Correlation between orthogonally projected matrices

Catherine Fraikin, Yurii Nesterov and Paul Van Dooren
Université catholique de Louvain
Department of Mathematical Engineering
B-1348 Louvain-la-Neuve, Belgium
fraikin@inma.ucl.ac.be; nesterov@inma.ucl.ac.be; vdooren@inma.ucl.ac.be

Abstract— **We consider the problem of finding the optimal correlation between two projected matrices U^*AU and V^*BV. The square matrices A and B may be of different dimensions, but the isometries U and V have a common column dimension k. The correlation is measured by the real function $c(U,V) = \Re \operatorname{tr}(U^*AUV^*B^*V)$, which we maximize of the isometries $U^*U = V^*V = I_k$.**

This problem can be viewed as an extension of the generalized numerical range of two matrices, which are now allowed to be of different dimension. We discuss several properties of this optimization problem, characterize its extremal points and propose an algorithm converging to such an extremal point.

Keywords— **Correlation, Trace maximization, Generalized numerical range, Isometry**

I. Introduction

The problem of projection of matrices in lower-dimensional subspaces is of great interest for a large range of applications. The projection of matrices provides an easier visualization and comprehension of the initial problem and is often used to reduce its complexity. Moreover the correlation between these projections can reveal some particularities inherent to the data which can then be analyzed and interpreted.

We consider the correlation between two projected matrices U^*AU and V^*BV, where A and B are respectively of dimensions $m \times m$ and $n \times n$. The isometries U and V have a common column dimension k and satisfy thus the constraint $U^*U = V^*V = I_k$, where I_k denotes the identity matrix of dimension $k \leq \min(m,n)$. The correlation is expressed as the real function

$$c(U,V) = \Re \operatorname{tr}(U^*AUV^*B^*V). \tag{1}$$

This is a generic problem which can be linked to various applications treated in the literature and which has been studied extensively in a variety of contexts for particular dimensions of the projection and of the matrices. A first field of application lies in the analysis of graphs. The notion of graph similarity matrix, which is a matrix that expresses how similar the nodes of two graphs are, has recently been introduced in [1]. For undirected graphs, this

similarity matrix is the correlation of the graph adjacency matrices projected in a one-dimensional subspace. The graph similarity matrix is e.g. useful for the development of efficient Web search engines or the automatic extraction of synonyms in a dictionary. Another important task in graph analysis is that of graph matching which is a fundamental problem in pattern recognition and in shape and image analysis. A popular class of methods in graph matching are the spectral methods in which spectral properties of characteristic matrices are used to compare the graphs. The spectral method developed in [2] combines a projection technique and a clustering algorithm to match the graphs in a lower-dimensional subspace. It can be shown that the step of projection used by the authors is equivalent to maximize (1) for symmetric matrices A and B. A second field of applications concerns experiments in quantum mechanics and in particular the task of maximizing the signal intensity in coherent spectroscopy (see e.g. [3] and [4]). From a mathematical point of view, this problem is equivalent to optimizing an expression similar to (1) where all the matrices are square.

In the linear algebra literature, problem (1) has also been largely studied for particular cases and dimensions and hence constitutes an extension of a variety of known problems. For the case where all the matrices are square, this problem corresponds to the generalized numerical range (or B-numerical range). See e.g. [5] and [6] for a survey on the properties of the generalized numerical range. For the scalar case which corresponds to a one-dimensional projection, the expression (1) is equivalent to the product of the field of values of two matrices (see e.g. [7]). In this paper we extend these existing notions to a more general case of dimensions of matrices. We treat also the complex and real cases.

From a numerical point of view, many algorithms exist to maximize (1) for particular dimensions of the matrices (e.g. [3] and [8]). We develop here a simple recursive algorithm valid for the general case, i.e. for complex or real problems and for all dimensions of the matrices. Characterizations of the fixed points of the function and of the algorithm are presented.

The paper is organized as follows. In Section 2, we introduce some notations. In Section 3, we define the problem considered in the paper which consists of maximizing an

expression similar to (1) for a general case of dimensions of the matrices. We recall some important results from the literature that we can link to our problem. The first one concerns of square matrices and appears in the field of the generalized numerical range and in the context of semi-definite programming relaxations. The second case is about scalar projections and is linked to the field of values of matrices. The main new results are in Section 4 where we characterize the fixed points of the problem. Then we focus on the case of Hermitian and normal matrices and we give lower and upper bounds for the optimal value. In Section 5, we propose a simple algorithm to solve the optimization problem. The last Section 6 summarizes the results and describes some directions for future research.

II. NOTATIONS

In this section, we introduce some notations used in the paper. The first part treats of the complex and Hermitian inner product of matrices, while the second part summarizes some definitions about derivatives of functions with matrix arguments.

A. Inner product

Let $\mathbb{R}$ and $\mathbb{C}$ denote the real and complex field, respectively. $\mathbb{R}^{m \times n}$ and $\mathbb{C}^{m \times n}$ denote the set of all $m \times n$ real and complex matrices. X^T, $\bar{X}$ and X^* represent respectively the transpose, the complex conjugate and the complex conjugate transpose of X. The inner product between matrices is defined as follows. For $X, Y \in \mathbb{R}^{m \times n}$, the *real inner product* is denoted by

$$\langle X, Y \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij}$$

and can be linked to the trace function of a matrix:

$$\langle X, Y \rangle = \mathrm{tr}(XY^T) = \mathrm{tr}(X^T Y).$$

For complex matrices $X, Y \in \mathbb{C}^{m \times n}$, the *complex inner product* is defined by

$$\langle X, Y \rangle_C = \sum_{i=1}^{m} \sum_{j=1}^{n} \bar{X}_{ij} Y_{ij}$$

and can be linked again to the trace

$$\langle X, Y \rangle_C = \mathrm{tr}(XY^*).$$

We introduce also the *Hermitian inner product* as the real part of the complex inner product:

$$\langle X, Y \rangle_H = \Re(\langle X, Y \rangle_C) = \langle \Re(X), \Re(Y) \rangle + \langle \Im(X), \Im(Y) \rangle.$$

B. Derivatives

Let $f(X) : \mathbb{C}^{m \times n} \to \mathbb{R}$ be a differentiable real-valued function with matrix argument X. Then the first-order approximation of f at a point X can be expressed as

$$f(X + \Delta) = f(X) + \langle \nabla f(X), \Delta \rangle_H + o(\|\Delta\|) \tag{2}$$

where the derivative $\nabla f(X)$ is the $m \times n$ matrix whose (i, j) entry is $\frac{\partial f(X)}{\partial X_{i,j}}$. To differentiate an expression $f(X)$ with respect to a complex variable X, the more general complex derivative called *complex Gradient Matrix* for a real function f of a complex matrix X is used. This gradient ∇f is defined by (2) and allows us to estimate some derivatives of inner products we need in the paper in order to calculate gradients.

III. MAIN KNOWN RESULTS

For $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$, we consider the following problem

$$\max_{\substack{U^*U = I_k \\ V^*V = I_k}} \langle U^*AU, V^*BV \rangle_H \tag{3}$$

where $U \in \mathbb{C}^{m \times k}$ and $V \in \mathbb{C}^{n \times k}$ with $k \leq \min(m, n)$. For real matrices A and B, U and V must also be real. This problem has largely been studied for particular dimensions of m, n and k. The first section contains results existing for $m = n = k$ and the second one for $k = 1$.

A. Square matrices U and V

For $m = n = k$, U and V are square matrices and the problem we consider is equivalent to:

$$\max_{Q^*Q = I} \langle Q^*AQ, B \rangle_H \tag{4}$$

where the optimization depends only on the unitary matrix Q which corresponds to the product UV^*. This problem has been studied in a variety of contexts. In the rest of the section, we connect our problem to the *B*-numerical range and present results in the field of semi-definite programming relaxations in the aim to provide bounds on the problem. To link the notations used in the literature for this problem with (4), we point out that

$$\langle Q^*AQ, B \rangle_H = \Re(\mathrm{tr}(AQB^*Q^*)) = \Re(\mathrm{tr}(B^*Q^*AQ)).$$

1) B-numerical range: The problem (4) is equivalent to maximizing the real part of the *B*-numerical range of A (or generalized numerical range) introduced by [9] and defined by

$$W_B(A) := \{\mathrm{tr}(B^*Q^*AQ) : Q \text{ is unitary}\}.$$

The *B*-numerical range has been studied by many authors in the last few decades (see e.g. [5] and [6] for a survey on the properties of the generalized numerical range) and has many domains of applications, e.g. in quantum dynamics for the study of the efficiency of polarization or coherence transfer between quantized states under unitary transformations (e.g. [3] and [4]). Some authors have also used the numerical range to study problems on norms of operators (see [5]).

2) Semidefinite programming relaxations:

a) Real matrices: In the case of real matrices A and B and by adding the redundant constraints $QQ^T = I$, the problem (4) becomes

$$v_P = \max_{\substack{Q^T Q = I \\ QQ^T = I}} \operatorname{tr}(AQB^T Q^T). \qquad (5)$$

By a reasoning similar to the one developed in [8], we construct the following dual Lagrangian:

$$v_D = \min \operatorname{tr} S + \operatorname{tr} T \qquad (6)$$
$$\text{s.t.} \quad \tfrac{B \otimes A}{2} + \tfrac{B^T \otimes A^T}{2} - S \otimes I - I \otimes T \preceq 0$$
$$S = S^T$$
$$T = T^T$$

where the two symmetric matrices of Lagrange multipliers S and T are used to relax the constraints $Q^T Q = I$ and $QQ^T = I$ and $\otimes$ denotes the Kronecker product. The redundant constraint $QQ^T = I$ is added in order to close the duality gap for symmetric matrices A and B. Indeed, it is proved that strong duality holds for symmetric matrices, i.e. $v_P = v_D$. For non-symmetric matrices, strong duality does not hold, but this method provides an upper bound v_D for the problem we consider, i.e. $v_P \leq v_D$.

b) Complex matrices: A complex matrix $A = A_R + jA_I$ of dimension $n \times n$ can be represented by a real matrix $\tilde{A}$ of dimension $2n \times 2n$ of the form:

$$\tilde{A} = \begin{pmatrix} A_R & A_I \\ -A_I & A_R \end{pmatrix}. \qquad (7)$$

For a Hermitian matrix A, $\tilde{A}$ is symmetric while for a unitary matrix Q, $\tilde{Q}$ is orthogonal. The following theorem results from this representation.

Theorem 1: Suppose that we represent the matrices A, B and $Q \in \mathbb{C}^{n \times n}$ by the matrices $\tilde{A}$, $\tilde{B}$ and $\tilde{Q} \in \mathbb{C}^{2n \times 2n}$ by use of the expression (7). Then we obtain the following link between the two trace functions:

$$2\Re \operatorname{tr}(AQB^*Q^*) = \operatorname{tr}(\tilde{A}\tilde{Q}\tilde{B}^T\tilde{Q}^T). \qquad (8)$$

By Theorem 1, the optimization problem (4) in term of complex matrices is equivalent to maximizing

$$\max_{\tilde{Q}^T \tilde{Q} = I} \frac{1}{2}\operatorname{tr}(\tilde{A}\tilde{Q}\tilde{B}\tilde{Q}^T). \qquad (9)$$

expressed in term of real matrices. The dual method developed previously for real matrices can then be applied in the same way and provides an upper bound for the problem.

B. One-dimensional matrices U and V

When k equals one, the matrices U and V are reduced to vectors u and v and the problem (3) becomes

$$\max_{\substack{u^*u = 1 \\ v^*v = 1}} \langle u^*Au, v^*Bv \rangle_H. \qquad (10)$$

This problem is related to the notion of the field of values. The field of values of a matrix A (also known as the numerical range) is defined by [7]

$$F(A) := \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1\}.$$

The problem is then reduced to obtaining the maximum of the Hermitian product of the elements from the fields of values of A and of B. The field of values is known to be a convex subset of the complex plane while the product of two fields of values $F(A)F(B)$ is generally not a convex set.

In the real and Hermitian cases, we obtain the exact optimal value of the function while in the complex case, we can only derive some upper bounds for this optimal value.

1) Hermitian case: For a Hermitian matrix A_H, the field of values is a real interval and any point α of the field of value can be be bounded by

$$\lambda_{min}(A_H) \leq \alpha \leq \lambda_{max}(A_H)$$

with $\lambda_{min}(A_H)$ and $\lambda_{max}(A_H)$ the smallest and largest eigenvalues of A_H. This interval is the smallest that contains $F(A_H)$. The solution of (10) is then the product of the adequate extremal (smallest and largest) eigenvalues of the Hermitian matrices A_H and B_H depending on their signs. The solutions u and v providing the optimum are the eigenvectors of A_H and B_H corresponding to the eigenvalues providing the solution, respectively.

2) General complex case: For complex matrices A and B, the field of values is a set of complex values. Any point $\alpha + j\beta$ of $F(A)$ and $\gamma + j\delta$ of $F(B)$ satisfies

$$\begin{aligned} \lambda_{min}(A_H) &\leq \alpha \leq \lambda_{max}(A_H), \\ \lambda_{min}(A_S) &\leq \beta \leq \lambda_{max}(A_S), \\ \lambda_{min}(B_H) &\leq \gamma \leq \lambda_{max}(B_H), \\ \lambda_{min}(B_S) &\leq \delta \leq \lambda_{max}(B_S), \end{aligned}$$

where $A_H = \frac{A+A^*}{2}$, $B_H = \frac{B+B^*}{2}$ represent the Hermitian parts of A and B, $A_S = \frac{A-A^*}{2j}$, $B_S = \frac{B-B^*}{2j}$ the skew-Hermitian parts of A and B. These intervals define the smallest boxes containing $F(A)$ and $F(B)$. These bounds on the fields of values provide an upper bound for the problem (10).

3) Real case: For a real matrix A, the field of value could be complex in general. The real field of values associated with a real square matrix A is defined by [7]

$$F_R(A) := \{x^T Ax : x \in \mathbb{R}^n, x^T x = 1\}.$$

If we notice that $F_R(A) = F_R(A_S)$, with $A_S = \frac{A+A^T}{2}$ the symmetric part of A, then it is sufficient to consider only the symmetric part of the matrix in order to study the real field of values. The solution of (10), for A, B, u and v real, is then the product of the adequate extremal eigenvalues of the symmetric parts of A and B depending on their signs. The solutions u and v are the eigenvectors of A_S and B_S corresponding to the eigenvalues forming the optimum.

In the particular case of real symmetric matrices, it can be shown that our problem is linked to the concept of the similarity matrix S introduced in [1] which expresses how similar the vertices of two graphs are.

IV. The general case

In this section we provide some results obtained for the general problem

$$\max_{\substack{U^*U = I_k \\ V^*V = I_k}} \langle U^*AU, V^*BV \rangle_H \tag{11}$$

where $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, $U \in \mathbb{C}^{m \times k}$ and $V \in \mathbb{C}^{n \times k}$ with $k \leq \min(m,n)$. This problem is equivalent to maximizing the correlation between two projected matrices A and B subject to isometry constraints.

We derive first the expressions for the fixed points of the optimization problem. Then we consider some particular cases, i.e. when one matrix is Hermitian, when the two matrices are normal and finally when $k = \min(m,n)$, i.e. U or V is a square matrix. An upper and a lower bound to the general problem are also obtained by decomposing the problem into the sum of two Hermitian problems.

A. Fixed points

We consider the optimization problem (11) which is an optimization problem of a continuous function on a compact domain. There always exists a solution U and V optimizing the function such that the first order conditions are satisfied. These first-order derivative optimality conditions can be derived from the Lagrangian

$$F = \langle U^*AU, V^*BV \rangle_H + \langle U^*A^*U, V^*B^*V \rangle_H \\ + \langle X, (I - U^*U) \rangle_H + \langle Y, (I - V^*V) \rangle_H \tag{12}$$

where X and Y are Hermitian matrices of Lagrange multipliers. By taking a particular coordinate system, the first order conditions can be expressed by:

$$U\Sigma = A^*U(V^*BV) + AU(V^*B^*V), \tag{13}$$

$$V\Sigma = B^*V(U^*AU) + BV(U^*A^*U), \tag{14}$$

with Σ a diagonal matrix.

B. Case where one matrix is Hermitian

In this section we consider the particular case where one of the matrices is Hermitian (e.g. $A = A^*$). The maximum of (11) is then achieved for matrices U and V corresponding respectively to the dominant eigenvectors of A and $(B + B^*)$. Moreover $U\Sigma V^*$ is exactly of rank k. In other words in this case the problem is decoupled regarding the matrices A and B. The optimum is characterized by

$$\Sigma = U^*AUV^*(B+B^*)V = \begin{pmatrix} \alpha_1\beta_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_k\beta_k \end{pmatrix}$$

where $\alpha_1, \cdots, \alpha_k$ and $\beta_1 \cdots, \beta_k$ are k real eigenvalues of A and $(B + B^*)$ ordered in such a way that the optimum of (11) equal to

$$\frac{1}{2}\text{tr}\Sigma = \frac{1}{2}\left(\sum_{i=1}^{k} \alpha_i\beta_i\right) \tag{15}$$

is optimal for all the combinations of the eigenvalues.

In practice, the maximal sum is obtained for adequate combinations of the eigenvalues and can be constructed via a simple procedure.

C. Sum of two Hermitian problems

A square matrix A can always be decomposed into

$$A = A_H + jA_S$$
$$B = B_H + jB_S$$

where the matrices

$$A_H = \frac{A + A^*}{2}, \quad A_S = \frac{A - A^*}{2j}$$

are Hermitian matrices. The objective function can then also be decomposed into a sum of two Hermitian problems

$$\frac{1}{2}\langle U^*A_HU, V^*B_HV \rangle_H + \frac{1}{2}\langle U^*A_SU, V^*B_SV \rangle_H. \tag{16}$$

From (15), an upper bound for the optimal value can be obtained for (16), expressed as follows:

$$\frac{1}{2}\left(\sum_{i=1}^{k} \alpha_i^H\beta_i^H\right) + \frac{1}{2}\left(\sum_{i=1}^{k} \alpha_i^S\beta_i^S\right)$$

where α_i^H and β_i^H represent the eigenvalues of A_H and B_H ordered in such a way that $\frac{1}{2}\left(\sum_{i=1}^{k} \alpha_i^H\beta_i^H\right)$ is maximal over all possible combinations, and where α_i^S and β_i^S are the eigenvalues of A_S and B_S ordered such that $\frac{1}{2}\left(\sum_{i=1}^{k} \alpha_i^S\beta_i^S\right)$ is maximal.

A lower bound can also be found by choosing the matrices U and V optimizing one of the two Hermitian problems and by calculating the value of (16) for this pair of matrices which is the same for the two problems.

D. Case of two normal matrices

In case of normal matrices A and B (i.e. $AA^* = A^*A$ and $BB^* = B^*B$), the optimal value for the optimization function can be found for $k = 1$ and $k = m = n$. In general, for $k \leq \min(m,n)$, we can only provide an upper bound for the problem. The following developments are based on the fact that all normal matrices are diagonalizable under unitary transformation. We can thus transform the matrices A and B to diagonal matrices D_A and D_B by unitary transformations. We decompose them into real and imaginary parts

$$D_A = D_{A_R} + jD_{A_I}$$
$$D_B = D_{B_R} + jD_{B_I}$$

where the subscripts $_R$ and $_I$ denote respectively the real and imaginary parts of the matrices.

1) One-dimensional case: For $k = 1$ and by using the diagonalization of the normal matrices, (11) can be expressed as follows:

$$\max_{\substack{u^*u = 1 \\ v^*v = 1}} \langle u^*D_Au, v^*D_Bv \rangle_H.$$

This problem is equivalent to

$$\max \Re(\sum_{i=1}^{n} \mu_i \alpha_i)(\sum_{i=1}^{m} \nu_i \beta_i)$$
$$\text{s.t.} \quad \sum_i \mu_i = 1$$
$$\sum_i \nu_i = 1$$
$$\mu_i \geq 0, \ \nu_i \geq 0$$

where α_i and β_i are the eigenvalues of A and B respectively, $\mu_i = |u_i|^2$ and $\nu_i = |v_i|^2$ are nonnegative real numbers. We optimize then the real part of the products of convex combinations of the eigenvalues of A and B. This problem is a bilinear form with respect to μ_i and ν_i. If we fix μ_i the problem is linear in ν_i and we resolve a linear programming problem. The feasible set forms a polyhedron and the optimal solution is situated on a vertex of this polyhedron (or on a face of the polyhedron). We apply then the same reasoning for μ_i to obtain the optimal solution. The problem is then equivalent to finding the indices i and j maximizing

$$\max_{i,j} \Re(\alpha_i \beta_j).$$

If we separate the real and imaginary parts of α_i and β_i the optimal value of the function becomes

$$\max_{i,j} (\alpha_{i_R} \beta_{j_R} + \alpha_{i_I} \beta_{j_I}), \tag{17}$$

where the subscripts $_R$ and $_I$ denote respectively the real and imaginary parts. This problem can be solved in $O(mn)$ operations.

2) Square matrices: For $k = n$ and by using again the diagonalization of the matrices, (11) becomes:

$$\max_{Q^*Q = I} \ \langle Q^* D_{A_R} Q, D_{B_R} \rangle + \langle Q^* D_{A_I} Q, D_{B_I} \rangle.$$

By developing the first term and using a result of Birkhoff (see [10]), the value at the optimum can be shown to be

$$\sum_{k,l=1}^{n} (\alpha_{k_R} \beta_{l_R} + \alpha_{k_I} \beta_{l_I}) \tag{18}$$

where α_{k_R}, β_{l_R}, α_{k_I} and β_{l_I} are the elements of $diag(D_{A_R})$, $diag(D_{B_R})$, $diag(D_{A_I})$ and $diag(D_{B_I})$, ordered in such a way that (18) is maximal. The optimal value is thus obtained for an adequate combination of the real and imaginary parts of the eigenvalues of A and B, which is a linear programming problem.

3) General case: For $1 \leq k \leq min(m,n)$ we optimize

$$\max_{\substack{U^*U = I_k \\ V^*V = I_k}} \langle U^* D_{A_R} U, V^* D_{B_R} V \rangle + \langle U^* D_{A_I} U, V^* D_{B_I} V \rangle.$$

An upper bound to this problem is then

$$\sum_{i,j=1}^{k} (\alpha_{i_R} \beta_{j_R}) + \sum_{i,j=1}^{k} (\alpha_{i_I} \beta_{j_I}) \tag{19}$$

where the elements α_{i_R} and β_{i_R} in d_{A_R} and d_{B_R} are ordered in such a way that $\sum_{i,j=1}^{k}(\alpha_{i_R}\beta_{j_R})$ is maximal and the elements α_{i_I} and β_{i_I} in d_{A_I} and d_{B_I} ordered such that $\sum_{i,j=1}^{k}(\alpha_{i_I}\beta_{j_I})$ is maximal. This problem is combinatorial and differs from (18).

E. Matrices U and V of maximal size

For $k = \min(m,n)$, the problem (11) is equivalent to

$$\max_{Q^*Q = I} \ \langle Q^* A Q, B \rangle_H$$

where $Q = UV^*$ is an isometry of dimension $m \times n$. The general problem is thus reduced to an optimization problem with only one variable Q.

V. ALGORITHM

In this section we present an iterative algorithm to solve the problem (11) whose fixed points satisfy the first-order derivative conditions (13) and (14) for U and V. This iteration is the following:

$$U_{i+1}\Sigma_{i+1}V_{i+1}^* - U_{\perp_{i+1}}\Sigma_{-i+1}V_{\perp_{i+1}}^* = AU_iV_i^*B^* + A^*U_iV_i^*B \tag{20}$$

where $U_\perp$ and $V_\perp$ are orthogonal complements of U and V and Σ is a diagonal matrix. The subscript $i+1$ represents the new stage of the iteration while the subscript i is the current stage. $U_{i+1}\Sigma_{i+1}V_{i+1}^*$ is the best approximation of rank k of $AU_iV_i^*B^* + A^*U_iV_i^*B$.

In practice (20) is realized by application of the SVD algorithm until convergence. At this moment, the solution satisfies

$$U\Sigma V^* = AUV^*B^* + A^*UV^*B + U_\perp \Sigma_- V_\perp^*. \tag{21}$$

The convergence is not proved but in all our experiments the process always converged linearly to a solution.

A. Relation to the optimization problem

In this part we show that the iteration is equivalent to solving the optimization problem

$$\max_{\substack{U^*U = I_k \\ V^*V = I_k}} \langle U^* A U, V^* B V \rangle_H \tag{22}$$

whose fixed points are expressed by (13) and (14), i.e.

$$U\Sigma = A^*U(V^*BV) + AU(V^*B^*V), \tag{23}$$
$$V\Sigma = B^*V(U^*AU) + BV(U^*A^*U). \tag{24}$$

The expression (21) implies (23) and (24) by simply right and left multiplying left- and right-hand sides of (21) by V and U^*. Then a fixed point exists for the iterative algorithm.

Conversely we can prove that (23) and (24) imply (21). To show this we multiply (23) by V^*:

$$U\Sigma V^* = A^*UV^*BVV^* + AUV^*B^*VV^*$$
$$= N(I - V_\perp V_\perp^*)$$

with $N = (A^*UV^*BVV^* + AUV^*B^TVV^*)$. Here we used $VV^* = I - V_\perp V_\perp^*$ because V is an isometry, then its projector VV^* can be linked to its complementary projector $V_\perp V_\perp^*$. Similarly by multiplying (24) by U^* and taking the transpose, we obtain

$$U\Sigma V^* = (I - U_\perp U_\perp^*)N.$$

An orthogonal projector $P_\perp$ of N satisfies the property $P_\perp(P_\perp N) = P_\perp N$. This property implies

$$NV_\perp V_\perp^* = U_\perp U_\perp^* NV_\perp V_\perp^* = U_\perp U_\perp^* N.$$

Therefore

$$U\Sigma V^* = N(I - V_\perp V_\perp^*) = (I - U_\perp U_\perp^*)N$$

lies in the orthogonal complement of U and also of V. That proves the existence of a fixed point of (20).

B. Non-convexity

For arbitrary matrices A and B the set

$$\{\langle U^*AU, V^*BV\rangle_H : U, V \text{ isometries}\}$$

is in general not convex. Local minima and local maxima can exist and then the algorithm may not always converge to the global optimum. We can notice the non-convexity of the set by taking a particular example as follows. In case of square matrices of the same dimensions, we optimize the B-numerical range of A, as defined previously. In general $W_B(A)$ is not a convex set, e.g. [11] gave an example in which B is normal but not Hermitian and where $W_B(A)$ is not convex. It is then easy to choose the coordinates for U and V to generate a non-convex example.

VI. CONCLUSION

In this paper, we analyze the correlation between projections in lower-dimensional subspace of matrices submitted to isometry constraints. Our problem provides a method to project simultaneously the matrices in subspaces of arbitrary dimensions and can be applied to both real and complex matrices. We indicate that it is an extension of various problems found in the literature. Many applications can arise from this formulation.

We present some mathematical properties of the problem and we characterize the maximal correlation for particular matrices such as Hermitian or normal matrices. In general only an upper bound can be found.

We develop an iterative algorithm providing the optimum and we characterize the fixed points. This algorithm is very simple to implement and is based on the SVD. Because this problem is not convex, the analysis of convergence and stability of the fixed points is difficult.

Investigations of mathematical properties and applications of the correlation between projected matrices can be pursued in several directions. A deeper analysis of the convergence of the algorithm is worthwhile to consider. We outline a few possible improvements and future research directions. The first possible improvement concerns the convergence of the algorithm. Experimentally we observe a linear convergence to the optimum but this convergence has not yet been proved and remains an important point to develop in the future. Secondly, because the problem is not convex, the analysis of the stability of the fixed points and the study of their basins of attraction is not easy to obtain. This last point is thus a delicate but interesting task to explore. From a more applied point of view, another topic of interest is to investigate how the mathematical concepts proposed here can be used, possibly in modified form, for applications in various areas. We believe that the problem considered in this paper gives rise to interesting mathematical problems but also to useful applications in different areas.

REFERENCES

[1] V. D. Blondel, A. Gajardo, M. Heymans, P. Senellart, and P. Van Dooren, "A measure of similarity between graph vertices: applications to synonym extraction and Web searching," *SIAM Rev.*, vol. 46, no. 4, pp. 647–666, 2004.

[2] T. Caelli and S. Kosinov, "An eigenspace projection clustering method for inexact graph matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 4, pp. 515–519, 2004.

[3] U. Helmke, K. Hüper, J. B. Moore, and T. Schulte-Herbrüggen, "Gradient flows computing the C-numerical range with applications in NMR spectroscopy," *J. Global Optim.*, vol. 23, no. 3-4, pp. 283–308, 2002.

[4] C. K. Li and J. Woerdeman, "A lower bound on the C-numerical radius of nilpotent matrices appearing in coherent spectroscopy," *SIAM J. Matrix Anal. Appl.*, vol. accepted for publication, 2005.

[5] C. K. Li and N. K. Tsing, "Norms that are invariant under unitary similarities and the C-numerical radii," *Linear Multilinear Algebra*, vol. 24, pp. 209–222, 1989.

[6] C. K. Li and L. Rodman, "Multiplicative preservers of C-numerical ranges and radii," *Linear Multilinear Algebra*, vol. 52, pp. 265–279, 2004.

[7] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. New York, NY, USA: Cambridge University Press, 1991.

[8] K. Anstreicher and H. Wolkowicz, "On Lagrangian relaxation of quadratic matrix constraints," *SIAM J. Matrix Anal. Appl.*, vol. 22, no. 1, pp. 41–55, 2000.

[9] M. Goldberg and E. G. Straus, "Elementary inclusion relations for generalized numerical ranges," *Linear Algebra Appl.*, vol. 18, pp. 1–24, 1977.

[10] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY, USA: Cambridge University Press, 1985.

[11] R. Westwick, "A theorem on numerical range," *Linear Multilinear Algebra*, vol. 2, pp. 311–315, 1975.