1) smallest eigenvalues are

$$
\lambda_1 = \dots = \lambda_{k-1} = \frac{\left(1-|r|\right)^2}{1+|r|^2} << 1 \tag{A9}
$$

and the largest eigenvalue is

$$
\lambda_k = \frac{2k-(k-1)\left(1-|r|\right)^2}{2k-(k-1)\left(1+|r|^2\right)} \tag{A10}
$$

Note that for $|r| = 1$, $\lambda_k = k$ and $\lambda_1 = \dots = \lambda_{k-1} = 0$, as expected. Detailed numerical analysis shows that the largest and the smallest eigenvalues are predicted quite accurately, and that the prediction accuracy is worse for the eigenvalues in between. However, when computing the capacity these inaccuracies tend to compensate each other so that the capacity is predicted quite accurately.