SIAM J. MATRIX ANAL. APPL.
Vol. 12, No. 3, pp. 401–425, July 1991

001

THE RESTRICTED SINGULAR VALUE DECOMPOSITION: PROPERTIES AND APPLICATIONS*

BART L. R. DE MOOR† AND GENE H. GOLUB‡

Abstract. The *restricted singular value decomposition* (RSVD) is the factorization of a given matrix, relative to two other given matrices. It can be interpreted as the *ordinary singular value decomposition* with different inner products in row and column spaces. Its properties and structure, as well as its connection to generalized eigenvalue problems, canonical correlation analysis, and other generalizations of the singular value decomposition, are investigated in detail.

Applications that are discussed include the analysis of the extended shorted operator, unitarily invariant norm minimization with rank constraints, rank minimization in matrix balls, the analysis and solution of linear matrix equations, rank minimization of a partitioned matrix, and the connection with generalized Schur complements, constrained linear and total linear least squares problems with mixed exact and noisy data, including a generalized Gauss–Markov estimation scheme.

Key words. generalized SVD, generalized matrix inverses, (total) linear least squares, (generalized) Schur complements, matrix balls, shorted operator

AMS(MOS) subject classifications. 15A09, 15A18, 15A21, 15A24, 65F20

1. Introduction.

The *ordinary singular value decomposition* (OSVD) has a long history with original contributions by Beltrami (1873) [2], Sylvester (1889) [26], Autonne (1902) [1], Eckart and Young (1936) [12] and many others (see, e.g., the references in [15], [21], [27]). It has become an important tool in the analysis and numerical solution of numerous problems arising in such diverse applications as psychometrics, statistics, signal processing, and system theory. Not only does it allow for an elegant problem formulation, but at the same time it provides geometrical and algebraic insight together with an immediate numerically robust implementation [15].

Recently, several generalizations to the OSVD have been proposed and their properties analysed. The one that is best known is the *generalized* SVD as introduced by Paige and Saunders in 1981 [22], which we propose to rename as the *Quotient* SVD (QSVD) [8]. Another example is the *Product* SVD (PSVD) as proposed by Fernando and Hammarling in 1987 [14] and further analysed in [10]. The third one is the *Restricted* SVD (RSVD), introduced in its explicit form by Zha in [32] and further developed and discussed in this paper. In [8] we have proposed a standardized nomenclature for the singular value decomposition and its generalizations. This set of names has the advantage of being *alphabetic* and *mnemonic*, **O-P-Q-R-SVD**. For the structure and properties of the OSVD, PSVD, and QSVD, we also refer to [8].

The RSVD, which is the main subject of this paper, applies for a given triplet of matrices A, B, C of compatible dimensions (Theorem 1). In essence, the RSVD provides a factorization of the matrix A, relative to matrices B and C. It could be

* Received by the editors June 8, 1989; accepted for publication (in revised form) September 12, 1990. Part of this work was supported by the United States Army under contract DAAL03-87-K-0095.

† Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94 B-3001, Leuven, Belgium (demoor@esat.kuleuven.ac.be). This author was a visiting research associate at the Computer Science Department and the Department of Electrical Engineering (Information Systems Laboratory) of Stanford University, where he was supported by an Advanced Research Fellowship in Science and Technology of the North Atlantic Treaty Organization (NATO) Science Fellowships Program and by a grant from IBM. He is now a research associate of the Belgian National Fund for Scientific Research (NFWO).

‡ Department of Computer Science, Stanford University, Stanford, California 94305 (na.golub@na-net.stanford.edu).

considered as the OSVD of the matrix A, but with different (possibly nonnegative-definite) inner products in its column and in its row space. It will be shown that the RSVD not only allows for an elegant treatment of algebraic and geometric problems in a wide variety of applications, but that its structure provides a powerful tool in simplifying proofs and derivations that are algebraically rather complicated.

Soon after the present paper was completed, Zha and de Moor discovered that the RSVD is only one of the three possible SVD-like factorizations for three matrices. Similar generalizations of the OSVD are not only limited to two or three matrices, but can be derived for 4, 5, $\cdots$, i.e., any number of matrices of compatible dimensions. The PSVD and the QSVD serve as basic building blocks in this infinite tree of generalizations of the OSVD. For instance, the RSVD which is analysed in this paper can also be considered as a double QSVD. This is the reason why we have called it the **QQ-SVD** in [11], where the complete structure of this tree of generalizations is also developed in detail.

This paper is organised as follows. In §2, the main structure of the RSVD is analysed in terms of the ranks of the concatenation of certain matrices. The factorization is related to a generalized eigenvalue problem (§2.2.1). A variational characterization is provided in §2.2.2. A generalized dyadic decomposition is explored in §2.2.3 together with a geometrical interpretation. It is shown how the RSVD contains other generalizations of the OSVD, such as the PSVD and the QSVD, as special cases in §2.2.4. In §3, several applications are discussed:

— *Rank minimization* and the *extended shorted operator* are the subject of §3.1, as well as *unitarily invariant norm minimization with rank constraints* and the relation with *matrix balls.* We also investigate a certain linear matrix equation which is directly related to the Moore–Penrose pseudo-inverse of a matrix.

— The *low rank approximation of a partitioned matrix* when only one of its blocks can be modified is explored in §3.2, together with *total least squares with mixed exact and noisy data and linear constraints.* While the role of the Schur complement and its close connection to least squares estimation is well understood, it will be shown in this section that there exists a similar relation between constrained total linear least squares solutions and a *generalized Schur complement.*

— *Generalized Gauss–Markov models*, possibly with constraints, are discussed in §3.3 and it is shown how the RSVD simplifies the solution of linear least squares problems with constraints.

In §4 the main conclusions are presented together with some perspectives. Let us conclude this Introduction by referring to the reports mentioned in [9] for a detailed constructive proof of the main theorem of this paper.

Notation, conventions, and abbreviations. Throughout the paper, capitals denote matrices. The lower case letters $i, j, k, l, m, n, p, q, r$ are nonnegative integers. Other lower case letters denote vectors. The set of real numbers is denoted by $\Re$. Scalars (possibly complex) are denoted by Greek letters. The matrices A $(m \times n)$, B $(m \times p)$, C $(q \times n)$ are given matrices. Their ranks will be denoted by r_a, r_b, r_c. D is a $p \times q$ matrix. M is the matrix with A, B, C, D^* as its blocks: $M = \begin{pmatrix} A & B \\ C & D^* \end{pmatrix}$. We shall also frequently use the following ranks: $r_{ac} = \operatorname{rank}\begin{pmatrix} A \\ C \end{pmatrix}, r_{abc} = \operatorname{rank}\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}$, $r_{ab} = \operatorname{rank}(\ A \ \ B\ )$. A^t is the transpose of a (possibly complex) matrix A and $\overline{A}$ is the complex conjugate of A. A^* denotes the complex conjugate transpose of a (complex) matrix: $A^* = \overline{A}^t$. The matrix A^{-*} represents the inverse of A^*. I_k is the $k \times k$

identity matrix. The subscript is omitted when the dimensions are clear from the context. Identity vectors with the ith component equal to 1 and all others zero, are denoted by e_i $(m \times 1)$. A matrix X is called an $A(i,j,\cdots)$-inverse of the matrix A if it satisfies equation $i,j,\cdots$ of the following:

1. $AXA = A$,
2. $XAX = X$,
3. $(AX)^* = AX$,
4. $(XA)^* = XA$.

An $A(1)$ inverse is also called an inner inverse and denoted by A^-. The $A(1,2,3,4)$ inverse is the Moore–Penrose pseudo-inverse denoted by A^+ and it is unique. We shall also need the following lemmas.

LEMMA 1 (inner inverse of a factored matrix). *Let the matrix A be factored as*

$$A = P^{-*}\begin{pmatrix} D_a & 0 \\ 0 & 0 \end{pmatrix} Q^{-1}$$

where D_a is square $r_a \times r_a$ nonsingular. Then, every inner inverse A^- can be written as

$$A^- = Q\begin{pmatrix} D_a^{-1} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} P^* \tag{1}$$

where Z_{12}, Z_{21}, Z_{22} are arbitrary matrices. Conversely, every matrix A^- of this form is an inner inverse of A.

For a detailed discussion of generalized inverses, we refer to [21]. The matrices U_a $(m\times m)$, V_a $(n\times n)$, V_b $(p\times p)$, U_c $(q\times q)$ are unitary, i.e., $U_aU_a^* = I_m = U_a^*U_a, V_aV_a^* = I_n = V_a^*V_a$, $V_bV_b^* = I_p = V_b^*V_b$, $U_cU_c^* = I_q = U_c^*U_c$. The matrices P $(m \times m)$ and Q $(n\times n)$ are square nonsingular. The nonzero elements of the diagonal matrices S_1, S_2, and S_3, which appear in the theorems, are denoted by α_i, β_i, and γ_i. The vector a_i denotes the ith column of the matrix A. The range (column space) of the matrix A is denoted by $\mathbf{R}(A) = \{y|y = Ax\}$. The row space of A is denoted by $\mathbf{R}(A^*)$. The null space of the matrix A is represented as $\mathbf{N}(A) = \{x|Ax = 0\}$. The symbol $\bigcap$ denotes the intersection of two vector spaces. We shall use the following well-known result.

LEMMA 2 (the dimension of the intersection of subspaces).

$$\dim(\mathbf{R}(A)\bigcap\mathbf{R}(B)) = r_a + r_b - r_{ab}$$
$$\dim(\mathbf{R}(A^*)\bigcap\mathbf{R}(C^*)) = r_a + r_c - r_{ac}.$$

$\|A\|$ is any unitarily invariant matrix norm while $\|A\|_F$ is the Frobenius norm: $\|A\|_F^2 = \operatorname{trace}(AA^*)$. The norm of the vector a is denoted by $\|a\|_2$ where $\|a\|_2^2 = a^*a$. Moreover, we will adopt the following convention for block matrices: Any (possibly rectangular) block of zeros is denoted by 0, the precise dimensions being obvious from the block dimensions. The symbol I represents a matrix block corresponding to the square identity matrix of appropriate dimensions. Whenever a dimension indicated by an integer in a block matrix is zero, the corresponding block row or block column should be omitted and all expressions and equations in which a block matrix of that block row or block column appears, can be disregarded. An equivalent formulation would be that we allow $0 \times n$ or $n \times 0$ $(n \neq 0)$ blocks to appear in matrices. This permits an elegant treatment of several cases at once. Finally, we would like to introduce the term *quasi-diagonal* matrix for a matrix, the block rows and block columns of which are a permutation of a diagonal matrix.

2. The restricted singular value decomposition (RSVD). The idea of a generalization of the OSVD for three matrices is implicit in the S, T-singular value decomposition of Van Loan [30] via its relation to a generalized eigenvalue problem. Zha [32] introduced an explicit formulation of the RSVD constructing it through the use of several OSVDs and QSVDs (see also [9]). For the sake of brevity, we have omitted our constructive proof based on a sequence of OSVDs and PSVDs. It can be found in [9]. In this section, we first state the main theorem (§2.1), which describes the structure of the RSVD, followed by a discussion of the main properties in §2.2, including the connection to generalized eigenvalue problems, a generalized dyadic decomposition, geometrical insights, and the demonstration that the RSVD contains the OSVD, the PSVD, and the QSVD as special cases.

2.1. The RSVD theorem. With the notation and conventions of §1, we have the following theorem.

THEOREM 1 (the restricted singular value decomposition). *Every triplet of matrices A $(m \times n)$, B $(m \times p)$, and C $(q \times n)$ can be factorized as*

$$\begin{aligned} A &= P^{-*} S_a Q^{-1}, \\ B &= P^{-*} S_b V_b^*, \\ C &= U_c S_c Q^{-1}, \end{aligned}$$

where P $(m \times m)$ and Q $(n \times n)$ are square nonsingular, and V_b $(p \times p)$ and U_c $(q \times q)$ are unitary. S_a $(m \times n)$, S_b $(m \times p)$, and S_c $(q \times n)$ are real quasi-diagonal matrices with nonnegative elements and the following block structure:

$$\begin{pmatrix} S_a & S_b \\ S_c & \end{pmatrix} = \begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\ \\ 1\\2\\3\\4 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1&2&3&4&5&6&1&2&3&4 \end{array} \\ \left(\begin{array}{cccccccccc} S_1 & 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & & & & \\ I & 0 & 0 & 0 & 0 & 0 & & & & \\ 0 & I & 0 & 0 & 0 & 0 & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & & & & \\ 0 & 0 & 0 & 0 & S_3 & 0 & & & & \end{array} \right) \end{array}$$

The block dimensions of the matrices S_a, S_b, S_c are the following.

	Block columns of S_a and S_c	*Block columns of S_b*
1.	$r_{abc} + r_a - r_{ac} - r_{ab}$	$r_{abc} + r_a - r_{ac} - r_{ab}$
2.	$r_{ab} + r_c - r_{abc}$	$r_{ac} + r_b - r_{abc}$
3.	$r_{ac} + r_b - r_{abc}$	$p - r_b$
4.	$r_{abc} - r_b - r_c$	$r_{ab} - r_a$
5.	$r_{ac} - r_a$	
6.	$n - r_{ac}$	
	Block rows of S_a and S_b	*Block rows of S_c*
1.	$r_{abc} + r_a - r_{ab} - r_{ac}$	$r_{abc} + r_a - r_{ab} - r_{ac}$
2.	$r_{ab} + r_c - r_{abc}$	$r_{ab} + r_c - r_{abc}$
3.	$r_{ac} + r_b - r_{abc}$	$q - r_c$
4.	$r_{abc} - r_b - r_c$	$r_{ac} - r_a$
5.	$r_{ab} - r_a$	
6.	$m - r_{ab}$	

The matrices S_1, S_2, S_3 are square nonsingular diagonal with positive diagonal elements.

Let $\alpha_i, \beta_j, \gamma_k$ be the diagonal elements of the matrices S_1, S_2, S_3. We propose to call the following triplets of numbers the *restricted singular value triplets*:

- $r_{abc} + r_a - r_{ab} - r_{ac}$ triplets of the form $(\alpha_i, 1, 1)$ with $\alpha_i > 0$. By convention, they will be ordered as

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_{r_{abc}+r_a-r_{ab}-r_{ac}} > 0.$$

- $r_{ab} + r_c - r_{abc}$ triplets of the form $(1, 0, 1)$.
- $r_{ac} + r_b - r_{abc}$ triplets of the form $(1, 1, 0)$.
- $r_{abc} - r_b - r_c$ triplets of the form $(1, 0, 0)$.
- $r_{ab} - r_a$ triplets of the form $(0, \beta_j, 0)$, $\beta_j > 0$ (elements of S_2).
- $r_{ac} - r_a$ triplets of the form $(0, 0, \gamma_k)$, $\gamma_k > 0$ (elements of S_3).
- $\min(m - r_{ab}, n - r_{ac})$ trivial triplets $(0, 0, 0)$.

We propose to call the factorization of a matrix triplet, as described in Theorem 1, the *restricted* singular value decomposition because the RSVD allows us to analyse matrix problems that can be stated in terms of the matrices $A + BDC$ and

$$M = \begin{pmatrix} A & B \\ C & D^* \end{pmatrix},$$

in which the matrices B and C represent certain *restrictions* on the type of operations that are allowed. Typically, we are interested in the ranks of these matrices as the matrix D is modified. The rank of the matrix $A + BDC$ can only be reduced by modifications that belong to the column space of B and the row space of C. It will be shown how the rank of M can be analysed via a generalized Schur complement, which is of the form $D^* - CA^-B$, where again, C and B represent certain restrictions and A^- is an inner inverse of A. Moreover, the RSVD yields the restriction of the linear operator represented by the matrix A to the column space of B and the row space of C. Finally, the RSVD can be interpreted as an OSVD but with certain restrictions on the inner products to be used in the column and row space of the matrix A (see §2.2.1).

Some algorithmic issues related to the RSVD are discussed in [11], [13], [29], [28], and [33], though a full portable and documented algorithm for the RSVD is still to be developed.

2.2. Properties of the RSVD. The OSVD, as well as the PSVD and the QSVD, can all be related to a certain (generalized) eigenvalue problem. It comes as no surprise that this is also the case for the RSVD. First, the generalized eigenvalue problem for the RSVD will be analysed in §2.2.1 and we shall point out an interesting connection with canonical correlation analysis. A variational characterization of the RSVD is provided in §2.2.2. A generalized dyadic decomposition and some geometrical properties are investigated in §2.2.3. In §2.2.4, it is shown how the OSVD, PSVD, and QSVD are special cases of the RSVD.

2.2.1. Relation to a generalized eigenvalue problem. Consider the generalized eigenvalue problem

$$(2) \qquad \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} BB^* & 0 \\ 0 & C^*C \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \lambda.$$

Let p_i be the ith column of P and q_i the ith column of Q. Obviously, the column vector $(p_i^* \;\; q_i^*)^*$ is a generalized eigenvector of the pencil (2). There are four types of generalized eigenvalues (finite nonzero, zero, infinite, and arbitrary), which can be related to the restricted singular value triplets of Theorem 1.

Note that if $BB^* = I_m$ and $C^*C = I_n$, the eigenvalues λ are $\pm$ the singular values of the matrix A. In the case that the matrices BB^* and C^*C are nonsingular, it can be shown that the generalized eigenvalue problem (2) is equivalent to a singular value decomposition. It follows from (2) that

$$\begin{aligned} Aq_i &= BB^* p_i \lambda_i, \\ A^* p_i &= C^*C q_i \lambda_i. \end{aligned}$$

If BB^* and C^*C are both nonsingular, then there exist square nonsingular matrices W_b and W_c (for example, the Cholesky decomposition) such that $BB^* = W_b^* W_b$ and $C^*C = W_c^* W_c$. Then, we have that

$$\begin{aligned} (W_b^{-*} A W_c^{-1})(W_c q_i) &= (W_b p_i)\lambda_i, \\ (W_c^{-*} A^* W_b^{-1})(W_b p_i) &= (W_c q_i)\lambda_i. \end{aligned}$$

From Theorem 1, it follows that $P^*(BB^*)P = S_b S_b^t$ and $Q^*(C^*C)Q = S_c^t S_c$. Hence, if BB^* is nonsingular, the column vectors of P are orthogonal with respect to the inner product provided by the positive-definite matrix BB^*. A similar observation applies for the column vectors of Q with respect to C^*C. The BB^*-orthogonality of the vectors p_i and the C^*C-orthogonality of the vectors q_i implies that the vectors $W_b p_i$ and $W_c q_i$ are (multiples of) the left and right singular vectors of the matrix $W_b^{-*} A W_c^{-1}$.

Consider the RSVD of the matrix triplet (A^*B, A^*, B) and its related generalized eigenvalue problem:

$$\begin{pmatrix} 0 & A^*B \\ B^*A & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} A^*A & 0 \\ 0 & B^*B \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \sigma.$$

This is nothing more than the eigenvalue problem that arises in *canonical correlation analysis* (principal angles and vectors between subspaces; see, e.g., [3], [15]). There exist applications where the matrices BB^* and C^*C are (almost) singular (see, e.g., [13], [18]). The matrices BB^* and C^*C can be (sample) covariance matrices that are (almost) singular. This is, for instance, the case in [18], where a generalized type of canonical correlation analysis is required, allowing singular covariance matrices. Another example is generalized Gauss–Markov estimation as described in §3.3. It is in these situations that the RSVD may provide essential insight into the geometry of the singularities and at the same time yield a numerically robust and elegant implementation of the solution by avoiding the explicit solution (with its "implicit squaring") of the generalized eigenvalue problem.

2.2.2. A variational characterization. Let $\phi(x, y) = x^*Ay$ be a bilinear form of 2 vectors x and y. We wish to maximize $\phi(x, y)$ over all vectors x, y subject to $x^*BB^*x = 1$ and $y^*C^*Cy = 1$. It follows directly from the RSVD that a solution exists only if one of the following situations occurs:

- $r_{abc} + r_a - r_{ab} - r_{ac} \neq 0$. In this case, the maximum is equal to the largest diagonal element of S_1 and the optimizing vectors are $x = p_1$ (first column vector of P) and $y = q_1$ (first column vector of Q) so that $\phi(p_1, q_1) = \alpha_1$.

- $r_{abc} + r_a - r_{ab} - r_{ac} = 0$. The norm constraints on x and y can only be satisfied if

$$r_{ac} + r_b - r_{abc} > 0 \quad \text{or} \quad r_{ab} - r_a > 0$$

and

$$r_{ab} + r_c - r_{abc} > 0 \quad \text{or} \quad r_{ac} - r_a > 0.$$

In either case, the maximum is 0.

If none of these conditions is satisfied, there is *no* solution.

Assume that the maximum is achieved for the vectors $x_1 = p_1$ and $y_1 = q_1$. Then, other extrema of the objective function $\phi(x, y) = x^* A y$, constrained to lie in subspaces that are BB^*-orthogonal to p_1 and C^*C-orthogonal to q_1, can be found in an obvious recursive manner. All of these extrema are then generated by the columns of the matrices P and Q.

2.2.3. A generalized dyadic decomposition and geometrical properties. Denote $P' = P^{-*}$ and $Q^{-1} = Q'^*$. Then, with an appropriate partitioning of the matrices P', Q', U_c, and V_b, corresponding to the diagonal structure of the matrices S_a, S_b, S_c of Theorem 1, it is straightforward to obtain the following sums:

$$\begin{aligned} A &= P_1' S_1 Q'^*_1 + P_2' Q'^*_2 + P_3' Q'^*_3 + P_4' Q'^*_4, \\ B &= P_1' V_{b1}^* + P_3' V_{b2}^* + P_5' S_2 V_{b4}^*, \\ C &= U_{c1} Q'^*_1 + U_{c2} Q'^*_2 + U_{c4} S_3 Q'^*_5. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbf{R}(P_1') + \mathbf{R}(P_3') &= \mathbf{R}(A) \bigcap \mathbf{R}(B), \\ \mathbf{R}(Q'^*_1) + \mathbf{R}(Q'^*_2) &= \mathbf{R}(A^*) \bigcap \mathbf{R}(C^*). \end{aligned}$$

The decomposition of A can be interpreted as a decomposition relative to $\mathbf{R}(B)$ and $\mathbf{R}(C^*)$: The four terms of this decomposition can be classified geometrically as follows:

	in $\mathbf{R}(B)$	not in $\mathbf{R}(B)$
in $\mathbf{R}(C^*)$	$P_1' S_1 Q'^*_1$	$P_2' Q'^*_2$
not in $\mathbf{R}(C^*)$	$P_3' Q'^*_3$	$P_4' Q'^*_4$

Obviously, the term $P_1' S_1 Q'^*_1$ represents the *restriction* of the linear operator represented by the matrix A to the column space of the matrix B **and** the row space of the matrix C, while the term $P_4' Q_4'^*$ is the restriction of A to the orthogonal complements of $\mathbf{R}(B)$ and $\mathbf{R}(C^*)$.

Also, we find that

$$\begin{aligned} \mathbf{R}(B^*) &= \mathbf{R}(V_{b1}^*) + \mathbf{R}(V_{b2}^*) + \mathbf{R}(V_{b4}^*), \\ \mathbf{R}(C) &= \mathbf{R}(U_{c1}) + \mathbf{R}(U_{c2}) + \mathbf{R}(U_{c4}), \end{aligned}$$

and

$$\begin{aligned} BV_{b3} = 0 \quad &\Longrightarrow \quad \mathbf{N}(B) = \mathbf{R}(V_{b3}), \\ U_{c3}^* C = 0 \quad &\Longrightarrow \quad \mathbf{N}(C^*) = \mathbf{R}(U_{c3}). \end{aligned}$$

Finally, some of the block dimensions in the RSVD of the matrix triplet (A, B, C) can be related to geometrical interpretations by repeated application of Lemma 2.

$$\dim\left[\mathbf{R}\begin{pmatrix} A \\ C \end{pmatrix}\bigcap\mathbf{R}\begin{pmatrix} B \\ 0 \end{pmatrix}\right] = r_{ac} + r_b - r_{abc},$$

$$\dim[\,\mathbf{R}(A \;\; B)^* \bigcap \mathbf{R}(C \;\; 0)^*\,] = r_{ab} + r_c - r_{abc},$$

$$\dim[\,\mathbf{R}(A) \bigcap \mathbf{R}(B)\,] = r_a + r_b - r_{ab},$$

$$\dim[\,\mathbf{R}(A^*) \bigcap \mathbf{R}(C^*)\,] = r_a + r_c - r_{ac}.$$

It is easy to show that

$$\mathbf{R}(Q_6') = \mathbf{N}(A) \bigcap \mathbf{N}(C),$$

$$\mathbf{R}(P_6') = \mathbf{N}(A^*) \bigcap \mathbf{N}(B^*).$$

Hence Q_6' provides a basis for the common null space of A and C, which is of dimension $n - r_{ac}$, while P_6' provides a basis for the common null space of A^* and B^*, which is of dimension $m - r_{ab}$.

2.2.4. Relation to (generalized) SVDs. The RSVD reduces to the OSVD, the PSVD, or the QSVD for special choices of the matrices A, B, and/or C. For the precise structure of the PSVD and the QSVD, we refer to [8].

THEOREM 2 (special cases of the RSVD).

1. RSVD *of* (A, I_m, I_n) *is an* OSVD *of* A.
2. RSVD *of* (I_m, B, C) *is a* PSVD *of* (B^*, C).
3. RSVD *of* (A, B, I_n) *is a* QSVD *of* (A, B).
4. RSVD *of* (A, I_m, C) *is a* QSVD *of* (A, C).

Proof. 1. $B = I_m, C = I_n$. Consider the RSVD of (A, I_m, I_n). By definition, $I_m = P^{-*}S_bV_b^*$ and $I_n = U_cS_cQ^{-1}$. This implies $P^{-*} = V_bS_b^{-1}$ and $Q^{-1} = S_c^{-1}U_c^*$. Hence, we find that $A = V_b(S_b^{-1}S_aS_c^{-1})U_c^*$, which is an OSVD of A.

2. $A = I_m$. Consider the RSVD of (I_m, B, C). Then $I_m = P^{-*}S_aQ^{-1}$, which implies $Q^{-1} = S_a^{-1}P^*$. Hence, $B^* = V_bS_b^tP^{-1}$, $C = U_c(S_cS_1^{-1})P^*$, which is nothing else than a PSVD of (B^*, C).

3. $C = I_n$. Consider the RSVD of (A, B, I_n). Then $I_n = U_cS_cQ^{-1}$, which implies $Q^{-1} = S_c^{-1}U_c^*$. Then, $A = P^{-*}(S_aS_c^{-1})U_c^*$, $B = P^{-*}S_bV_b^*$, which is (up to a diagonal scaling) a QSVD of the matrix pair (A, B).

4. $B = I_m$. The proof is similar to part 3. □

3. Applications. In this section, we shall first explore the use of the RSVD in the analysis of problems related to expressions of the form $A + BDC$ where A, B, C are given matrices. The connection with Mitra's concept of the extended shorted operator [20] and with matrix balls will be discussed, as will the solution of the matrix equation $BDC = A$, which led Penrose to rediscover the pseudo-inverse of a matrix [24], [25]. In §3.2, it is shown how the RSVD can be used to solve constrained total linear least squares problems with exact, noiseless rows and columns and the close connection to Carlson's generalized Schur complement [4] is emphasized. In §3.3, we discuss the application of the RSVD in the analysis and solution of generalized Gauss–Markov models, with and without constraints.

Throughout this section, we shall use a matrix E, defined as

$$E = V_b^*DU_c \tag{3}$$

with a block partitioning derived from the block structure of S_b and S_c as follows:

(4)

$$\begin{array}{c} \\ r_{abc}+r_a-r_{ab}-r_{ac} \\ r_{ac}+r_b-r_{abc} \\ p-r_b \\ r_{ab}-r_a \end{array} \begin{array}{c} \begin{array}{cccc} r_{abc}+r_a-r_{ab}-r_{ac} & r_{ab}+r_c-r_{abc} & q-r_c & r_{ac}-r_a \end{array} \\ \left(\begin{array}{cccc} E_{11} & E_{12} & E_{13} & E_{14} \\ E_{21} & E_{22} & E_{23} & E_{24} \\ E_{31} & E_{32} & E_{33} & E_{34} \\ E_{41} & E_{42} & E_{43} & E_{44} \end{array}\right) \end{array}.$$

3.1. On the structure of A + BDC. The RSVD provides geometrical insight into the structure of a matrix A relative to the column space of a matrix B and the row space of a matrix C. As will now be shown, it is an appropriate tool to analyse expressions of the form $A + BDC$ where D is an arbitrary $p \times q$ matrix. The RSVD allows us to analyse and solve the following questions:

1. What is the range of ranks of $A + BDC$ over all possible $p \times q$ matrices D (§3.1.1)?
2. When is the matrix D that minimizes the rank of $A + BDC$ unique (§3.1.2)?
3. When is the term BDC that minimizes rank$(A + BDC)$ unique? It will be shown how this corresponds to Mitra's extension of the shorted operator [20] in §3.1.3.
4. In the case of nonuniqueness, what is the minimum norm solution (for unitarily invariant norms) D that minimizes rank$(A + BDC)$ (§3.1.4)?
5. The reverse question is the following: Assume that $\|D\| \leq \delta$ where δ is a given positive real scalar. What is the minimum rank of $A + BDC$? This can be linked to rank minimization problems in so-called matrix balls (§3.1.5).
6. An extreme case occurs if we look for the (minimum norm) solution D to the linear matrix equation $BDC = A$. The RSVD provides the necessary and sufficient conditions for consistency and allows us to parameterize all solutions (§3.1.6).

3.1.1. The range of ranks of A+BDC. The range of ranks of $A + BDC$ for all possible matrices D is described in the following theorem.

THEOREM 3 (on the rank of $A + BDC$).

$$r_{ab} + r_{ac} - r_{abc} \leq \text{rank}(A + BDC) \leq \min(r_{ab}, r_{ac}).$$

For every number r in between these bounds, there exists a matrix D such that rank$(A + BDC) = r$.

Proof. The proof uses the RSVD structure of Theorem 1:

$$\begin{aligned} A + BDC &= P^{-*}S_aQ^{-1} + P^{-*}S_bV_b^*DU_cS_cQ^{-1} \\ &= P^{-*}(S_a + S_bES_c)Q^{-1}, \end{aligned}$$

where $E = V_b^*DU_c$. Because of the nonsingularity of P, Q, U_c, V_b, we have that rank$(A + BDC) = $ rank$(S_a + S_bES_c)$. Using elementary row and column operations and the block partitioning of E as in (4), it is easy to show that

$$\text{(5)} \qquad \text{rank}(A + BDC) = \text{rank} \begin{pmatrix} S_1 + E_{11} & 0 & 0 & 0 & E_{14}S_3 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ S_2E_{41} & 0 & 0 & 0 & S_2E_{44}S_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

the block dimensions of which are the same as those of S_a in Theorem 1. Obviously, a lower bound is achieved for $E_{11} = -S_1$, $E_{14} = 0$, $E_{41} = 0$, $E_{44} = 0$. The upper bound is achieved for almost every ("random") choice of $E_{11}, E_{14}, E_{41}, E_{44}$. □

Observe that, if $r_a = r_{ab} + r_{ac} - r_{abc}$, then there is no S_1 block in S_a and the minimum rank of $A + BDC$ will be r_a. Also observe that the minimum achievable rank, $r_{ab} + r_{ac} - r_{abc}$, is precisely the number of restricted singular values triplets of the form $(1,0,1)$, $(1,1,0)$, and $(1,0,0)$.

3.1.2. The unique rank minimizing matrix D. When is the matrix D that minimizes the rank of $A+BDC$ unique? The answer is given in the following theorem.

THEOREM 4. *Let D be such that* $\operatorname{rank}(A + BDC) = r_{ab} + r_{ac} - r_{abc}$ *and assume that $r_a > r_{ab} + r_{ac} - r_{abc}$. Then the matrix D that minimizes the rank of $A + BDC$ is unique if and only if:*

1. $r_c = q$,
2. $r_b = p$,
3. $r_{abc} = r_{ab} + r_c = r_{ac} + r_b$.

In the case where these conditions are satisfied, the matrix D is given as

$$D = V_b \begin{pmatrix} -S_1 & 0 \\ 0 & 0 \end{pmatrix} U_c^*.$$

Observe that the expression for the matrix D is nothing more than an OSVD!

Proof. It can be verified from the matrix in (5) that the rank of $A + BDC$ is independent of the block matrices E_{12}, E_{13}, E_{21}, E_{22}, E_{23}, E_{24}, E_{31}, E_{32}, E_{33}, E_{34}, E_{42}, E_{43}. Hence, the rank minimizing matrix D will not be unique, whenever one of the corresponding block dimensions is not zero, in which case it is parameterized by the blocks E_{ij} in

$$D = V_b \begin{pmatrix} -S_1 & E_{12} & E_{13} & 0 \\ E_{21} & E_{22} & E_{23} & E_{24} \\ E_{31} & E_{32} & E_{33} & E_{34} \\ 0 & E_{42} & E_{43} & 0 \end{pmatrix} U_c^*. \tag{6}$$

Setting the expressions for these block dimensions equal to zero results in the necessary conditions. The unique optimal matrix D is then given by $D = V_b E U_c^*$, where

$$E = \begin{array}{c} \\ p + r_a - r_{ab} \\ r_{ab} - r_a \end{array} \begin{array}{c} \begin{array}{cc} q + r_a - r_{ac} & r_{ac} - r_a \end{array} \\ \begin{pmatrix} E_{11} & E_{14} \\ E_{41} & E_{44} \end{pmatrix} \end{array} = \begin{pmatrix} -S_1 & 0 \\ 0 & 0 \end{pmatrix}. \qquad \square$$

3.1.3. On the uniqueness of BDC: The extended shorted operator. A question related to the one of §3.1.2 concerns the uniqueness of the product term BDC that minimizes the rank of $A + BDC$. As a matter of fact, this problem has received much attention in the literature where the term BDC is called the *extended shorted operator* and was introduced in [20]. It is an extension to rectangular matrices, of the shorting of an operator considered by Krein, Anderson, and Trapp only for positive operators (see [20] for references).

DEFINITION 1 (the extended shorted operator[1]). Let A $(m \times n)$, B $(m \times p)$, and C $(q \times n)$ be given matrices. A shorted matrix $\mathcal{S}(A|B, C)$ is any $m \times n$ matrix that satisfies the following conditions:

[1] We have slightly changed the notation that is used in [20].

1.
$$\mathbf{R}(\mathcal{S}(A|B,C)) \subseteq \mathbf{R}(B), \qquad \mathbf{R}(\mathcal{S}(A|B,C)^*) \subseteq \mathbf{R}(C^*).$$

2. If F is an $m \times n$ matrix satisfying $\mathbf{R}(F) \subseteq \mathbf{R}(B)$ and $\mathbf{R}(F^*) \subseteq \mathbf{R}(C^*)$, then
$$\operatorname{rank}(A - F) \geq \operatorname{rank}(A - \mathcal{S}(A|B,C)).$$

Hence, the shorted operator is a matrix whose column space belongs to the column space of B, whose row space belongs to the row space of C, and which minimizes the rank of $A - F$ over all matrices F, satisfying these conditions. From this, it follows that the shorted operator can be written as
$$\mathcal{S}(A|B,C) = BDC$$
for a certain $p \times q$ matrix D. This establishes the direct connection of the concept of extended shorted operator with the RSVD.

The shorted operator is not always unique, as can be seen from the following example. Let
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$
Then, all matrices of the form
$$S = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
minimize the rank of $A - S$, which equals 2, for arbitrary α and β.

Necessary conditions for uniqueness of the shorted operator can be found in a straightforward way from the RSVD.

THEOREM 5 (on the uniqueness of the extended shorted operator). *Let the* RSVD *of the matrix triplet* (A, B, C) *be given as in Theorem* 1. *Then*
$$\mathcal{S}(A|B,C) = P^{-*}\mathcal{S}(S_a|S_b,S_c)Q^{-1}.$$
The extended shorted operator $\mathcal{S}(A|B,C)$ *is unique if and only*

1. $r_{abc} = r_c + r_{ab}$,
2. $r_{abc} = r_b + r_{ac}$,

and is given by
$$\mathcal{S}(A|B,C) = P^{-*} \begin{pmatrix} -S_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} Q^{-1}.$$

Proof. It follows from Theorem 3 that the minimum rank of $A + BDC$ is $r_{ab} + r_{ac} - r_{abc}$, and that in this case $E_{11} = -S_1, E_{14} = 0, E_{41} = 0, E_{44} = 0$. A short computation shows that
$$BDC = P^{-*} \begin{pmatrix} -S_1 & E_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ E_{21} & E_{22} & 0 & 0 & E_{24}S_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & S_2E_{42} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} Q^{-1}.$$

Hence, the matrix BDC is unique if and only if the blocks E_{12}, E_{22}, E_{42}, E_{21}, E_{22}, and E_{24} do not appear in this decomposition. Setting the corresponding block dimensions equal to zero proves the theorem. $\square$

Observe that the conditions for uniqueness of the extended shorted operator BDC are less restrictive than the uniqueness conditions for the matrix D (Theorem 4). As a consequence of Theorem 5, we also obtain a parameterization of all shorted operators in the case where the uniqueness conditions are not satisfied. All possible shorted operators are then parameterized by the matrices E_{12}, E_{21}, E_{22}, E_{24}, E_{42}. Observe that the shorted operator is independent of the matrices E_{13}, E_{23}, E_{31}, E_{32}, E_{33}, E_{34}, E_{43}. The result of Theorem 5, derived via the RSVD, corresponds to Theorem 4.1 and Lemma 5.1 of [20]. Some connections with the generalized Schur complement and statistical applications of the shorted operator can also be found in [20].

3.1.4. The minimum norm solutions D that reduce the rank of A+BDC. Consider the problem of finding the matrix D of minimal (unitarily invariant) norm $\|D\|$ such that:

$$\operatorname{rank}(A + BDC) = r < r_a,$$

where r is a prescribed nonnegative integer.

It follows from Theorem 3 that a necessary condition for a solution to exist is that $r_a > r \geq r_{ab} + r_{ac} - r_{abc}$. Observe that if $r_a = r_{ab} + r_{ac} - r_{abc}$, no solution exists. In this case, there is no diagonal matrix S_1 in S_a of Theorem 1. Assume that the required rank r equals the minimal achievable: $r = r_{ab} + r_{ac} - r_{abc}$. Then, if the conditions of Theorem 4 are satisfied, the optimal D is unique and follows directly from the RSVD. The interesting case occurs whenever the rank minimizing D is not unique. Before examining matrices D that *minimize* the rank of $A+BDC$, note that, whenever $\min(r_{ab}, r_{ac}) - r_a > 0$, there exist many matrices that will *increase* the rank of $A + BDC$. In this case,

$$\inf_{\epsilon}\{\ \epsilon = \|D\| \mid \operatorname{rank}(A + BDC) > r_a\} = 0, \tag{7}$$

which implies that there exist arbitrarily "small" matrices D that will increase the rank.

THEOREM 6. *Consider all matrices D satisfying*

$$r_{ab} + r_{ac} - r_{abc} \leq r = \operatorname{rank}(A + BDC) < r_a$$

where r is a given integer and let $\|.\|$ be any unitarily invariant norm. A matrix D of minimal norm $\|D\|$ is given by

$$D = -V_b \begin{pmatrix} S_1^r & 0 \\ 0 & 0 \end{pmatrix} U_c^*$$

where S_1^r is a singular diagonal matrix

$$S_1^r = \begin{array}{c} \\ r + r_{abc} - r_{ab} - r_{ac} \\ r_a - r \end{array} \begin{array}{c} \begin{array}{cc} r + r_{abc} - r_{ac} - r_{ab} & r_a - r \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0 & S_d \end{pmatrix} \end{array}.$$

S_d contains the $r_a - r$ smallest diagonal elements of S_1.

Proof. From the RSVD of the matrix triplet A, B, C it follows that

$$\begin{aligned} A + BDC &= P^{-*}(S_a + S_b(V_b^* D U_c) S_c) Q^{-1} \\ &= P^{-*}(S_a + S_b E S_c) Q^{-1} \end{aligned}$$

with $\|E\| = \|V_b^* D U_c\| = \|D\|$. The result follows immediately from the partitioning of E as in (4) and from equation (5). □

We could use Theorem 6 to define the *restricted singular values* σ_k as

$$\sigma_k = \inf_{\epsilon}\{\ \epsilon = \sigma_{\max}(D) \mid \operatorname{rank}(A + BDC) = k - 1\ \}$$

where $\sigma_{\max}(.)$ denotes the maximum ordinary singular value. Because the rank of $A + BDC$ cannot be reduced below $r_{ab} + r_{ac} - r_{abc}$, there will be $r_{ab} + r_{ac} - r_{abc}$ infinite restricted singular values. There are $r_a + r_{abc} - r_{ab} - r_{ac}$ finite restricted singular values, corresponding to the diagonal elements of S_1. From (5), it can be seen that the diagonal elements of S_2 and S_3 can be used to increase the rank of $A+BDC$ to $\min(r_{ab}, r_{ac})$. However, from (7) it is obvious that $\min(r_{ac} - r_a, r_{ab} - r_a)$ restricted singular values will be zero.

3.1.5. The reverse problem: Given ||D||, what is the minimal rank of A+BDC? The results of §§3.1.3 and 3.1.4 allow us to obtain in a simple fashion the answer to the reverse question: Assuming we are given a positive real number δ such that $\|D\| \le \delta$, what is the minimum rank $r_{\min}$ of $A + BDC$?

The answer is an immediate consequence of Theorem 6. Note that the optimal matrix D is given as the product of three matrices, which form its OSVD! Hence, $\|D\| = \|S_1^\tau\|$ and the integer $r_{\min}$ can be determined as follows. Let S_i be the $i \times i$ diagonal matrix that contains the i smallest elements of S_1. Then,

$$r_{\min} = r_a - (\max_i \{\text{size}(S_i) \text{ such that } \|S_i\| \le \delta\}). \tag{8}$$

It is interesting to note that expressions of the form $A+BDC$ with restrictions on the norm of D can be related to the notion of *matrix balls*, which show up in the analysis of so-called completion problems [6].

DEFINITION 2 (matrix ball). For given matrices A $(m \times n)$, B $(m \times p)$, and C $(q \times n)$, the closed matrix ball $\mathcal{R}(A|B, C)$ with center A, left semiradius B, and right semiradius C is defined by

$$\mathcal{R}(A|B, C) = \{\ X \mid X = A + BDC \text{ where } \|D\|_2 \le 1\}.$$

Using Theorem 6 and (8), we can find all matrices of least rank within a certain given matrix ball by simply requiring that $\sigma_{\max}(D) \le 1$. The solution is obtained from the appropriate truncation of S_1^τ in Theorem 6. Since the solution of the completion problems investigated in [6] are described in terms of matrix balls, it follows that we can find the minimal rank solution in the matrix ball of all solutions of the completion problems, using the RSVD.

3.1.6. The matrix equation BDC = A. Consider the problem of investigating the consistency of, and, if consistent, finding a (minimum norm) solution to, the linear equation in the unknown matrix D:

$$BDC = A.$$

This equation has an historical significance because it led Penrose to rediscover what is now called the Moore–Penrose pseudo-inverse [21], [24]. Of course, this problem can be viewed as an extreme case of Theorems 3 and 6, with the prescribed integer $r = 0$.

THEOREM 7. *The matrix equation $BDC = A$ in the unknown matrix D is consistent if and only if*

$$r_{ab} = r_b, \qquad r_{ac} = r_c, \qquad r_{abc} = r_b + r_c.$$

All solutions are then given by

$$D = V_b \begin{pmatrix} S_1 & E_{13} & 0 \\ E_{31} & E_{33} & E_{34} \\ 0 & E_{43} & 0 \end{pmatrix} U_c^*$$

and a minimum norm solution corresponds to $E_{13} = 0$, $E_{31} = 0$, $E_{33} = 0$, $E_{34} = 0$, $E_{43} = 0$.

Proof. Let $E = V_b^* D U_c$ and partition E as in (4). The consistency of $BDC = A$ depends on whether the following is satisfied with equality

$$\begin{pmatrix} E_{11} & E_{12} & 0 & 0 & E_{14}S_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ E_{21} & E_{22} & 0 & 0 & E_{24}S_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ S_2E_{41} & S_2E_{42} & 0 & 0 & S_2E_{44}S_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} S_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Comparing the diagonal blocks, the conditions for consistency follow immediately as $r_{abc} = r_{ab} + r_c = r_{ac} + r_b = r_b + r_c$, which implies $r_{ab} = r_b$ and $r_{ac} = r_c$. These conditions express the fact that the column space of A should be contained in the column space of B and that the row space of A should be contained in the row space of C. If these conditions are satisfied, the matrix equation $BDC = A$ is consistent and the matrix $E = V_b^* D U_c$ is given by

$$E = \begin{array}{c} \\ r_a \\ p - r_b \\ r_b - r_a \end{array} \begin{array}{c} \begin{array}{ccc} r_a & q - r_c & r_c - r_a \end{array} \\ \begin{pmatrix} E_{11} & E_{13} & E_{14} \\ E_{31} & E_{33} & E_{34} \\ E_{41} & E_{43} & E_{44} \end{pmatrix} \end{array}.$$

The equation $BDC = A$ is equivalent to

$$\begin{pmatrix} E_{11} & E_{14}S_3 & 0 \\ S_2E_{41} & S_2E_{44}S_3 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} S_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This is solved for $E_{11} = S_1$, $E_{14} = 0$, $E_{41} = 0$, $E_{44} = 0$. Observe that the solution is independent of the blocks E_{13}, E_{31}, E_{33}, E_{34}, E_{43}. Hence, all solutions can be parameterized as

$$D = \begin{pmatrix} V_{b1} & V_{b3} & V_{b4} \end{pmatrix} \begin{pmatrix} S_1 & E_{13} & 0 \\ E_{31} & E_{33} & E_{34} \\ 0 & E_{43} & 0 \end{pmatrix} \begin{pmatrix} U_{c1}^* \\ U_{c3}^* \\ U_{c4}^* \end{pmatrix}.$$

The minimum norm solution follows immediately. $\square$

Penrose originally proved [21], [24] that a necessary and sufficient condition for $BDC = A$ to have a solution is:

$$BB^{-}AC^{-}C = A, \tag{9}$$

where B^{-} and C^{-} are inner inverses of B and C. All solutions D can then be written as

$$D = B^{-}AC^{-} + Z - BB^{-}ZC^{-}C, \tag{10}$$

where Z is an arbitrary $p \times q$ matrix. It requires a tedious though straightforward calculation to verify that our solution of Theorem 7 coincides with (10). In order to verify this, consider the RSVD of A, B, C and use Lemma 1 to obtain an expression for the inner inverses of B and C, which will contain arbitrary matrices. Using the block dimensions of S_a, S_b, S_c as in Theorem 1, it can be shown that the consistency conditions of Theorem 7 coincide with the consistency condition (9).

Before concluding this section, it is worth mentioning that all results of this section can be specialized for the case where either B or C equals the identity matrix. In this case, the RSVD specializes to the QSVD (Theorem 2) and mutatis mutandis, the same type of questions, now related to two matrices, can be formulated and solved using the QSVD such as shorted operators, minimum norm rank minimization, solution of the matrix equation $DC = A$, etc.

3.2. On low rank approximations of a partitioned matrix. In this section, the RSVD will be used to analyse and solve problems that can be stated in terms of the matrix[2] $M = \begin{pmatrix} A & B \\ D & D^* \end{pmatrix}$ where A, B, C, D are given matrices. The main results include the analysis of the (generalized) Schur complement [4] in terms of the RSVD (§3.2.1), the range of ranks of the matrix M as D is modified, and the analysis of the (non)unique matrix D that minimizes the rank of M (§3.2.2), and finally the solution of the constrained total least squares problem with exact and noisy data by imposing additional norm constraints on D (§3.2.3).

3.2.1. (Generalized) Schur complements and the RSVD. The notion of a Schur complement S of the matrix A in M (which is $S = D^* - CA^{-1}B$ when A is square nonsingular), can be generalized to the case where the matrix A is rectangular and/or rank deficient [4] as follows.

DEFINITION 3 ((Generalized) Schur complement). A generalized Schur complement of A in $M = \begin{pmatrix} A & B \\ C & D^* \end{pmatrix}$ is any matrix $S = D^* - CA^{-}B$ where A^{-} is an inner inverse of A.

In general, there are many generalized Schur complements, because from Lemma 1 we know that there are many inner inverses. However, the RSVD allows us to investigate the dependency of S on the choice of the inner inverse.

THEOREM 8. *The Schur complement $S = D^* - CA^{-}B$ is independent of A^{-} if and only if $r_a = r_{ab} = r_{ac}$. In this case, S is given by*

$$S = U_c \begin{pmatrix} E_{11}^* - S_1^{-1} & E_{21}^* & E_{31}^* \\ E_{12}^* & E_{22}^* & E_{32}^* \\ E_{13}^* & E_{23}^* & E_{33}^* \end{pmatrix} V_b^*.$$

[2] In order to keep the notation consistent with that of §3.1, we use the matrix D^*, which is the complex conjugate transpose of D in §3.1, as the lower right block of M. This allows us, for instance, to use the same matrix E as defined in (3) and (4).

Proof. Consider the factorization of A as in the RSVD. From Lemma 1, every inner inverse of A can be written as

$$A^- = Q \begin{pmatrix} S_1^{-1} & 0 & 0 & 0 & X_{15} & X_{16} \\ 0 & I & 0 & 0 & X_{25} & X_{26} \\ 0 & 0 & I & 0 & X_{35} & X_{36} \\ 0 & 0 & 0 & I & X_{45} & X_{46} \\ X_{51} & X_{52} & X_{53} & X_{54} & X_{55} & X_{56} \\ X_{61} & X_{62} & X_{63} & X_{64} & X_{65} & X_{66} \end{pmatrix} P^*$$

for certain block matrices X_{ij}, where the block dimensions correspond to the block dimensions of the matrix S_a^* of Theorem 1. It is straightforward to show that

$$CA^-B = U_c \begin{pmatrix} S_1^{-1} & 0 & 0 & X_{15}S_2 \\ 0 & 0 & 0 & X_{25}S_2 \\ 0 & 0 & 0 & 0 \\ S_3X_{51} & S_3X_{53} & 0 & S_3X_{55}S_2 \end{pmatrix} V_b^*.$$

Hence, this product is dependent on the blocks X_{15}, X_{25}, X_{51}, X_{53}, X_{55}. The corresponding block dimensions are 0 if and only if $r_a = r_{ab} = r_{ac}$. □

Observe that the theorem is equivalent with the statement that the (generalized) Schur complement $S = D^* - CA^-B$ is independent of the precise choice of A^- if and only if $\mathbf{R}(B) \subset \mathbf{R}(A)$ and $\mathbf{R}(C^*) \subset \mathbf{R}(A^*)$. This corresponds to Carlson's statement of the same result (Proposition 1 of [4]). In the case that these conditions are not satisfied, all possible generalized Schur complements are parameterized by the blocks X_{51}, X_{53}, X_{15}, X_{25}, and X_{55} as

$$S = U_c \begin{pmatrix} E_{11}^* - S_1^{-1} & E_{21}^* & E_{31}^* & E_{41}^* - X_{15}S_2 \\ E_{12}^* & E_{22}^* & E_{32}^* & E_{42}^* - X_{25}S_2 \\ E_{13}^* & E_{23}^* & E_{33}^* & E_{43}^* \\ E_{14}^* - S_3X_{51} & E_{24}^* - S_3X_{53} & E_{34}^* & E_{44}^* - S_3X_{55}S_2 \end{pmatrix} V_b^*. \tag{11}$$

3.2.2. How does the rank of M change with changing D? Define the matrix $M(\tilde{D}) = \left(\begin{smallmatrix} A & B \\ C & D^* - \tilde{D}^* \end{smallmatrix}\right)$. We shall also use $\bar{D} = D - \tilde{D}$. How can we modify the rank of $M(\tilde{D})$ by changing the matrix $\tilde{D}$? Before answering this question, we need to state the following (well-known) lemma.

LEMMA 3 (rank of a partitioned matrix and the Schur complement). *If A is square and nonsingular, then*

$$\operatorname{rank}\begin{pmatrix} A & B \\ C & D^* \end{pmatrix} = \operatorname{rank}(A) + \operatorname{rank}(D^* - CA^{-1}B).$$

Proof. Observe that:

$$\begin{pmatrix} A & B \\ C & D^* \end{pmatrix} = \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & D^* - CA^{-1}B \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}. \qquad \square$$

Thus we have Theorem 9.

THEOREM 9.

$$\operatorname{rank}\begin{pmatrix} A & B \\ C & D^* \end{pmatrix} = r_{ab} + r_{ac} - r_a + \operatorname{rank}\begin{pmatrix} E_{11}^* - S_1^{-1} & E_{21}^* & E_{31}^* \\ E_{12}^* & E_{22}^* & E_{32}^* \\ E_{13}^* & E_{23}^* & E_{33}^* \end{pmatrix}.$$

Proof. From the RSVD, it follows immediately that the required rank is equal to the rank of the matrix

$$\begin{pmatrix} S_a & S_b \\ S_c & E^* \end{pmatrix} = \left(\begin{array}{cccccc|cccc} S_1 & 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline I & 0 & 0 & 0 & 0 & 0 & E_{11}^* & E_{21}^* & E_{31}^* & E_{41}^* \\ 0 & I & 0 & 0 & 0 & 0 & E_{12}^* & E_{22}^* & E_{32}^* & E_{42}^* \\ 0 & 0 & 0 & 0 & 0 & 0 & E_{13}^* & E_{23}^* & E_{33}^* & E_{43}^* \\ 0 & 0 & 0 & 0 & S_3 & 0 & E_{14}^* & E_{24}^* & E_{34}^* & E_{44}^* \end{array}\right).$$

From the nonsingularity of S_2 and S_3, it follows that the rank is independent of E_{41}, E_{42}, E_{43}, $E_{14}, E_{24}, E_{34}, E_{44}$. The result then follows immediately from Lemma 3, taking into account the block dimensions of the matrices. □

A consequence of Theorem 9 is the following result.

COROLLARY 1. *The range of ranks r of M attainable by an appropriate choice of $\tilde{D}$ in $M = \left(\begin{smallmatrix} A & B \\ C & D^*-\tilde{D}^* \end{smallmatrix}\right)$ is*

$$r_{ab} + r_{ac} - r_a \le r \le \min(p + r_{ac}, q + r_{ab}).$$

The minimum is attained for

$$\tilde{D}^* = U_c \begin{pmatrix} E_{11}^* - S_1^{-1} & E_{21}^* & E_{31}^* & \tilde{E}_{41}^* \\ E_{12}^* & E_{22}^* & E_{32}^* & \tilde{E}_{42}^* \\ E_{13}^* & E_{23}^* & E_{33}^* & \tilde{E}_{43}^* \\ \tilde{E}_{14}^* & \tilde{E}_{24}^* & \tilde{E}_{34}^* & \tilde{E}_{44}^* \end{pmatrix} V_b^* \tag{12}$$

where the matrices $\tilde{E}_{14}$, $\tilde{E}_{24}$, $\tilde{E}_{34}$, $\tilde{E}_{41}$, $\tilde{E}_{42}$, $\tilde{E}_{43}$, and $\tilde{E}_{44}$ are arbitrary matrices.

Compare the expression of $\tilde{D}$ of Corollary 1 with the expression for the generalized Schur complement of A in M, as given by (11). Obviously, the set of matrices $\tilde{D}$ contains all generalized Schur complements, which are those matrices $\tilde{D}$ for which $\tilde{E}_{34} = E_{34}$ and $\tilde{E}_{43} = E_{43}$. If these blocks are not present in E, there are no matrices $\tilde{D}$, other than generalized Schur complements, that minimize the rank of M. Hence, we have proved the following theorem.

THEOREM 10. *The rank of $M(\tilde{D})$ is minimized for $\tilde{D}$ equal to a generalized Schur complement of A in M. The rank of $M(\tilde{D})$ is minimized* **only** *for $\tilde{D} = D^* - CA^-B$ where A^- is an inner inverse of A, if and only if $r_{ab} = r_a$ or $r_c = q$* **and** *$r_{ac} = r_c$ or $r_b = p$. If $r_a = r_{ab} = r_{ac}$, then the minimizing $\tilde{D}$ is unique.*

Proof. The fact that each generalized Schur complement minimizes the rank of $M(\tilde{D})$ follows directly from the comparison of $\tilde{D}$ in Corollary 2 with the expression for the generalized Schur complement in (11). The rank conditions follow simply from setting the block dimensions of E_{34} and E_{43} in (4) equal to 0. The condition for uniqueness of $\tilde{D}$ follows from Theorem 8. □

This theorem can also be found as Theorem 3 of [4], where it is proved via a different approach. Related results can be found in [7] and [31].

3.2.3. Total linear least squares with exact rows and columns. The nomenclature *total linear least squares* was introduced in [16]. The technique is an extension of least squares fitting in the case where there are errors in both the observation vector b and the data matrix A for overdetermined sets of linear equations

$Ax \approx b$. The analysis and solution is given completely in terms of the OSVD of the concatenated matrix $(A \;\; b)$. In the case where some of the columns of A are noise-free while the others contain errors, a mixed least squares–total least squares strategy was developed in [17]. The problem where some rows are also error-free was analysed via a Schur complement-based approach in [7]. One of the key canonical decompositions (Lemma 2 of [7]) and related results concerning rank minimization were described earlier in [4]. Another useful reference is [31]. We shall now show how the RSVD allows us to treat the general situation in an elegant way. Again, let the data matrix be given as $M = \left(\begin{smallmatrix} A & B \\ C & D^* \end{smallmatrix}\right)$ where A, B, C are free of error and only D is contaminated by noise. It is assumed that the data matrix is of full row rank.

The *constrained total linear least squares problem* is the following.

Find the matrix $\hat{D}$ and the nonzero vector x such that

$$\begin{pmatrix} A & B \\ C & \hat{D}^* \end{pmatrix} x = 0,$$

and $\|D - \hat{D}\|_F$ is minimized.

A slightly more general problem is the following.

Find the matrix $\hat{D}$ such that $\|D - \hat{D}\|_F$ is minimal and

$$\text{rank} \begin{pmatrix} A & B \\ C & \hat{D}^* \end{pmatrix} \leq r. \tag{13}$$

The error matrix $D - \hat{D}$ will be denoted by $\tilde{D}$. Assume that a solution x is found. By partitioning x conformally to the dimensions of A and B, we find that the vector x satisfies

$$\begin{aligned} Ax_1 + Bx_2 &= 0, \\ Cx_1 + \hat{D}^* x_2 &= 0. \end{aligned}$$

Hence, the total least squares problem can be interpreted as follows. The rows of A and B correspond to linear constraints on the solution vector x. The columns of the matrix C contain error-free (noiseless) data while those of the matrix D are corrupted by noise. In order to find a solution, we must modify the matrix D with minimum effort, as measured by the Frobenius norm of the "error matrix" $\tilde{D}$, into the matrix $\hat{D}$. Without the constraints imposed by matrices A and B, the problem reduces to a mixed linear–total linear least squares problem, as is analysed and solved in [17].

From the results in §3.2.2, we already know that a necessary condition for a solution to exist is $r \geq r_{ab} + r_{ac} - r_a$ (Corollary 1). The class of rank minimizing matrices $\tilde{D}$ is described by Corollary 1 when $r = r_{ab} + r_{ac} - r_a$. Theorem 9 shows how the generalized Schur complements of A in M form a subset of this set. From Corollary 1, it is straightforward to find the minimum norm matrix $\tilde{D}$ that reduces the rank of $M(\tilde{D})$ to $r = r_{ab} + r_{ac} - r_a$. It is given by

$$\tilde{D}^* = U_c \begin{pmatrix} E_{11}^* - S_1^{-1} & E_{21}^* & E_{31}^* & 0 \\ E_{12}^* & E_{22}^* & E_{32}^* & 0 \\ E_{13}^* & E_{32}^* & E_{33}^* & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} V_b^*.$$

The *minimum norm generalized Schur complement* that minimizes the rank of M is given by

$$S = U_c \begin{pmatrix} E_{11}^* - S_1^{-1} & E_{21}^* & E_{31}^* & 0 \\ E_{12}^* & E_{22}^* & E_{32}^* & 0 \\ E_{13}^* & E_{23}^* & E_{33}^* & E_{43}^* \\ 0 & 0 & E_{34}^* & 0 \end{pmatrix} V_b^* .$$

This corresponds to a choice of inner inverse in (11) given by $X_{15} = E_{41}^* S_2^{-1}$, $X_{25} = E_{42}^* S_2^{-1}$, $X_{51} = S_3^{-1} E_{14}^*$, $X_{53} = S_3^{-1} E_{24}^*$, $X_{55} = S_3^{-1} E_{44}^* S_2^{-1}$.

We shall now investigate two solution strategies, both of which are based on the RSVD. The first one is an immediate consequence of Theorems 6, but, while elegant and extremely simple, might be considered as suffering from some "overkill." It is a direct application of the insights obtained in analysing the sum $A + BDC$. The second one is less elegant but is more in the line of results reported in [4] and [7]. It exploits the insights obtained from analysing the partitioned matrix $M = \left(\begin{smallmatrix} A & B \\ C & D^* \end{smallmatrix}\right)$.

3.2.3.1. Constrained total linear least squares directly via the RSVD. It is straightforward to show that the constrained total least squares problem can be recast as a minimum norm problem as discussed in Theorem 6 as follows.

Find the matrix $\tilde{D}$ of minimum norm $\|\tilde{D}\|$ such that

$$\operatorname{rank}\left(\begin{pmatrix} A & B \\ C & D^* \end{pmatrix} + \begin{pmatrix} 0_{m\times q} \\ I_q \end{pmatrix} \tilde{D}^* \begin{pmatrix} 0_{p\times n} & I_p \end{pmatrix}\right) \leq r.$$

The solution is an immediate consequence of Theorem 6.

COROLLARY 2. *The solution of the constrained total linear least squares problem follows from the application of Theorem* 6 *to the matrix triplet* A', B', C' *where*

$$A' = \begin{pmatrix} A & B \\ C & D^* \end{pmatrix}, \qquad B' = \begin{pmatrix} 0_{m\times q} \\ I_q \end{pmatrix}, \qquad C' = \begin{pmatrix} 0_{p\times n} & I_p \end{pmatrix}.$$

Hence, all that we need is the RSVD of the matrix triplet (A', B', C') and the truncation of the matrix S_1 as described in Theorem 6. It is interesting to also apply Theorem 3 to the matrix triplet (A', B', C'):

$$\begin{aligned} r_{a'b'} &= \operatorname{rank}\begin{pmatrix} A & B & 0 \\ C & D^* & I_q \end{pmatrix} = r_{ab} + q, \\ r_{a'c'} &= \operatorname{rank}\begin{pmatrix} A & B \\ C & D^* \\ 0 & I_p \end{pmatrix} = r_{ac} + p, \\ r_{a'b'c'} &= \operatorname{rank}\begin{pmatrix} A & B & 0 \\ C & D^* & I_q \\ 0 & I_p & 0 \end{pmatrix} = r_a + p + q. \end{aligned}$$

Hence, from Theorem 3, the minimum achievable rank is $r_{a'b'} + r_{a'c'} - r_{a'b'c'} = r_{ab} + r_{ac} - r_a$, which corresponds precisely to the result from Corollary 1.

As a special case, consider the Golub–Hoffman–Stewart result [17] for the total linear least squares solution of $(A \;\; B)x \approx 0$, where A is noise-free and B is contaminated with errors. Instead of applying the QR-SVD-least squares solution as discussed in [17], we could as well achieve the mixed linear–total linear least squares solution from the following.

Minimize $\|\tilde{B}\|$ such that

$$\operatorname{rank}((A \;\; B) - \tilde{B}(0_{p\times n} \;\; I_p)) \leq r,$$

where r is a prespecified integer. This can be done directly via the QSVD of the matrix pair $((A \ \ B), (0_{p\times n} \ \ I_p))$ and it is not too difficult to provide another proof of the Golub–Hoffman–Stewart result derived in [17], now in terms of the properties of the QSVD.

As a matter of fact, the RSVD of the matrix triplet of Corollary 2 allows us to provide a geometrical proof of constrained total linear least squares, in the line of the Golub–Hoffman–Stewart result, taking into account the structure of the matrices B' and C'. We shall not, however, consider this any further in this paper.

3.2.3.2. Solution via RSVD–OSVD. While the solution to the constrained total least squares problem as presented in Corollary 2 is extremely simple, we might object to it because of the apparent "overkill" in computing the RSVD of the matrix triplet (A', B', C'), where B' and C' have an extremely simple structure (zeros and the identity matrix). It will now be shown that the RSVD, combined with the OSVD, may lead to a computationally simpler solution, which more closely follows the lines of the solution as presented in [7].

Using the RSVD, we find that

$$\begin{pmatrix} A & B \\ C & D^* \end{pmatrix} = \begin{pmatrix} P^{-*} & 0 \\ 0 & U_c \end{pmatrix} \begin{pmatrix} S_a & S_b \\ S_c & U_c^* D^* V_b \end{pmatrix} \begin{pmatrix} Q^{-1} & 0 \\ 0 & V_b^* \end{pmatrix}.$$

Let $E^* = U_c^* D^* V_b$. Since U_c and V_b are unitary matrices, the problem can be restated as follows.

Find $\hat{E}$ such that $\|E - \hat{E}\|_F$ is minimal and

$$\text{rank} \begin{pmatrix} S_a & S_b \\ S_c & \hat{E}^* \end{pmatrix} \leq r.$$

The constrained total least squares problem can now be solved as follows.

THEOREM 11 (RSVD–OSVD solution of constrained total least squares). *Consider the* OSVD

$$\begin{pmatrix} E_{11} - S_1^{-1} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{pmatrix} = \sum_{i=1}^{r_e} u_i^e \sigma_i^e (v_i^e)^*,$$

where r_e is the rank of this matrix. The modification of minimal Frobenius norm follows immediately from the OSVD *of this matrix by truncating its dyadic decomposition after $r - r_{ab} - r_{ac} + r_a$ terms. Let*

$$\hat{E} = \sum_{i=1}^{r - r_{ab} - r_{ac} + r_a} u_i^e \sigma_i^e (v_i^e)^*.$$

Then the optimal $\hat{D}$ is given by

$$\hat{D} = V_b \begin{pmatrix} \hat{E} & 0 \\ 0 & 0 \end{pmatrix} U_c^*.$$

Proof. From Theorem 9, it follows that the rank of $\left(\begin{smallmatrix} A & B \\ C & D^* \end{smallmatrix}\right)$ can be reduced by reducing the rank of the matrix

$$\begin{pmatrix} E_{11} - S_1^{-1} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{pmatrix}.$$

The matrix $\tilde{D}$ is then obtained from (12) by setting the blocks $\tilde{E}_{14}$, $\tilde{E}_{24}$, $\tilde{E}_{34}$, $\tilde{E}_{41}$, $\tilde{E}_{42}$, $\tilde{E}_{43}$, $\tilde{E}_{43}$ to 0 in order to minimize the Frobenius norm and then truncating the OSVD of the matrix above. $\square$

We conclude this section by pointing out that more results as well as algorithms to solve total least squares problems with and without constraints and given covariance matrices, can be found in [7], [28], [29], and [31].

3.3. Generalized Gauss–Markov models with constraints. Consider the problem of minimizing $\|y\|^2 + \|z\|^2 = y^*y + z^*z$ over all vectors x, y, z satisfying

$$b = Ax + By, \qquad z = Cx$$

where A, B, C, b are given.

This formulation is a generalization of the conventional least squares problem where $B = I_m$ and $C = 0$. The formulation above admits singular or ill-conditioned matrices B and C. The problem formulation as presented here could be considered as a "square root" version of the problem as follows.

Find x such that

$$\|b - Ax\|_{W_b} + \|x\|_{W_c}$$

is minimized, where $\|u\|_{W_b} = u^*W_bu$ and W_b and W_c are nonnegative-definite symmetric matrices.

In the case that BB^* is nonsingular, we can put $W_b = (BB^*)^{-1}$ and $W_c = C^*C$. The solution can then be obtained as follows.

Minimize $\|y\|^2 + \|z\|^2$ where

$$y^*y = (b - Ax)^*W_b(b - Ax),$$
$$z^*z = x^*C^*Cx.$$

Setting the derivative with respect to x equal to 0, results in

$$x = (A^*W_bA + W_c)^{-1}A^*W_bb. \tag{14}$$

In the case where $W_b = I_m$ and $C = 0$, (14) reduces to the classical least squares expression. For the more general case, we can see a connection with so-called regularization problems. Consider the case where $C \neq 0$ and $B = I_m$. If the matrix A is ill conditioned (because of so-called collinearities, which are (almost) linear dependencies among the columns of A), the addition of the term C^*C may possibly make the sum better suited for numerical inversion than the original product A^*A, hence stabilizing the solution x.

The matrix B acts as a "static" noise filter: Typically, it is assumed that the vector y is normally distributed with the covariance matrix $E(yy^*)$ being a multiple of the identity. The error vector By for the first equation can only be in a direction which is present in the column space of B. If the observation vector b has some component in a certain direction not present in the column space of B, this component should be considered as error-free. The matrix C represents a weighting on the components of x. It reflects possible a priori information concerning the unknown components of x or may reflect the fact that certain components of x (or linear combinations thereof) are more "likely" or less costly than others. The fact that we try to minimize $y^*y + z^*z$ reflects the intention to explain as much as possible (i.e., min y^*y) in terms of the data (columns of the matrix A), taking into account a priori knowledge of the geometrical distribution of the noise (the weighting W_b). The matrix C reflects the cost per component, expressing the preference (or prejudice?) of the modeller to use more of one variable in explaining the phenomenon than of another. In applications, however, typically the matrix A contains many more rows than columns, which corresponds to the fact that better results are to be expected if there are more equations (measurements) than unknowns. However, the condition that BB^* is nonsingular requires a priori knowledge concerning the statistics of the noise. Because typically this knowledge is rather limited, B will have fewer columns than rows, implying that BB^* is singular and (14) does not apply. In this case, however, the RSVD can be applied. It provides important geometrical information on the sensitivity of the solution. Inserting the RSVD of the matrix triplet (A, B, C), the problem can be rewritten as

$$\begin{aligned}(P^*b) &= S_a(Q^{-1}x) + S_b(V_b^*y),\\ (U_c^*z) &= S_c(Q^{-1}x).\end{aligned}$$

Define $b' = P^*b, x' = Q^{-1}x, y' = V_b^*y, z' = U_c^*z$. Then, with obvious partitionings of b', x', y', z', it follows that

$$\begin{aligned}
b'_1 &= S_1 x'_1 + y'_1, & z'_1 &= x'_1,\\
b'_2 &= x'_2, & z'_2 &= x'_2,\\
b'_3 &= x'_3 + y'_2, & z'_3 &= 0,\\
b'_4 &= x'_4, & z'_4 &= S_3 x'_5,\\
b'_5 &= S_2 y'_4, &&\\
b'_6 &= 0. &&
\end{aligned}$$

Observe that $b'_6 = 0$ is a *consistency* condition. It reflects the fact that b is not allowed to have a component in a direction that is not present in the column space of $(A\ B)$. The components of x'_2 and x'_4 can be estimated without error while the fact that $b'_5 = S_2 y'_4$ could be exploited to estimate the variance of the noise.

Most terms in the object function $y^*y + z^*z$ can now be expressed with the subvectors $x'_i, (i = 1, \cdots, 6)$,

$$\begin{aligned}y^*y + z^*z = {} & b'^*_1 b'_1 + x'^*_1 S_1^2 x'_1 - 2b'^*_1 S_1 x'_1 + b'^*_3 b'_3 + x'^*_3 x'_3 - 2b'^*_3 x'_3\\ & + y'^*_3 y'_3 + b'^*_5 S_2^{-2} b'_5 + x'^*_1 x'_1 + x'^*_5 S_3^2 x'_5 + b'^*_2 b'_2.\end{aligned}$$

The minimum solution follows from differentiation with respect to these vectors and results in

$$\begin{aligned}
x'_1 &= (I + S_1^2)^{-1} S_1 b'_1, & y'_1 &= (I + S_1^2)^{-1} b'_1, & z'_1 &= (I + S_1^2)^{-1} S_1 b'_1,\\
x'_2 &= b'_2, & y'_2 &= 0, & z'_2 &= b'_2,\\
x'_3 &= b'_3, & y'_3 &= 0, & z'_3 &= 0,\\
x'_4 &= b'_4, & y'_4 &= S_2^{-1} b'_5, & z'_4 &= 0,\\
x'_5 &= 0, &&&&\\
x'_6 &= \text{arbitrary}. &&&&
\end{aligned}$$

Statistical properties, such as (un)biasedness and consistency, can be analysed in the same spirit as in [23], where Paige has related the Gauss–Markov model without the z-equation, to the QSVD. Similarly, the RSVD also allows us to analyse the sensitivity of the solution. If, for instance, S_2 is ill conditioned, then the minimum of the object function will tend to be high, whenever b'_5 has strong components among the "weak" singular vectors of S_2, because of the term $b'^*_5 S_2^{-2} b'_5$.

A related problem is the following.

Minimize y^*y subject to $b = Ax + By$ and $Cx = c$ where A, B, C, b, c are given.

This is also a Gauss–Markov linear estimation problem as in [23], but now with constraints. The solution is again straightforward from the RSVD. With $b' = P^*b$, $x' = Q^{-1}x$, $y' = V_b^*y$, $c' = U_c^*c$, and an appropriate partitioning, we find

$$\begin{array}{ll} x'_1 = c'_1, & y'_1 = b'_1 - S_1 c'_1, \\ x'_2 = c'_2 = b'_2, & y'_2 = 0, \\ x'_3 = b'_3, & y'_3 = 0, \\ x'_4 = b'_4, & y'_4 = S_2^{-1} b'_5, \\ x'_5 = S_3^{-1} c'_4, & \\ x'_6 = \text{arbitrary}. & \end{array}$$

Observe that $c'_2 = b'_2$ and $c'_3 = 0$ are two *consistency* conditions.

4. Conclusions and perspectives. In this paper, we have derived a generalization of the OSVD, the *restricted singular value decomposition* (RSVD), which has the OSVD, PSVD, and QSVD as special cases. A constructive proof, based upon a sequence of OSVDs and PSVDs can be found in [9]. We have also analysed in detail its structural and geometrical properties and its relations to generalized eigenvalue problems and canonical correlation analysis. It was shown how the RSVD is a valuable tool in the analysis and solution of rank minimization problems with restrictions. First, we have shown how to study expressions of the form $A+BDC$ and find matrices D of minimum norm that minimize the rank. It was demonstrated how this problem is connected to the concept of shorted operators and matrix balls. Second, we have analysed in detail low rank approximations of a partitioned matrix, when only one of its blocks can be modified. The close relation with generalized Schur complements was discussed and it was shown how the RSVD permits us to solve constrained total linear least squares problems with mixed exact and noisy data. Third, it was demonstrated how the RSVD provides an elegant solution to Gauss–Markov models with constraints. The fact that the RSVD is only the tip of an iceberg of generalizations of the OSVD for 2, 3, 4, $\cdots$ matrices, is fully explored in [11].

Acknowledgments. We would like to thank Hongyuan Zha, Sabine Van Huffel and the referees for their detailed comments and suggestions, which allowed us to improve considerably an earlier version of this paper. We would also like to thank Dan Boley for providing us with an English translation of Beltrami's original paper.

REFERENCES

[1] L. Autonne, *Sur les groupes linéaires, réels et orthogonaux*, Bull. Sci. Math., France, 30 (1902), pp. 121–133.

[2] E. Beltrami, *Sulle funzioni bilineari*, in Giornale di Mathematiche, G. Battagline and E. Fergola, eds., 11 (1873), pp. 98–106.

[3] A. Björck and G. H. Golub, *Numerical methods for computing angles between linear subspaces*, Math. Comp., 27 (1973) pp. 579–594.

[4] D. Carlson, *What are Schur complements, anyway?*, Linear Algebra Appl., 74 (1986), pp. 257–275.

[5] J. S. Chipman, *Estimation and aggregation in econometrics: An application of the theory of generalized inverses*, in Generalized Inverses and Applications, Academic Press, New York, 1976, pp. 549–769.

[6] C. Davis, W. M. Kahan, and H. F. Weinberger, *Norm-preserving dilations and their application to optimal error bounds*, SIAM J. Numer. Anal., 19 (1982), pp. 445–469.

[7] J. W. Demmel, *The smallest perturbation of a submatrix which lowers the rank and constrained total least squares problems*, SIAM J. Numer. Anal., 24 (1987), pp. 199–206.

[8] B. De Moor and G. H. Golub, *Generalized singular value decompositions: A proposal for a standardized nomenclature*, Numerical Analysis Project Manuscript NA-89-03, Department of Computer Science, Stanford University, Stanford, CA, April 1989; also in ESAT-SISTA Report 1989-10, Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, April 1989.

[9] ———, *The restricted singular value decomposition: Properties and applications*, Numerical Analysis Project Manuscript NA-89-04, Department of Computer Science, Stanford University, Stanford, CA, April 1989; also in ESAT-SISTA Report 1989-09, Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, April 1989.

[10] B. De Moor, *On the structure and geometry of the product singular value decomposition*, Numerical Analysis Project Manuscript NA-89-05, Department of Computer Science, Stanford University, Stanford, CA, May 1989; also in ESAT-SISTA Report 1989-12, Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, May 1989.

[11] B. De Moor and H. Zha, *A tree of generalizations of the ordinary singular value decomposition*, ESAT-SISTA Report 1989-21 and revised version ESAT-SISTA Report 1990-11, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium; also in the special issue on Matrix Canonical Forms, Linear Algebra Appl., to appear.

[12] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.

[13] L. M. Ewerbring and F. T. Luk, *Canonical correlations and generalized* SVD: *Applications and new algorithms*, Proc. SPIE, Vol. 977, Real Time Signal Processing XI, paper 23, 1988.

[14] K. V. Fernando and S. J. Hammarling, *A product induced singular value decomposition for two matrices and balanced realisation*, in Linear Algebra in Signal Systems and Control, B. N. Datta, C. R. Johnson, M. A. Kaashoek, R. Plemmons, and E. Sontag, eds., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1988, pp. 128–140.

[15] G. H. Golub and C. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1983.

[16] ———, *An analysis of the total least squares problem*, SIAM J. Numer. Anal., 17 (1980), pp. 883–893.

[17] G. H. Golub, A. Hoffman, and G. W. Stewart, *A generalization of the Eckart–Young–Mirsky matrix approximation theorem*, Linear Algebra Appl., 88/89 (1987), pp. 317–327.

[18] W. E. Larimore, *Identification of nonlinear systems using canonical variate analysis*, in Proc. 26th Conference on Decision and Control, Los Angeles, CA, December 1987.

[19] C. Lawson and R. Hanson, *Solving Least Squares Problems*, Prentice–Hall, Englewood Cliffs, NJ, 1974.

[20] S. K. Mitra and M. L. Puri, *Shorted matrices—An extended concept and some applications*, Linear Algebra Appl., 42 (1982), pp. 57–79.

[21] M. Z. Nashed, ed., *Generalized Inverses and Applications*, Academic Press, New York, 1976.

[22] C. C. Paige and M. A. Saunders, *Towards a generalized singular value decomposition*, SIAM J. Numer. Anal., 18 (1981), pp. 398–405.

[23] C. C. Paige, *The general linear model and the generalized singular value decomposition,* Linear Algebra Appl., 70 (1985), pp. 269–284.

[24] R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406–413.

[25] ———, *On best approximate solutions of linear matrix equations*, Proc. Cambridge Philos. Soc., 52 (1956), pp. 17–19.

[26] J. J. Sylvester *Sur la réduction biorthogonale d'une forme linéo-linéaire à sa forme canonique*, Comptes Rendus, CVIII, 1889, pp. 651–653.

[27] J. Vandewalle and B. De Moor, *A variety of applications of the singular value decomposition,* in SVD and Signal Processing: Algorithms, Applications and Architectures, E. Deprettere, ed., North-Holland, Amsterdam, 1988, pp. 43–91.

[28] S. Van Huffel and J. Vandewalle, *Analysis and properties of the generalized total least squares problem $Ax = B$ when somme or all columns in A are subject to error*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 294–315.

[29] S. Van Huffel and H. Zha, *Restricted total least squares: A unified approach for solving (generalized) (total) least squares problems with(out) equality constraints*, ESAT-SISTA Report 1989-05, Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, March 1989.

[30] C. F. Van Loan, *Generalizing the singular value decomposition*, SIAM J. Numer. Anal., 13 (1976), pp. 76–83.

[31] G. A. WATSON, *The smallest perturbation of a submatrix which lowers the rank of the matrix*, IMA J. Numer. Anal., 8 (1988), pp. 295–303.

[32] H. ZHA, *Restricted* SVD *for matrix triplets and rank determination of matrices*, Scientific Report 89-2, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, Germany, 1989.

[33] ———, *A numerical algorithm for computing the* RSVD *for matrix triplets*, Scientific Report 89-1, Konrad-Zuse-Zentrum für Informationstechnik, Berlin, Germany, 1989.