MAXIMIZING THE DETERMINANT FOR A SPECIAL CLASS OF BLOCK-PARTITIONED MATRICES

OTILIA POPESCU, CHRISTOPHER ROSE, AND DIMITRIE C. POPESCU

An analytical solution is found for the maximum determinant of a block-partitioned class of matrices with constant trace for each block. As an immediate application of this result, the maximum determinant of a sum of Kronecker products is derived.

1. Introduction

The problem of maximizing the determinant of a matrix may arise in different areas, including information and communication theory [4, 5]. The more recent reference [4] presents an overview of the applications of the determinant maximization problem along with an algorithm for determinant maximization with linear matrix inequality constraints. In our particular case, this problem occurs in the context of sum capacity maximization for a communication system with multiple transmitters and multiple receivers [3].

For an $N \times N$ positive definite matrix $\mathbf{X}$ with constant trace, it is easily shown that $|\mathbf{X}|$ is maximized when $\mathbf{X}$ is a scaled identity matrix (Lemma 2.2).

To extend this result, we introduce a block-structured subclass of positive definite matrices ($\mathcal{M}^+$) with blocks of constant traces.

Definition 1.1. Define the class of matrices

$$
\begin{aligned}
\mathcal{D}_{N,J} = \{\mathbf{Q} \mid &\mathbf{Q} \in \mathcal{M}^+, \ \mathbf{Q} = \{\mathbf{Q}_{ij}\}, \\
&i,j = 1,\ldots,J, \ \mathbf{Q}_{ij} \in \mathbb{R}^{N \times N}, \ \mathbf{Q}_{ij} = \mathbf{Q}_{ji}, \ \mathrm{tr}[\mathbf{Q}_{ij}] = E_{ij}\}.
\end{aligned}
\tag{1.1}
$$

For the class of matrices defined above, we state the following theorem, which is our main result.

THEOREM 1.2. *Let*

$$
\mathbf{Q} = \begin{bmatrix}
\mathbf{Q}_{11} & \mathbf{Q}_{12} & \cdots & \mathbf{Q}_{1J} \\
\mathbf{Q}_{21} & \mathbf{Q}_{22} & & \mathbf{Q}_{2J} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{Q}_{J1} & \mathbf{Q}_{J2} & \cdots & \mathbf{Q}_{JJ}
\end{bmatrix},
\tag{1.2}
$$

where $\mathbf{Q} \in \mathfrak{Q}_{N,J}$. Let the trace of each subblock be fixed and equal to $\mathrm{tr}[\mathbf{Q}_{ij}] = E_{ij}$. Then the determinant of $\mathbf{Q}$ is maximized when each block is a scaled identity matrix of the form

$$\mathbf{Q}_{ij} = \frac{E_{ij}}{N}\mathbf{I}_N \tag{1.3}$$

and the maximum value of the determinant is

$$\max_{\mathbf{Q} \in \mathfrak{Q}_{N,J}} |\mathbf{Q}| = \frac{1}{N^{NJ}} |\mathbf{E}|^N, \tag{1.4}$$

where the elements of the square matrix $\mathbf{E}$ are $\{E_{ij}\}$.

2. Proof of the main result

We prove Theorem 1.2 by induction. First we prove the result for $J = 2$ and then show that the result holds for $J + 1$ assuming that it holds for J.

2.1. $J = 2$. Let $\mathbf{A}$ and $\mathbf{C}$ be $N \times N$ positive definite matrices ($\in \mathcal{M}_N^+$) with fixed traces $E_{\mathbf{A}}$ and $E_{\mathbf{C}}$, respectively. Let $\mathbf{B}$ be an $N \times N$ matrix with trace $E_{\mathbf{B}}$. Finally, we define $\mathbf{Q} \in \mathcal{M}^+$ as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{bmatrix}. \tag{2.1}$$

The matrix $\mathbf{Q}$ may be factored as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{A} & 0 \\ \mathbf{B} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B}^\top \\ 0 & \mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top \end{bmatrix} \tag{2.2}$$

implying

$$|\mathbf{Q}| = |\mathbf{A}|\,|\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top|. \tag{2.3}$$

The term $\mathbf{Q}/\mathbf{A} \equiv \mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top$ is called Schur complement of $\mathbf{A}$ in $\mathbf{Q}$ [1] and it can be easily shown that it is positive definite when $\mathbf{Q}$ is positive definite.

LEMMA 2.1. *If $\mathbf{Q} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \in \mathcal{M}^+$, then $(\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top) \in \mathcal{M}^+$.*

Proof. The matrix $\mathbf{Q}$ is positive definite. Therefore,

$$\begin{bmatrix} \mathbf{x}_1^\top & \mathbf{x}_2^\top \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} > 0 \tag{2.4}$$

for all $\mathbf{x}_i \in \mathbb{R}^N$ such that $|\mathbf{x}_1|^2 + |\mathbf{x}_2|^2 > 0$. Expanding, we have

$$\mathbf{x}_1^\top \mathbf{A}\mathbf{x}_1 + \mathbf{x}_2^\top \mathbf{B}\mathbf{x}_1 + \mathbf{x}_1^\top \mathbf{B}^\top \mathbf{x}_2 + \mathbf{x}_2^\top \mathbf{C}\mathbf{x}_2 > 0. \tag{2.5}$$

If we let $\mathbf{x}_1 = -\mathbf{A}^{-1}\mathbf{B}^\top\mathbf{x}_2$, we have

$$\mathbf{x}_2^\top\mathbf{C}\mathbf{x}_2 - \mathbf{x}_2^\top\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top\mathbf{x}_2 = \mathbf{x}_2^\top(\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top)\mathbf{x}_2 > 0 \tag{2.6}$$

for all $\mathbf{x}_2 \in \mathbb{R}^N$, which implies $(\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top) \in \mathcal{M}^+$ and completes the proof. $\qquad\square$

Another important result is the following lemma.

LEMMA 2.2. *For any $n \times n$ positive definite matrix $\mathbf{X}$ with constant trace $\mathrm{tr}[\mathbf{X}] = \alpha$, the determinant is maximized when $\mathbf{X} = (\alpha/n)\mathbf{I}_n$.*

Proof. Applying Hadamard inequality [1], the determinant of an $n \times n$ matrix $\mathbf{X}$ is maximized when the matrix is diagonal, that is, eigenvalues of the matrix are the diagonal elements. If $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, $\sum_i a_i = \alpha$, is the vector of eigenvalues of $\mathbf{X}$, from majorization theory [2], the vector $\mathbf{a}^* = (\alpha/n, \alpha/n, \ldots, \alpha/n)$, with all elements equal, is majorized by any other vector $\mathbf{a}$. Also, a majorization result says that if g is a continuous nonnegative function on $I \subset \mathbb{R}$, a function $\phi(\mathbf{x}) = \prod_{i=1}^n g(x_i)$ is Schur-concave (convex) on I^n if and only if $\log g$ is concave (convex) on I. In our case, $\log x$ is a concave function on $\mathbb{R}^+$ and $\det(\mathbf{X}) = \prod_{i=1}^n a_i$ is a Schur-concave function and its maximum is attained for $\mathbf{a}^*$. Having all eigenvalues equal is equivalent to saying that $\mathbf{X}$ is a scaled identity matrix, under its trace constraint. $\qquad\square$

This lemma shows that the determinant of a matrix with constant trace is maximized when the matrix is a scaled identity matrix. Maximizing the trace of the matrix implies that the determinant is absolutely maximized when the matrix is a scaled identity matrix with maximum trace. Further more, if a positive definite matrix can be written as the difference of two other positive definite matrices, $\mathbf{X} = \mathbf{X}_1 - \mathbf{X}_2$, it is obvious that $\mathrm{tr}[\mathbf{X}]$ is maximized when $\mathrm{tr}[\mathbf{X}_1]$ is maximized and $\mathrm{tr}[\mathbf{X}_2]$ is minimized. In our case, we consider

$$\mathbf{X} = \mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top. \tag{2.7}$$

Then we can say

$$\max_{\mathbf{A},\mathbf{C},\mathbf{B}} |\mathbf{C} - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top| = \max_{\mathbf{A},\mathbf{C}} \left[\frac{1}{N}\left(\mathrm{tr}[\mathbf{C}] - \min_{\mathbf{B}}\mathrm{tr}[\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top] \right) \right]^N. \tag{2.8}$$

We turn our attention to minimizing $\mathrm{tr}[\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top]$ in $\mathbf{B}$. Consider that

$$\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top = \mathbf{U}\mathbf{D}\mathbf{V}^\top\mathbf{\Psi}\mathbf{\Omega}^{-1}\mathbf{\Psi}^\top\mathbf{V}\mathbf{D}\mathbf{U}^\top, \tag{2.9}$$

where $\mathbf{A} = \mathbf{\Psi}\mathbf{\Omega}\mathbf{\Psi}^\top$ is an eigendecomposition and $\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ is the usual singular value decomposition. We then have

$$\mathrm{tr}[\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top] = \mathrm{tr}[\mathbf{D}\mathbf{V}^\top\mathbf{\Psi}\mathbf{\Omega}^{-1}\mathbf{\Psi}^\top\mathbf{V}\mathbf{D}] = \mathrm{tr}[\mathbf{D}^2\mathbf{P}^\top\mathbf{\Omega}^{-1}\mathbf{P}] \tag{2.10}$$

after defining $\mathbf{V}^\top\mathbf{\Psi} = \mathbf{P}^\top$. This expression reduces to

$$\mathrm{tr}[\mathbf{D}^2\mathbf{P}^\top\mathbf{\Omega}^{-1}\mathbf{P}] = \sum_{ij} \frac{p_{ij}^2 d_j^2}{\omega_i}, \tag{2.11}$$

where p_{ij} are the elements of matrix $\mathbf{P}$. Since $\mathbf{P}$ is the product of two unitary matrices, it is a unitary matrix itself so that

$$\sum_i p_{ij}^2 = \sum_j p_{ij}^2 = 1. \tag{2.12}$$

So, the problem of minimizing $\operatorname{tr}[\mathbf{BA}^{-1}\mathbf{B}^\top]$ over all $\mathbf{B}$ becomes

$$\min_{\mathbf{B}} \operatorname{tr}\left[\mathbf{D}^2\mathbf{P}^\top\mathbf{\Omega}^{-1}\mathbf{P}\right] = \min_{\{p_{ij}\},\{d_i\}} \sum_{ij} \frac{p_{ij}^2 d_j^2}{\omega_i}. \tag{2.13}$$

From (2.12), p_{ij}^2, $j = 1,2,\ldots,N$, as well as p_{ij}^2, $i = 1,2,\ldots,N$, can be considered as probability mass functions. Therefore, we can rewrite the optimization problem as

$$\min_{\{p_{ij}\},\{d_i\}} \sum_i \frac{E_i[d^2]}{\omega_i}, \tag{2.14}$$

where

$$E_i[d^2] = \sum_j p_{ij}^2 d_j^2. \tag{2.15}$$

Assume for now that the d_j^2 and the ω_i are fixed and ordered from the smallest to the largest. We first note that

$$\sum_i E_i[d^2] = \sum_j d_j^2 \tag{2.16}$$

and

$$d_1^2 \leq E_i[\cdot] \leq d_N^2. \tag{2.17}$$

For the moment, imagine that one could choose any $E_i[\cdot]$ subject to the constraints of (2.16) and (2.17). That is, the underlying structure of the $E_i[\cdot]$ as generated from squared entries of a unitary matrix is ignored for the time being. We then note that if $E_N[\cdot] \neq d_N^2$, we can always reduce $\operatorname{tr}[\mathbf{D}^2\mathbf{P}^\top\mathbf{\Omega}^{-1}\mathbf{P}]$ by reducing one of the other $E_i[\cdot]$ which multiplies another $1/\omega_i > 1/\omega_N$ by some amount and adding this amount to $E_N[\cdot]$. If no such greater $1/\omega_i$ exists, then we can still increase $E_N[\cdot]$ to its maximum with no penalty to the objective. For completeness, please note that if $E_N[\cdot] < d_N^2$, then the sum, constraint on $E_i[\cdot]$ in (2.16) guarantees the existence of at least one $E_i[\cdot]$ from which mass can be "borrowed."

Thus, at the minimum, we will require $E_N[\cdot] = d_N^2$. However, to do so implies that $p_{NN} = 1$ which implies that $p_{Nj} = 0$, $j \neq N$, and that $p_{iN} = 0$, $i \neq N$—further implying that $E_i[\cdot] \leq d_{N-1}^2$, $i < N$. Proceeding recursively, we see that to minimize $\operatorname{tr}[\mathbf{D}^2\mathbf{P}^\top\mathbf{\Omega}^{-1}\mathbf{P}]$, we will have

$$p_{ij} = \begin{cases} 1, & i = j, \ j = 1,2,\ldots,N, \\ 0, & \text{otherwise}, \end{cases} \tag{2.18}$$

so that

$$\min_{\mathbf{B}} \mathrm{tr}\left[\mathbf{D}^2 \mathbf{P}^\top \mathbf{\Omega}^{-1} \mathbf{P}\right] = \min_{\{d_j\}} \sum_i \frac{d_i^2}{\omega_i}. \tag{2.19}$$

Equation (2.18) implies that the right singular basis set of $\mathbf{B}$ and the eigenvector matrix of $\mathbf{A}$ should be *identical*, $\mathbf{V} = \mathbf{\Psi}$ in (2.9), and that the singular values of $\mathbf{B}$ should be ordered from the smallest to the largest magnitude if the eigenvalues of $\mathbf{A}$ are ordered from the smallest to the largest. If $\mathbf{B}$ is symmetric, this condition is strengthened to $\mathbf{B}$ and $\mathbf{A}$ sharing the same eigenvector matrix with the eigenvalues of $\mathbf{B}$ being identical to the singular values.

Now, consider the optimization over the set $\{d_j\}$ of (2.19). To obtain constraints for the optimization, consider that

$$\mathrm{tr}[\mathbf{B}] = \mathrm{tr}\left[\mathbf{U}\mathbf{D}\mathbf{V}^\top\right] = \mathrm{tr}\left[\mathbf{D}\mathbf{V}^\top\mathbf{U}\right] = \mathrm{tr}[\mathbf{D}\mathbf{P}] = E_{\mathbf{B}}, \tag{2.20}$$

where $\mathbf{P}$ is a unitary matrix obtained from the multiplication of two unitary matrices $\mathbf{U}$ and $\mathbf{V}^\top$. For such a matrix, we must have

$$\sum_j p_{ij} p_{kj} = \delta_{ik} \tag{2.21}$$

and we note for later use that no entry p_{ij} can have magnitude larger than 1. Proceeding, we have

$$\mathrm{tr}[\mathbf{D}\mathbf{P}] = \sum_i d_i p_{ii} = E_{\mathbf{B}}. \tag{2.22}$$

Forming the auxiliary function for the optimization problem, we have

$$L(\mathbf{D}) = \sum_i \frac{d_i^2}{\omega_i} + \alpha \left(\sum_i p_{ii} d_i - E_{\mathbf{B}} \right). \tag{2.23}$$

The cross partials are zero and the second partials are positive since all $\omega_i > 0$. So the problem is convex with a unique minimum. At the extremal, we must have

$$\frac{d_i}{\omega_i} = -\frac{\alpha}{2} p_{ii}, \tag{2.24}$$

which implies

$$\sum_i p_{ii} d_i = E_{\mathbf{B}} = -\frac{\alpha}{2} \sum_i \omega_i p_{ii}^2 = -\frac{\alpha}{2} \rho^2, \tag{2.25}$$

where we define $\rho^2 = \sum_i \omega_i p_{ii}^2$. We then have

$$-\frac{\alpha}{2} = \frac{E_{\mathbf{B}}}{\rho^2}, \tag{2.26}$$

which further results in

$$\frac{d_i}{\omega_i} = \frac{E_{\mathbf{B}}}{\rho^2} p_{ii}. \tag{2.27}$$

Using this result, we reevaluate the objective

$$\sum_i \frac{d_i^2}{\omega_i} = \frac{E_{\mathbf{B}}}{\rho^2} \sum_i d_i p_{ii} = \frac{E_{\mathbf{B}}^2}{\rho^2} \tag{2.28}$$

and we next seek to maximize ρ^2.

Now, consider that since $|p_{ij}| \leq 1$, we must have

$$\sum_i \omega_i p_{ii}^2 \leq \sum_i \omega_i = E_{\mathbf{A}} \tag{2.29}$$

with equality if and only if $|p_{ii}| = 1$. Furthermore, since p_{ij} are obtained from the product of two unitary matrices, setting all $|p_{ii}| = 1$ implies that all $p_{ii} = 1$ and thence that $\mathbf{V}^\top \mathbf{U} = \mathbf{I}$. Therefore, the optimum $\mathbf{B}$ should be a *symmetric matrix* and the singular values d_i are therefore the eigenvalues of $\mathbf{B}$. The set of d_i should satisfy

$$d_i = \frac{E_{\mathbf{B}}}{E_{\mathbf{A}}} \omega_i, \tag{2.30}$$

which implies that $\mathbf{B}$ is a scaled version of $\mathbf{A}$,

$$\mathbf{B} = \frac{E_{\mathbf{B}}}{E_{\mathbf{A}}} \mathbf{A}, \tag{2.31}$$

since we have already determined that $\mathbf{V} = \mathbf{\Psi}$. The minimum trace is then

$$\min_{\mathbf{B}} \text{tr}\left[\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top\right] = \text{tr}\left[\frac{E_{\mathbf{B}}^2}{E_{\mathbf{A}}^2}\mathbf{A}\right] = \frac{E_{\mathbf{B}}^2}{E_{\mathbf{A}}}. \tag{2.32}$$

We summarize the result as a theorem.

THEOREM 2.3. *Given two square $N \times N$ matrices, $\mathbf{A} \in \mathcal{M}^+$ with $\text{tr}[\mathbf{A}] = E_A$, and $\mathbf{B}$ with $\text{tr}[\mathbf{B}] = E_B$, the minimum of $\text{tr}[\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top]$ in $\mathbf{B}$ is achieved when $\mathbf{B}$ is a scaled version of $\mathbf{A}$, $\mathbf{B} = (E_B/E_A)\mathbf{A}$, and*

$$\min_{\mathbf{B}} \text{tr}\left[\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^\top\right] = \frac{E_{\mathbf{B}}^2}{E_{\mathbf{A}}}. \tag{2.33}$$

Returning to the original determinant maximization problem, we now have

$$\max_{\mathbf{Q} \in \mathcal{D}_{N,2}} |\mathbf{Q}| = \max_{\mathbf{A},\mathbf{C} \in \mathcal{M}_N^+} \left\{ |\mathbf{A}| \left| \mathbf{C} - \frac{E_{\mathbf{B}}^2}{E_{\mathbf{A}}^2}\mathbf{A} \right| \right\} \tag{2.34}$$

or

$$\max_{\mathbf{Q} \in \mathcal{D}_{N,2}} |\mathbf{Q}| = \max_{\mathbf{A} \in \mathcal{M}_N^+} \left\{ |\mathbf{A}| \max_{\mathbf{C} \in \mathcal{M}_N^+} \left| \mathbf{C} - \frac{E_{\mathbf{B}}^2}{E_{\mathbf{A}}^2}\mathbf{A} \right| \right\}. \tag{2.35}$$

We note that $\mathbf{C} - (E_{\mathbf{B}}^2/E_{\mathbf{A}}^2)\mathbf{A} \in \mathcal{M}^+$, which implies that $\mathbf{C} = (E_{\mathbf{B}}^2/E_{\mathbf{A}}^2)\mathbf{A} + \mathbf{X}$ where $\mathbf{X} \in \mathcal{M}^+$. Therefore, we can always choose the optimum $\mathbf{C}$ such that $\mathbf{X} = \mathbf{C} - (E_{\mathbf{B}}^2/E_{\mathbf{A}}^2)\mathbf{A}$ is a scaled identity matrix. This implies that the inner maximization does not depend on the shape of $\mathbf{A}$ and that we can choose $\mathbf{A}$ as a scaled identity matrix to maximize $|\mathbf{A}|$ in the outer maximization. This in turn makes $\mathbf{B}$ and the optimal $\mathbf{C}$ scaled identity matrices as well.

So,

$$\max_{\mathbf{Q} \in \mathcal{D}_{N,2}} |\mathbf{Q}| = \frac{1}{N^{2N}} (E_{\mathbf{A}} E_{\mathbf{C}} - E_{\mathbf{B}}^2)^N \tag{2.36}$$

and this maximum can be achieved when $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are all scaled identity matrices. We summarize this result as a theorem.

THEOREM 2.4. *Let*

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \tag{2.37}$$

be a positive definite matrix and let $\mathbf{B}$ *be an* $N \times N$ *matrix. Under the constraints* $\mathrm{tr}[\mathbf{A}] = E_{\mathbf{A}}$, $\mathrm{tr}[\mathbf{C}] = E_{\mathbf{C}}$, *and* $\mathrm{tr}[\mathbf{B}] = E_{\mathbf{B}}$,

$$\begin{vmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{vmatrix} \leq \begin{vmatrix} \dfrac{E_{\mathbf{A}}}{N}\mathbf{I}_N & \dfrac{E_{\mathbf{B}}}{N}\mathbf{I}_N \\ \dfrac{E_{\mathbf{B}}}{N}\mathbf{I}_N & \dfrac{E_{\mathbf{C}}}{N}\mathbf{I}_N \end{vmatrix} = \frac{1}{N^{2N}} (E_{\mathbf{A}} E_{\mathbf{C}} - E_{\mathbf{B}}^2)^N \tag{2.38}$$

with equality when the matrices $\mathbf{A}$, $\mathbf{C}$, *and* $\mathbf{B}$ *are all scaled identity matrices. If*

$$\mathbf{E} = \begin{bmatrix} E_{\mathbf{A}} & E_{\mathbf{B}} \\ E_{\mathbf{B}} & E_{\mathbf{C}} \end{bmatrix}, \tag{2.39}$$

then

$$\max \begin{vmatrix} \mathbf{A} & \mathbf{B}^\top \\ \mathbf{B} & \mathbf{C} \end{vmatrix} = \frac{1}{N^{2N}} |\mathbf{E}|^N. \tag{2.40}$$

2.2. The recursive extension for $J > 2$. Let $\mathbf{Q}$ be as defined in Theorem 1.2. Let the trace constraints of the subblocks be $\mathrm{tr}[\mathbf{Q}_{ij}] = E_{ij}$ and let $\mathbf{E}$ be the $J \times J$ matrix containing these traces. Following an inductive line of reasoning, we assume that

$$\max |\mathbf{Q}| = \frac{1}{N^{NJ}} |\mathbf{E}|^N \tag{2.41}$$

when all the submatrices of $\mathbf{Q}$ are scaled identity matrices. Now we seek to prove the result for $J + 1$.

So, define $\mathbf{X} \in \mathcal{D}_{N,(J+1)}$ as

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B}_1^\top & \cdots & \mathbf{B}_J^\top \\ \mathbf{B}_1 & & & \\ \vdots & & \mathbf{Q} & \\ \mathbf{B}_J & & & \end{bmatrix}. \tag{2.42}$$

We then have

$$|\mathbf{X}| = |\mathbf{A}| \left| \mathbf{Q} - \begin{bmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_J \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} \mathbf{B}_1^\top & \cdots & \mathbf{B}_J^\top \end{bmatrix} \right|, \tag{2.43}$$

which we rewrite as

$$|\mathbf{X}| = |\mathbf{A}| |\mathbf{Q} - \mathbf{Z}|, \tag{2.44}$$

where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_J \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} \mathbf{B}_1^\top & \cdots & \mathbf{B}_J^\top \end{bmatrix}. \tag{2.45}$$

Now, we know that $\mathbf{Z} \in \mathcal{M}^{0+}$ by Lemma 2.1, and we also know that it has rank at most N. We may therefore rewrite it as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{u} & \bar{\mathbf{u}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{0} & \\ 0 & & & \end{bmatrix} \begin{bmatrix} \mathbf{u}^\top \\ \bar{\mathbf{u}}^\top \end{bmatrix}, \tag{2.46}$$

where $\mathbf{u}$ is an orthonormal basis which spans the column space of

$$\begin{bmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_J \end{bmatrix} \tag{2.47}$$

and $\bar{\mathbf{u}}$ is its complement such that

$$\begin{bmatrix} \mathbf{u}^\top \\ \bar{\mathbf{u}}^\top \end{bmatrix} \begin{bmatrix} \mathbf{u} & \bar{\mathbf{u}} \end{bmatrix} = \mathbf{I} \tag{2.48}$$

and $\boldsymbol{\Omega} \in \mathcal{M}^{0+}$ is an $N \times N$ diagonal matrix with $\mathrm{tr}[\mathbf{Z}] = \mathrm{tr}[\boldsymbol{\Omega}]$.

Now, consider that

$$|\mathbf{Q} - \mathbf{Z}| = \left| \begin{bmatrix} \mathbf{u}^\top \\ \bar{\mathbf{u}}^\top \end{bmatrix} (\mathbf{Q} - \mathbf{Z}) \begin{bmatrix} \mathbf{u} & \bar{\mathbf{u}} \end{bmatrix} \right| = \begin{vmatrix} \mathbf{u}^\top \mathbf{Q} \mathbf{u} - \boldsymbol{\Omega} & \mathbf{u}^\top \mathbf{Q} \bar{\mathbf{u}} \\ \bar{\mathbf{u}}^\top \mathbf{Q} \mathbf{u} & \bar{\mathbf{u}}^\top \mathbf{Q} \bar{\mathbf{u}} \end{vmatrix}, \tag{2.49}$$

which we can reduce to

$$|\mathbf{Q} - \mathbf{Z}| = \left| \bar{\mathbf{u}}^\top \mathbf{Q} \bar{\mathbf{u}} \right| |\mathbf{H} - \boldsymbol{\Omega}|, \tag{2.50}$$

where

$$\mathbf{H} = \mathbf{u}^\top \mathbf{Q} \mathbf{u} - \mathbf{u}^\top \mathbf{Q} \bar{\mathbf{u}} (\bar{\mathbf{u}}^\top \mathbf{Q} \bar{\mathbf{u}})^{-1} \bar{\mathbf{u}}^\top \mathbf{Q} \mathbf{u} \tag{2.51}$$

and $\mathbf{H} \in \mathcal{M}^+$.

We can write $\mathbf{\Omega} = \alpha\mathbf{\Theta}$, where $\alpha \geq 0$ and $\mathbf{\Theta} \in \mathcal{M}^{0+}$ has unit trace. Then we have

$$|\mathbf{H} - \mathbf{\Omega}| = |\mathbf{H}|\,\big|\mathbf{I} - \alpha\mathbf{H}^{-1/2}\mathbf{\Theta}\mathbf{H}^{-1/2}\big| = |\mathbf{H}|\,|\mathbf{I} - \alpha\mathbf{\Lambda}|, \qquad (2.52)$$

where

$$\mathbf{H}^{-1/2}\mathbf{\Theta}\mathbf{H}^{-1/2} = \mathbf{\Psi}\mathbf{\Lambda}\mathbf{\Psi}^{\top}. \qquad (2.53)$$

Now, for any given $\mathbf{\Lambda}$ (and therefore any given $\mathbf{H}$ and $\mathbf{\Theta}$), reducing α will *strictly* increase the determinant in (2.52) since $(\mathbf{I} - \alpha\mathbf{\Lambda}) \in \mathcal{M}^{+}$. Thus, for $|\mathbf{X}|$ in (2.42) to be maximized, we must have α minimized.

Proceeding, we have $\alpha = \text{tr}[\mathbf{\Omega}] = \text{tr}[\mathbf{Z}]$. Expanding $\text{tr}[\mathbf{Z}]$, we have

$$\text{tr}[\mathbf{Z}] = \sum_i \text{tr}\big[\mathbf{B}_i \mathbf{A}^{-1} \mathbf{B}_i^{\top}\big], \qquad (2.54)$$

and from Theorem 2.3, we know that $\text{tr}[\mathbf{Z}]$ may be minimized when each $\mathbf{B}_i$ is a scaled version of $\mathbf{A}$. Defining $\text{tr}[\mathbf{B}_i] = \beta_i$, at this minimum, we have

$$\mathbf{Z} = \begin{bmatrix} \beta_1 \mathbf{I} \\ \vdots \\ \beta_J \mathbf{I} \end{bmatrix} \mathbf{A} \begin{bmatrix} \beta_1 \mathbf{I} & \cdots & \beta_J \mathbf{I} \end{bmatrix} \qquad (2.55)$$

so that

$$\mathbf{\Omega} = \left(\frac{\sum_i \beta_i^2}{\text{tr}[\mathbf{A}]}\right)\left(\frac{\mathbf{A}}{\text{tr}[\mathbf{A}]}\right) = \alpha^*\mathbf{\Theta}, \qquad (2.56)$$

where

$$\alpha^* = \frac{\sum_i \beta_i^2}{\text{tr}[\mathbf{A}]} \qquad (2.57)$$

is the minimum possible trace of $\mathbf{Z}$ and

$$\mathbf{\Theta} = \frac{\mathbf{A}}{\text{tr}[\mathbf{A}]}. \qquad (2.58)$$

If $\mathbf{\Theta} \in \mathcal{M}^{+}$ and $\alpha > \alpha^*$, then we can choose $\mathbf{A} = \text{tr}[\mathbf{A}]\mathbf{\Theta}$ and increase $|\mathbf{H} - \alpha\mathbf{\Theta}|$ by replacing α with α^*. Thus, for all $\mathbf{\Theta} \in \mathcal{M}^{+}$, we must have minimum trace $\mathbf{Z}$ if $|\mathbf{Q} - \mathbf{Z}|$ is maximized. However, if $\mathbf{\Theta} \in \mathcal{M}^{0+}$ and $\alpha > \alpha^*$, then we cannot choose $\mathbf{A}$ directly since $\mathbf{A} \in \mathcal{M}^{+}$. In this case, we can form

$$\left(\frac{\sum_i \beta_i^2}{\text{tr}[\mathbf{A}]}\right)\left(\frac{\mathbf{A}}{\text{tr}[\mathbf{A}]}\right) = \alpha^*\mathbf{\Theta} + \epsilon\mathbf{F}, \qquad (2.59)$$

where $\mathbf{F}$ is a symmetric matrix with $\text{tr}[\mathbf{F}] = 0$ which makes $\alpha\mathbf{\Theta} + \epsilon\mathbf{F} \in \mathcal{M}^{+}$ for $0 < \epsilon < \epsilon'$. We now show that replacement of $\alpha\mathbf{\Theta}$ by $\alpha^*\mathbf{\Theta} + \epsilon\mathbf{F}$ can always increase $|\mathbf{H} - \alpha\mathbf{\Theta}|$ for suitably small ϵ.

First we expand

$$|\mathbf{H} - \alpha\boldsymbol{\Theta} - \epsilon\mathbf{F}| = |\mathbf{H} - \alpha\boldsymbol{\Theta}|\,|\mathbf{I} - \epsilon(\mathbf{H} - \alpha\boldsymbol{\Theta})^{-1/2}\mathbf{F}(\mathbf{H} - \alpha\boldsymbol{\Theta})^{-1/2}|$$

$$= |\mathbf{H} - \alpha\boldsymbol{\Theta}|\prod_{i=1}^{N}(1 - \epsilon\nu_i), \tag{2.60}$$

where the $\{\nu_i\}$ are the eigenvalues of $\epsilon(\mathbf{H} - \alpha\boldsymbol{\Theta})^{-1/2}\mathbf{F}(\mathbf{H} - \alpha\boldsymbol{\Theta})^{-1/2}$. Since $\prod_{i=1}^{N}(1 - \epsilon\nu_i)$ is a continuous function in ϵ and

$$\prod_{i=1}^{N}(1 - \epsilon\nu_i)|_{\epsilon=0} = 1, \tag{2.61}$$

$\prod_{i=1}^{N}(1 - \epsilon\nu_i)$ can be made arbitrarily close to 1. That is,

$$|\mathbf{H} - \alpha\boldsymbol{\Theta}| - \delta \leq |\mathbf{H} - \alpha\boldsymbol{\Theta} - \epsilon\mathbf{F}| \leq |\mathbf{H} - \alpha\boldsymbol{\Theta}| + \delta \tag{2.62}$$

for any $\delta > 0$ through suitable choice of ϵ.

Now, since reduction of α *strictly* increases $|\mathbf{H} - \alpha\boldsymbol{\Theta}|$, we have

$$|\mathbf{H} - \alpha\boldsymbol{\Theta}| < |\mathbf{H} - \alpha^*\boldsymbol{\Theta}|, \tag{2.63}$$

and for suitably small $\delta > 0$, we also have

$$|\mathbf{H} - \alpha^*\boldsymbol{\Theta}| - \delta \leq |\mathbf{H} - \alpha^*\boldsymbol{\Theta} - \epsilon\mathbf{F}| \leq |\mathbf{H} - \alpha^*\boldsymbol{\Theta}| + \delta. \tag{2.64}$$

So $|\mathbf{H} - \alpha\boldsymbol{\Theta}|$ can always be strictly increased if $\alpha > \alpha^*$.

Therefore, if there exist some $\boldsymbol{\Theta}^*$ and $\mathbf{Q}^*$ which ostensibly minimize $|\mathbf{Q} - \mathbf{Z}|$, then unless $\mathrm{tr}[\mathbf{Z}]$ is minimized, this "maximum" could be further increased—a contradiction. Therefore, for $|\mathbf{Q} - \mathbf{Z}|$ to be maximized, $\alpha = \mathrm{tr}[\mathbf{Z}]$ must be minimized.

With $\mathrm{tr}[\mathbf{Z}]$ minimized, we have

$$\mathbf{Z}_{ij} = \frac{\beta_i\beta_j}{(\mathrm{tr}[\mathbf{A}])^2}\mathbf{A}, \qquad \mathrm{tr}[\mathbf{Z}_{ij}] = \frac{\beta_i\beta_j}{\mathrm{tr}[\mathbf{A}]}. \tag{2.65}$$

Since $(\mathbf{Q} - \mathbf{Z}) \in \mathcal{D}_{N,J}$, we obtain by the assumption for J that

$$|\mathbf{Q} - \mathbf{Z}| \leq \frac{1}{N^{NJ}}\begin{vmatrix} E_{11} - \dfrac{\beta_1^2}{\mathrm{tr}[\mathbf{A}]} & E_{12} - \dfrac{\beta_1\beta_2}{\mathrm{tr}[\mathbf{A}]} & \cdots & E_{1J} - \dfrac{\beta_1\beta_J}{\mathrm{tr}[\mathbf{A}]} \\ E_{21} - \dfrac{\beta_1\beta_2}{\mathrm{tr}[\mathbf{A}]} & E_{22} - \dfrac{\beta_2^2}{\mathrm{tr}[\mathbf{A}]} & & \vdots \\ \vdots & & \ddots & \vdots \\ E_{J1} - \dfrac{\beta_1\beta_J}{\mathrm{tr}[\mathbf{A}]} & \cdots & \cdots & E_{JJ} - \dfrac{\beta_J^2}{\mathrm{tr}[\mathbf{A}]} \end{vmatrix}. \tag{2.66}$$

From Theorem 2.3, we note that at the minimum trace point, $\text{tr}[\mathbf{Z}_{ij}]$ does not depend on the spectral shape of $\mathbf{A}$, but on its trace only. Therefore, once we have chosen $\mathbf{B}_i$ to minimize $\text{tr}[\mathbf{Z}]$ and thereby allow the maximum value of $|\mathbf{Q} - \mathbf{Z}|$, we are free to choose the shape of $\mathbf{A}$. Since choosing $\mathbf{A} = (\text{tr}[\mathbf{A}]/N)\mathbf{I}$ maximizes $|\mathbf{A}|$, we then have

$$\max |\mathbf{X}| \le \frac{\text{tr}[\mathbf{A}]^N}{N^N} \frac{1}{N^{NJ}} \begin{vmatrix} E_{11} - \dfrac{\beta_1^2}{\text{tr}[\mathbf{A}]} & E_{12} - \dfrac{\beta_1\beta_2}{\text{tr}[\mathbf{A}]} & \cdots & E_{1J} - \dfrac{\beta_1\beta_J}{\text{tr}[\mathbf{A}]} \\ E_{21} - \dfrac{\beta_1\beta_2}{\text{tr}[\mathbf{A}]} & E_{22} - \dfrac{\beta_2^2}{\text{tr}[\mathbf{A}]} & & \vdots \\ \vdots & & \ddots & \vdots \\ E_{J1} - \dfrac{\beta_1\beta_J}{\text{tr}[\mathbf{A}]} & \cdots & \cdots & E_{JJ} - \dfrac{\beta_J^2}{\text{tr}[\mathbf{A}]} \end{vmatrix}^N \tag{2.67}$$

with equality when all the submatrices of $\mathbf{X}$ are scaled identity matrices.

If we note that

$$\begin{vmatrix} \text{tr}[\mathbf{A}] & \beta_1 & \beta_2 & \cdots & \beta_J \\ \beta_1 & E_{11} & E_{12} & \cdots & E_{1J} \\ \beta_2 & E_{21} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \beta_J & E_{J1} & \cdots & \cdots & E_{JJ} \end{vmatrix}$$

$$= \text{tr}[\mathbf{A}] \left| \begin{bmatrix} E_{11} & E_{12} & \cdots & E_{1J} \\ E_{21} & E_{22} & & E_{2J} \\ \vdots & & \ddots & \vdots \\ E_{J1} & E_{J2} & \cdots & E_{JJ} \end{bmatrix} - \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_J \end{bmatrix} \frac{1}{\text{tr}[\mathbf{A}]} \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_J \end{bmatrix} \right| \tag{2.68}$$

$$= \text{tr}[\mathbf{A}] \begin{vmatrix} E_{11} - \dfrac{\beta_1^2}{\text{tr}[\mathbf{A}]} & E_{12} - \dfrac{\beta_1\beta_2}{\text{tr}[\mathbf{A}]} & \cdots & E_{1J} - \dfrac{\beta_1\beta_J}{\text{tr}[\mathbf{A}]} \\ E_{21} - \dfrac{\beta_1\beta_2}{\text{tr}[\mathbf{A}]} & E_{22} - \dfrac{\beta_2^2}{\text{tr}[\mathbf{A}]} & & \vdots \\ \vdots & & \ddots & \vdots \\ E_{J1} - \dfrac{\beta_1\beta_J}{\text{tr}[\mathbf{A}]} & \cdots & \cdots & E_{JJ} - \dfrac{\beta_J^2}{\text{tr}[\mathbf{A}]} \end{vmatrix},$$

then, when all blocks in $\mathbf{X}$ are scaled identity matrices, we have

$$|\mathbf{X}| = \frac{1}{N^{N(J+1)}} \begin{vmatrix} \text{tr}[\mathbf{A}] & \beta_1 & \cdots & \cdots & \beta_J \\ \beta_1 & E_{11} & E_{12} & \cdots & E_{1J} \\ \beta_2 & E_{21} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \beta_J & E_{J1} & \cdots & \cdots & E_{JJ} \end{vmatrix}^N, \tag{2.69}$$

which proves the $(J+1)$ case, and through induction completes the proof.

3. Application to maximum determinant of a sum of Kronecker products

Given two matrices $\mathbf{A} \in \mathbb{R}^{B \times B}$ fixed and $\mathbf{X} \in \mathbb{R}^{N \times N}$ with constant trace $\text{tr}[\mathbf{X}] = \alpha$, the determinant of their Kronecker product $\mathbf{R} = \mathbf{A} \otimes \mathbf{X}$ is given by

$$|\mathbf{R}| = |\mathbf{A}|^N |\mathbf{X}|^B. \tag{3.1}$$

Since $|\mathbf{A}|$ is assumed constant, in order to maximize $|\mathbf{R}|$, we need to maximize $|\mathbf{X}|$, and under the trace constraint, this is done when $\mathbf{X} = (\alpha/N)\mathbf{I}_N$ (Lemma 2.2).

Considering now a *sum* of Kronecker products,

$$\mathbf{R} = \sum_{k=1}^{L} \mathbf{A}_k \otimes \mathbf{X}_k = [\mathbf{R}_{ij}], \quad \mathbf{R}_{ij} = \sum_{k=1}^{L} a_k(i,j)\mathbf{X}_k, \ i,j = 1,\dots,L, \tag{3.2}$$

where we assume that $\mathbf{A}_k = [a_k(i,j)]$ are $B \times B$ symmetric constant matrices and $\mathbf{X}_k$ are $N \times N$ symmetric matrix with constant trace, $\text{tr}[\mathbf{X}_k] = \alpha_k$. To satisfy the conditions of Theorem 1.2, we require that $\mathbf{R}$ be a positive definite matrix and that all blocks on the main diagonal, $\mathbf{R}_{ii}$, be also positive definite matrices. Such matrices appear in the context of capacity maximization for multibase communication systems [3]. Thus, the matrix $\mathbf{R}$ belongs to the class $\mathcal{D}_{N,B}$ and the associated matrix of traces $\mathbf{E}$ has elements

$$E_{ij} = \sum_{k=1}^{L} a_k(i,j)\alpha_k. \tag{3.3}$$

As a direct application of Theorem 1.2 to matrix $\mathbf{R}$, we have the following corollary.

COROLLARY 3.1. *The determinant of a sum of Kronecker products, as described above, is maximized when the whole sum is equal to a single Kronecker product between the trace matrix and an identity matrix. Thus, maximum* $\det(\mathbf{R})$ *is achieved when*

$$\mathbf{R} = \sum_{k=1}^{L} \mathbf{A}_k \otimes \mathbf{X}_k = \mathbf{E} \otimes \mathbf{I}_N. \tag{3.4}$$

If we are interested in finding the set of matrices $\{\mathbf{X}_k\}$ that maximize $\det(\mathbf{R})$, we note that (3.4) is equivalent (because of its symmetry) to a linear system with $B(B+1)/2$ equations and L unknowns:

$$\sum_{k=1}^{L} a_k(i,j)\mathbf{X}_k = E_{ij}, \quad i = 1,\dots,B, \ j \geq i. \tag{3.5}$$

Depending on the relative magnitudes of B and L, this system might have one or more solutions, but one of these solutions is always $\mathbf{X}_k = (\alpha_k/N)\mathbf{I}_N$.

References

[1] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[2] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Mathematics in Science and Engineering, vol. 143, Academic Press, New York, 1979.

[3] O. Popescu and C. Rose, *Interference avoidance and sum capacity for multibase systems*, Proceedings 39th Allerton Conference on Communication, Control, and Computing (Monticello, Ill), University of Illinois at Urbana-Champaign, Illinois, 2001, pp. 1036–1049.

[4] L. Vandenberghe, S. Boyd, and S.-P. Wu, *Determinant maximization with linear matrix inequality constraints*, SIAM J. Matrix Anal. Appl. **19** (1998), no. 2, 499–533.

[5] H. S. Witsenhausen, *A determinant maximization problem occurring in the theory of data communication*, SIAM J. Appl. Math. **29** (1975), no. 3, 515–522.

Otilia Popescu: WINLAB, Rutgers University, 73 Brett Road, Piscataway, NJ 08854-8058, USA
E-mail address: otilia@winlab.rutgers.edu

Christopher Rose: WINLAB, Rutgers University, 73 Brett Road, Piscataway, NJ 08854-8058, USA
E-mail address: crose@winlab.rutgers.edu

Dimitrie C. Popescu: Department of Electrical Engineering, University of Texas at San Antonio, 6900 North Loop 1604 West, San Antonio, TX 78249-0669, USA
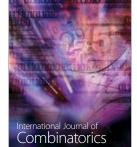*E-mail address*: dpopescu@utsa.edu