Connections Between Quadratic Transform for Fractional Programming and Schur Complement

Kaiming Shen[1], Kareem M. Attiah[2], Yannan Chen[3], and Wei Yu[2]
[1]School of Science and Engineering, The Chinese University of Hong Kong (Shenzhen), China
[2]Department of Electrical and Computer Engineering, University of Toronto, ON, Canada
[3]State Key Laboratory of Radio Frequency Heterogeneous Integration, Shenzhen University, China
E-mails: shenkaiming@cuhk.edu.cn, kattiah@ece.utoronto.ca, chenyannan@szu.edu.cn, weiyu@comm.utoronto.ca

***Abstract*—This paper shows that there are intimate connections between the quadratic transform technique for solving fractional programming (FP) problems and the Schur complement technique in matrix analysis. We show that the quadratic transform technique is related to two aspects of the Schur complement: (i) the linear matrix inequality (LMI) condition for positive semidefiniteness and (ii) the matrix determinant formula. In terms of the LMI condition, we establish that the quadratic transform and the Schur complement condition imply each other. This connection leads to generalizations of the quadratic transform in FP and the Schur complement LMI. Furthermore, the above connection allows us to provide new interpretations of the auxiliary variable in the quadratic transform and allows us to rederive the Schur complement determinant formula.**

I. Introduction

This paper explores a curious connection between two seemingly unrelated topics—quadratic transform for solving fractional programming (FP) problems [1] and the Schur complement in matrix analysis [2], both of which have been extensively used in the communication system design. Specifically, we show that the quadratic transform for FP can be obtained from the Schur complement, and vice versa. Furthermore, this connection leads to new interpretations and generalizations of results in FP and Schur complement.

FP is a class of mathematical optimizations that involve ratio terms. It plays a key role in communications and signal processing, because many key performance metrics are fractionally structured, e.g., the signal-to-interference-plus-noise ratio (SINR). While the classical FP methods (including Dinkelbach's algorithm [3]) are typically limited to the single-scalar-ratio case, the recent development of the so-called quadratic transform for FP [4] is able to handle a broader range of problems with multiple ratios and with matrix variables, as motivated by optimization problems involving multiple-input multiple-output (MIMO) networks.

The following theorem states the general result on the quadratic transform for FP. This result is useful for developing algorithms for optimizing ratios, because it decouples the numerator and denominator of each fractional term.

Theorem 1 (Quadratic Transform for FP [1]): Consider k pairs of matrix functions $A_i : \mathcal{X} \to \mathbb{R}^{n\times m}$ and $B_i : \mathcal{X} \to \mathbb{S}^n_{++}$, for $i = 1,\ldots,k$, where $\mathcal{X}$ is a nonempty constraint set, and $\mathbb{S}^n_{++}$ denotes the set of $n \times n$ positive definite matrices. The sum-of-traces-of-matrix-ratio FP problem

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{k} \mathrm{Tr}(A_i^T(x)B_i^{-1}(x)A_i(x)) \tag{1}$$

is equivalent to

$$\begin{aligned}\underset{x\in\mathcal{X},Y_1,\ldots,Y_k}{\text{maximize}} \quad & \sum_{i=1}^{k} \mathrm{Tr}(Y_i^T A_i(x) + A_i^T(x)Y_i - Y_i^T B_i(x)Y_i)\\ \text{subject to} \quad & Y_i \in \mathbb{R}^{n\times m},\ \text{for } i=1,\ldots,k \end{aligned} \tag{2}$$

in the sense that the two problems have the same solution for x and their optimal objective values are equal.

The proof of the above theorem is based on explicitly finding the optimal Y_i by completing-the-square, then substituting the optimal Y_i back into the objective function [4]. This paper shows that there is an alternative way of deriving the same transform based on the Schur complement in matrix analysis. In the rest of the paper, we drop the arguments in the matrix functions, e.g., $A_i(x)$ is written as A_i, whenever doing so causes no confusion.

The notion of Schur complement in linear algebra has a long history [5]. Consider a square matrix

$$M = \begin{bmatrix} C & A^T \\ A & B \end{bmatrix} \tag{3}$$

with the blocks $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{S}^n_{++}$, and $C \in \mathbb{S}^m_{+}$, where $\mathbb{S}^m_{+}$ denotes the set of $m \times m$ positive semidefinite matrices. The square matrix $C - A^T B^{-1} A$ is referred to as the *Schur complement* of the block C of the matrix M.

A classic result concerning the Schur complement is a linear matrix inequality (LMI) condition for the positive semidefiniteness of $C - A^T B^{-1} A$. A proof can be found in [6].

Theorem 2 (Schur Complement and LMI): When $B \succ 0$, the following two conditions are equivalent:

$$M \succeq 0 \quad \Leftrightarrow \quad C - A^T B^{-1} A \succeq 0. \tag{4}$$

Thus, the positive semidefiniteness of the Schur complement can be converted to an LMI in M, which is potentially easier to deal with for constrained optimization. There is also a second important result on the Schur complement concerning

The work of Kaiming Shen and Yannan Chen was supported by the NSFC under Grant 12426306. The work of Kareem Attiah and Wei Yu was supported by an NSERC Discovery Grant RGPIN-2023-04697.

the matrix determinant, which will be discussed in Section V.

The Schur complement has seen fruitful applications in information theory, e.g., in Wyner-Ziv coding [7], the relation between the mutual information and the minimum mean-squared error (MMSE) [8], the dirty-paper coding [9], and the uplink-downlink duality [10]. Furthermore, quantum information theory relies heavily on linear algebra; the Schur complement is a useful mathematical tool in this domain [11].

Recently, both FP and Schur complement have been used in the algorithm design for joint sensing and communications [12]–[18], indicating there may be a connection between the two, but such connection has not been properly explored. The present paper aims to develop a unified theory between the two. Our main results are three-fold:

- We show that Theorem 1 implies Theorem 2. Further, by extending this connection to the multiple-ratio case, we obtain a generalization of the LMI condition in (4).
- We show that Theorem 2 implies Theorem 1. This allows us to devise a generalized quadratic transform that accounts for the generalized inverse.
- We show that the auxiliary variable in quadratic transform can be reinterpreted in terms of an MMSE estimation problem, and use this connection to rederive the Schur complement determinant formula.

In the rest of the paper, we denote by $(\cdot)^+$ the generalized inverse, I the identity matrix, 0 the zero matrix, and $h(\cdot)$ the differential entropy.

II. From Quadratic Transform to Schur Complement

We start by rederiving the equivalence relation in (4) from Theorem 1. Let $m=1$ in Theorem 1, so each $A_i : \mathcal{X} \to \mathbb{R}^{n\times m}$ is just a vector function $a_i : \mathcal{X} \to \mathbb{R}^n$. Theorem 1 states that the sum-of-ratios FP problem

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{k} a_i^T B_i^{-1} a_i \tag{5}$$

is equivalent to

$$\underset{x\in\mathcal{X},\ y_1,\ldots,y_k\in\mathbb{R}^n}{\text{maximize}} \quad \sum_{i=1}^{k}(y_i^T a_i + a_i^T y_i - y_i^T B_i y_i). \tag{6}$$

We can now show the following chain of equivalences:

$$\begin{aligned}
&C - A^T B^{-1} A \succeq 0\\
&\overset{(a)}{\Leftrightarrow} q^T(C - A^T B^{-1} A) q \geq 0,\ \forall q \in \mathbb{R}^m\\
&\overset{(b)}{\Leftrightarrow} q^T C q - \sup_{y\in\mathbb{R}^n}\{y^T A q + (Aq)^T y - y^T B y\} \geq 0,\ \forall q\\
&\Leftrightarrow q^T C q - y^T A q - (Aq)^T y + y^T B y \geq 0,\ \forall q,\ \forall y\\
&\Leftrightarrow \begin{bmatrix} q^T & -y^T \end{bmatrix} \begin{bmatrix} C & A^T \\ A & B \end{bmatrix} \begin{bmatrix} q \\ -y \end{bmatrix} \geq 0,\ \forall q,\ \forall y\\
&\Leftrightarrow u^T \begin{bmatrix} C & A^T \\ A & B \end{bmatrix} u \geq 0,\ \forall u = \begin{bmatrix} q \\ -y \end{bmatrix} \in \mathbb{R}^{m+n}\\
&\overset{(c)}{\Leftrightarrow}\ M \succeq 0,
\end{aligned} \tag{7}$$

where (a) follows from the definition of positive semidefiniteness, (b) follows from Theorem 1 with $m=1$ as discussed above and $k=1$, and (c) follows from the definition of positive semidefiniteness. Thus, Theorem 2 follows from Theorem 1.

It is interesting to note that only the single-ratio FP (with $k=1$) is needed to prove (4). It turns out that if we utilize the multiple-ratio result of Theorem 1, the above connection between FP and the Schur complement would imply the following more general chain of equivalences:

$$\begin{aligned}
&C - \sum_{i=1}^{k} A_i^T B_i^{-1} A_i \succeq 0\\
&\Leftrightarrow q^T\left(C - \sum_{i=1}^{k} A_i^T B_i^{-1} A_i\right) q \geq 0,\ \forall q \in \mathbb{R}^m\\
&\Leftrightarrow q^T C q - \sum_{i=1}^{k} \sup_{y_i}\{y_i^T A_i q + (A_i q)^T y_i - y_i^T B_i y_i\} \geq 0,\ \forall q\\
&\Leftrightarrow q^T C q - \sum_{i=1}^{k} (y_i^T A_i q + (A_i q)^T y_i - y_i^T B_i y_i) \geq 0,\ \forall q, y_i\\
&\Leftrightarrow \begin{bmatrix} q \\ -y_1 \\ -y_2 \\ \vdots \\ -y_k \end{bmatrix}^T \begin{bmatrix} C & A_1^T & A_2^T & \cdots & A_k^T \\ A_1 & B_1 & 0 & \cdots & 0 \\ A_2 & 0 & B_2 & & \\ \vdots & \vdots & & \ddots & \\ A_k & 0 & & & B_k \end{bmatrix} \begin{bmatrix} q \\ -y_1 \\ -y_2 \\ \vdots \\ -y_k \end{bmatrix} \geq 0, \forall q, y_i
\end{aligned}$$

where all the blank blocks are filled with zero matrices. The resulting generalization of the Schur complement based LMI is stated in the following theorem.

Theorem 3 (Generalized Schur Complement Based LMI): Consider the matrix

$$M^{(k)} = \begin{bmatrix} C & A_1^T & A_2^T & \cdots & A_k^T \\ A_1 & B_1 & 0 & \cdots & 0 \\ A_2 & 0 & B_2 & & \\ \vdots & \vdots & & \ddots & \\ A_k & 0 & & & B_k \end{bmatrix} \tag{8}$$

with the blocks $A_i \in \mathbb{R}^{n\times m}$, $B_i \in \mathbb{S}_{++}^n$, and $C \in \mathbb{R}^{m\times m}$. The following equivalence relation holds:

$$M^{(k)} \succeq 0 \quad \Leftrightarrow \quad C - \sum_{i=1}^{k} A_i^T B_i^{-1} A_i \succeq 0. \tag{9}$$

Thus, by utilizing the multiple-ratio quadratic transform result in FP, a more general version of the Schur complement LMI relation can be established.

III. From Schur Complement to Quadratic Transform

We now discuss the reverse direction and show that Theorem 1 follows from Theorem 2. To ease notation, we start with the single-ratio case of (1):

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \operatorname{Tr}(A^T B^{-1} A). \tag{10}$$

The above problem can be readily rewritten as

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \inf_{C\in\mathbb{S}_+^m} \text{Tr}(C) \tag{11a}$$
$$\text{subject to} \quad C - A^T B^{-1} A \succeq 0. \tag{11b}$$

The above problem reformulation rewrites the objective in terms of a Schur complement. By Theorem 2, we convert the positive semidefiniteness constraint (11b) to LMI as

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \inf_{C\in\mathbb{S}_+^m} \text{Tr}(C) \tag{12a}$$
$$\text{subject to} \quad M \succeq 0, \tag{12b}$$

where M is defined as in (3). Note that problem (12) consists of two parts: the inner minimization over C for fixed x and the outer maximization over x. We proceed to find the Lagrangian dual of the inner problem.

The Lagrangian function of the inner minimization is

$$L(C,\Lambda) = \text{Tr}(C) - \text{Tr}(\Lambda M), \tag{13}$$

where $\Lambda \in \mathbb{S}_+^{m+n}$ is the Lagrange multiplier. This Λ can be partitioned into four blocks:

$$\Lambda = \begin{bmatrix} W & V^T \\ V & Z \end{bmatrix} \succeq 0, \tag{14}$$

where $W \in \mathbb{S}_+^m$, $Z \in \mathbb{S}_+^n$, and $V \in \mathbb{R}^{n\times m}$. Substituting (14) into $L(C,\Lambda)$, we have

$$\begin{aligned} L(C,\Lambda) &= \text{Tr}(C) - \text{Tr}(WC + V^T A + A^T V) - \text{Tr}(ZB) \\ &= \text{Tr}\big((I_m - W)C - V^T A - A^T V\big) - \text{Tr}(ZB), \end{aligned} \tag{15}$$

where I_m is the $m \times m$ identity matrix.

The dual function is now given by

$$g(\Lambda) = \inf_{C\in\mathbb{S}_+^m} L(C,\Lambda) \tag{16}$$
$$= \begin{cases} -\text{Tr}(V^T A + A^T V) - \text{Tr}(ZB) & \text{if } W = I_m \\ -\infty & \text{otherwise.} \end{cases} \tag{17}$$

To ensure that the value of $g(\Lambda)$ is finite, we must have

$$W = I_m. \tag{18}$$

After plugging $W = I_m$ into Λ in (14) and by using the following chain of equivalence relations:

$$\begin{aligned} \begin{bmatrix} I_m & V^T \\ V & Z \end{bmatrix} \succeq 0 &\overset{(a)}{\Leftrightarrow} \begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix} \begin{bmatrix} I_m & V^T \\ V & Z \end{bmatrix} \begin{bmatrix} 0 & I_m \\ I_n & 0 \end{bmatrix} \succeq 0 \\ &\Leftrightarrow \begin{bmatrix} Z & V \\ V^T & I_m \end{bmatrix} \succeq 0 \\ &\overset{(b)}{\Leftrightarrow} Z - VV^T \succeq 0, \end{aligned} \tag{19}$$

where (a) is due to the matrix similarity and (b) follows from Theorem 2, we arrive at the dual problem of the inner minimization in (12):

$$\underset{V,Z}{\text{maximize}} \quad -\text{Tr}(V^T A + A^T V) - \text{Tr}(ZB) \tag{20a}$$
$$\text{subject to} \quad Z - VV^T \succeq 0. \tag{20b}$$

Because $B \succ 0$, the optimal solution of Z is

$$Z^\star = VV^T. \tag{21}$$

As a consequence, the dual problem reduces to

$$\underset{V\in\mathbb{R}^{n\times m}}{\text{maximize}} \quad -\text{Tr}(V^T A + A^T V + V^T BV). \tag{22}$$

This dual problem is formulated with respect to the inner minimization problem in (12). Because the inner minimization problem is convex with strong duality, it is equivalent to its dual. Replacing the inner minimization by the dual problem, we convert (12) to

$$\underset{x\in\mathcal{X},V}{\text{maximize}} \quad -\text{Tr}(V^T A + A^T V + V^T BV) \tag{23a}$$
$$\text{subject to} \quad V \in \mathbb{R}^{n\times m}. \tag{23b}$$

Further, substituting $V = -Y$ into the above problem yields

$$\underset{x\in\mathcal{X},Y}{\text{maximize}} \quad \text{Tr}(Y^T A + A^T Y - Y^T BY) \tag{24a}$$
$$\text{subject to} \quad Y \in \mathbb{R}^{n\times m}, \tag{24b}$$

which is exactly the single-ratio case of the quadratic transform in Theorem 1.

The above result can be immediately extended to multiple ratios. The main idea is to rewrite (1) as

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \inf_{C_i\in\mathbb{S}_+^m} \left\{ \sum_{i=1}^k \text{Tr}(C_i) \right\} \tag{25a}$$
$$\text{subject to} \quad C_i - A_i^T B_i^{-1} A_i \succeq 0, \text{ for } i = 1,\ldots,k. \tag{25b}$$

After each constraint $C_i - A_i^T B_i^{-1} A_i \succeq 0$ is converted to an LMI, a Lagrange multiplier Λ_i is introduced for each LMI, and can be further determined as

$$\Lambda_i = \begin{bmatrix} I_m & V_i^T \\ V_i & V_i V_i^T \end{bmatrix}. \tag{26}$$

By substituting the resulting dual problem into the inner minimization, we arrive exactly at the multiple-ratio quadratic transform in (2). So Theorem 2 implies Theorem 1.

The key observation here is that the auxiliary variable Y of the quadratic transform in Theorem 1 is exactly the negative of the off-diagonal block V of the Lagrange multiplier Λ.

We remark that the above Lagrangian dual connection was foreshadowed in [17] for the scalar-ratio FP case of (10) with $m = 1$ in the context of an integrated sensing and communication problem. The proof technique of [17] uses the Schur complement, while earlier work [18] for a similar problem setting uses the quadratic transform for FP. This indicates a potential connection between the two (although the problem setting of [17], [18] is for the minimization FP, rather than the maximization.)

We have shown that Theorem 1 can be used to prove Theorem 2 and vice versa, but caution that there are applications of quadratic transform in more complicated optimization problems for which it is difficult to extend the above connection. Examples of such applications include the mixed-max-and-min FP and the log-ratio FP as treated in [1].

IV. Further Extensions

A. Minimization FP

We present extensions of the connection between the quadratic transform for FP and the Schur complement developed in the previous section. First, consider the following FP problem of minimizing the trace of a matrix ratio, as opposed to the earlier maximization case in (10):

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \text{Tr}(A^TB^{-1}A). \tag{27}$$

Again, by introducing an auxiliary variable $C \in \mathbb{S}_+^m$, we formulate an optimization problem with a constraint in the form of a Schur complement:

$$\underset{x\in\mathcal{X},C\in\mathbb{S}_+^m}{\text{minimize}} \quad \text{Tr}(C) \tag{28a}$$

$$\text{subject to} \quad C - A^TB^{-1}A \succeq 0. \tag{28b}$$

However, there is a crucial difference between the minimization FP and the maximization FP. For the minimization FP, after taking the Lagrangian dual of (28) and repeating the same procedure as before, we obtain

$$\underset{Y}{\text{maximize}} \quad \inf_{x\in\mathcal{X}} \text{Tr}(Y^TA + A^TY - Y^TBY) \tag{29a}$$

$$\text{subject to} \quad Y \in \mathbb{R}^{n\times m}. \tag{29b}$$

In contrast, if we directly decouple the matrix ratio $A^TB^{-1}A$ using the quadratic transform of Theorem 1, the minimization FP problem (27) is recast into

$$\underset{x\in\mathcal{X}}{\text{minimize}} \quad \sup_{Y} \text{Tr}(Y^TA + A^TY - Y^TBY) \tag{30a}$$

$$\text{subject to} \quad Y \in \mathbb{R}^{n\times m}. \tag{30b}$$

The two can be the same, but only if the minimization and the maximization can be interchanged. Thus, there is also a connection between the minimization FP and Schur complement just as the maximization FP case, but it requires additional conditions that ensure min-max is equal to max-min, e.g., Sion's minimax condition [19].

B. Generalized Inverse

The connection between the Schur complement and the quadratic transform allows advances in one area to be mapped to the other area. For example, the Schur complement can be extended to the case of the so-called generalized inverse [2]. This means that the quadratic transform technique for solving FP can also be extended for the generalized inverse. In this section, we develop such an extension of Theorem 1.

For any matrix $D \in \mathbb{R}^{m\times n}$, its generalized inverse, denoted by D^+, is an $n \times m$ matrix such that $DD^+D = D$. The Moore–Penrose inverse is a special case of the generalized inverse. The following is a result on the Schur complement for the generalized inverse [2].

Theorem 4 (Schur Complement Based LMI with Generalized Inverse [2]): When $B \succeq 0$ and $(I_n - BB^+)A = 0$, we have

$$M \succeq 0 \quad \Leftrightarrow \quad C - A^TB^+A \succeq 0. \tag{31}$$

This theorem requires a new condition $(I_n - BB^+)A = 0$, which turns out to be the condition needed in order to ensure that the Schur complement $C - A^TB^+A$ is well defined.

Now, let $A : \mathcal{X} \to \mathbb{R}^{n\times m}$ and $B : \mathcal{X} \to \mathbb{S}_+^n$ be a pair of matrix functions, where $B(x)$ is not necessarily invertible. Consider an FP involving generalized inverse, under the condition $(I_n - BB^+)A = 0$, as formulated below:

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \text{Tr}(A^TB^+A). \tag{32}$$

The above problem can be rewritten as

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \inf_{C\in\mathbb{S}_+^m} \text{Tr}(C) \tag{33a}$$

$$\text{subject to} \quad C - A^TB^+A \succeq 0. \tag{33b}$$

By Theorem 4, the constraint can be replaced by $M \succeq 0$. In this case, we can then use the same Lagrangian procedure to find its dual, which would lead to an extension of quadratic transform for generalized inverse. Observe that the replacement of B^{-1} by B^+ does not impact the Lagrangian procedure, so we would arrive at the same reformulation as in (24). The above result can be readily extended to the multiple-ratio case, as stated in the following theorem.

Theorem 5 (Quadratic Transform for FP with Generalized Inverse): Consider k pairs of matrix functions $A_i : \mathcal{X} \to \mathbb{R}^{n\times m}$ and $B_i : \mathcal{X} \to \mathbb{S}_+^n$, for $i = 1, \ldots, k$. In particular, assume that $(I_n - B_iB_i^+)A_i = 0$ for each i. The generalized sum-of-traces-of-matrix-ratio FP problem

$$\underset{x\in\mathcal{X}}{\text{maximize}} \quad \sum_{i=1}^{k} \text{Tr}(A_i^TB_i^+A_i) \tag{34}$$

is equivalent to

$$\underset{x\in\mathcal{X},Y_1,\ldots,Y_k}{\text{maximize}} \quad \sum_{i=1}^{k} \text{Tr}(Y_i^TA_i + A_i^TY_i - Y_i^TB_iY_i) \tag{35a}$$

$$\text{subject to} \quad Y_i \in \mathbb{R}^{n\times m}, \text{ for } i = 1, \ldots, k \tag{35b}$$

in the sense that the two problems have the same solution for x and their optimal objective values are equal.

V. Connection to Schur Complement Determinant Formula

One of the main goals of this paper is to give interpretations to the auxiliary variable Y_i in the quadratic transform for FP in Theorem 1. On the one hand, as seen in Section III, Y_i corresponds to the negative of the off-diagonal block of the Lagrangian dual variable in a Schur complement reformulation of the original FP problem. On the other hand, in Section II the role of Y_i can be seen in the chain of equivalences (7), where for the single-ratio case, y is the vector that optimizes a quadratic form as shown in line (b) of (7). This optimization of the quadratic form can be interpreted as an MMSE estimation. In this section, we show that this connection gives a way of rederiving the Schur complement determinant formula.

Consider an MMSE estimation problem on a joint Gaussian distribution for u:

$$u = \begin{bmatrix} q \\ -y \end{bmatrix} \sim \mathcal{N}(0, J), \text{ where } J = \begin{bmatrix} F & G^T \\ G & H \end{bmatrix} \succ 0 \quad (36)$$

with $F \in \mathbb{S}^m_{++}$, $G \in \mathbb{R}^{n\times m}$, and $H \in \mathbb{S}^n_+$. Treat q as the observation. Consider the problem of estimating y based on q. The MMSE estimator is given by

$$\hat{y} = \mathbb{E}[y|q] = \Sigma_{yq}\Sigma_{qq}^{-1}q = -GF^{-1}q, \quad (37)$$

where Σ_{yq} and Σ_{qq} are the respective covariance matrices.

In the meanwhile, in line (b) of (7), the optimal y is

$$y^\star = B^{-1}Aq. \quad (38)$$

It turns out that if we identify $J = M^{-1}$, then $y^\star$ and $\hat{y}$ coincide, i.e., the optimal auxiliary variable y of the quadratic transform amounts to an MMSE estimator. This is because the $(2,1)$-th block of MJ is $AF+BG=0$ when $J=M^{-1}$, thus we have $B^{-1}A = -GF^{-1}$.

The above interpretation of y in terms of an MMSE estimation problem gives a way of rederiving the Schur complement determinant formula. Assuming $J = M^{-1}$, the entropy of u can be computed as

$$h(u) = \frac{1}{2}\log\Big((2\pi e)^{m+n}\det(M^{-1})\Big). \quad (39)$$

Next, let us compute $h(u)$ in another way. Denote by ξ the MMSE error in the above estimation problem:

$$\xi = y - y^\star = y - B^{-1}Aq. \quad (40)$$

Because $M = \begin{bmatrix} C & A^T \\ A & B \end{bmatrix}$ and $u = \begin{bmatrix} q \\ -y \end{bmatrix}$, it holds that

$$\begin{aligned} u^TMu &= q^TCq - y^TAq - (Aq)^Ty - y^TBy \\ &= \begin{bmatrix} q^T & -\xi^T \end{bmatrix} \begin{bmatrix} C - A^TB^{-1}A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} q \\ -\xi \end{bmatrix}. \end{aligned} \quad (41)$$

Since $u \sim \mathcal{N}(0, M^{-1})$, we have

$$\begin{bmatrix} q \\ -\xi \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} C - A^TB^{-1}A & 0 \\ 0 & B \end{bmatrix}^{-1}\right), \quad (42)$$

from which we see that

$$q \sim \mathcal{N}(0, (C - A^TB^{-1}A)^{-1}) \text{ and } \xi \sim \mathcal{N}(0, B^{-1}), \quad (43)$$

and that q is independent of ξ, which is expected because the MMSE error ξ must be orthogonal to the observation q by the MMSE estimation theory. Then,

$$\begin{aligned} h(u) &\overset{(a)}{=} h(q) + h(\xi) \\ &\overset{(b)}{=} \frac{1}{2}\log\Big((2\pi e)^m \det\big((C - A^TB^{-1}A)^{-1}\big)\Big) \\ &\quad + \frac{1}{2}\log\Big((2\pi e)^n \det\big(B^{-1}\big)\Big), \end{aligned} \quad (44)$$

where (a) follows since q and ξ are independent, and (b) follows by (43). Combining (39) and (44) establishes the Schur complement determinant formula.

Theorem 6 (Schur Complement Determinant Formula): When $B \succ 0$, we have

$$\det(M) = \det(B)\cdot\det(C - A^TB^{-1}A). \quad (45)$$

VI. Conclusion

This paper shows that the quadratic transform technique for solving FP is closely connected to the Schur complement in the sense that they can be derived from each other. The auxiliary variable in FP is the dual variable associated with an LMI in the Schur complement reformulation of the FP; it can also be interpreted in the context of an MMSE estimation problem. This connection leads to new generalizations and an alternative proof of the Schur complement determinant formula.

References

[1] K. Shen and W. Yu, "Quadratic transform for fractional programming in signal processing and machine learning," *IEEE Signal Process. Mag.*, vol. 42, no. 3, pp. 14–34, May 2025.

[2] F. Zhang, *The Schur Complement and Its Applications*. Springer Press, 2005.

[3] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[4] K. Shen and W. Yu, "Fractional programming for communication systems—Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, Mar. 2018.

[5] J. Schur, "Über potenzreihen die im inneren des einheitskreises beschränkt sind," *J. Reine U. Angewandte Mathematik*, pp. 205–232, 1917.

[6] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[7] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[8] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[9] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Writing on colored paper," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2001, p. 322.

[10] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 361–374, Feb. 2006.

[11] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge University Press, 2017.

[12] J. Zou, S. Sun, C. Masouros, Y. Cui, Y.-F. Liu, and D. Ng, "Energy-efficient beamforming design for integrated sensing and communications systems," *IEEE Trans. Commun.*, vol. 72, no. 6, pp. 3766–3782, Jun. 2024.

[13] M. Yuan, D. He, H. Yin, H. Wang, F. Liu, and Z. Wang, "Hybrid beamforming for mmWave integrated sensing and communication with multi-static cooperative localization," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 771–786, Jul. 2025.

[14] P. Liu, S. Xu, S. Tang, X. Wang, F. Xia, and W. Yuan, "Sensing-assisted secure communications: A rate-splitting approach," *IEEE Internet Things J.*, vol. 12, no. 20, pp. 42 876–42 890, Oct. 2025.

[15] Y. Tian, D. Wang, C. Huang, and W. Zhang, "Performance trade-off of integrated sensing and communications for multi-user backscatter systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17 310–17 323, Nov. 2024.

[16] Y. Chen, L. Zhao, and K. Shen, "Mixed max-and-min fractional programming for wireless networks," *IEEE Trans. Signal Process.*, vol. 72, pp. 337–351, Dec. 2023.

[17] K. M. Attiah and W. Yu, "Beamforming design for integrated sensing and communications using uplink-downlink duality," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2024.

[18] M. Zhu, L. Li, S. Xia, and T.-H. Chang, "Information and sensing beamforming optimization for multi-user multi-target MIMO ISAC systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Jun. 2023.

[19] M. Sion, "On general minimax theorems," *Pacific J. Math.*, pp. 171–176, 1958.