# Matrix Fisher Inequalities for non-invertible linear systems

Christophe Vignat
Université de Marne-la-Vallée, ENST URA 820
77 454 Marne-la-Vallée cedex, France
and E.E.C.S., University of Michigan
Email: vignat@univ-mlv.fr.

Jean-François Bercher
ESIEE, Laboratoire Signaux et Télécoms
Cité Descartes BP 99
93 162 Noisy-le-Grand Cedex
Email: bercherj@esiee.fr

*Abstract* — **In this paper, we show how Fisher inequalities first derived by Zamir can be proved and extended using a simplified approach. Cases of equality are detailed to reveal interesting links between notions of gaussianity and invertibility.**

## I. Introduction

In [1], R. Zamir extended an important series of Fisher Information Inequalities (FII), first derived in [2], to the case of non-invertible linear systems. In this paper, we provide an alternate elementary derivation of Zamir's inequalities, and show how the concept of extractable component [1] characterizing the cases of equality is related to the notion of minimum norm solution of the linear system.

## II. A new derivation of Zamir's F.I.I.

In this paper, we consider a linear system with $(n \times 1)$ random input vector $X$ and $(m \times 1)$ random output vector $Y$, represented by a full row rank $m \times n$ matrix $A$, as

$$Y = AX.$$

The probability densities $f_X$ and $f_Y$ are supposed to satisfy regularity conditions [2], and we denote $\phi_X$ and $\phi_Y$ their respective score (log derivative) functions. The following first theorem extends [3, Lemma 1].

**Theorem 1** *The best estimate (in the minimum mean square error sense) of $\phi_X(X)$ from observations $Y$ is*

$$\hat{\phi}_X(X) = A^T \phi_Y(Y) \tag{1}$$

We propose here a simple proof of this theorem by showing that, for any multivariate function $h : R^m \to R^n$

$$E_X \left( \phi_X(X) - \hat{\phi}_X(Y) \right)^T h(Y) = 0 \tag{2}$$

where $\hat{\phi}_X(Y) = A^T \phi_Y(Y)$. This result follows from elementary algebraic computation rules involving the score functions. The proof follows then from the classical orthogonality property of the MMSE estimate. Observe that theorem 1 extends [3, Lemma 1] since the components of $X$ are not supposed independent here.

Next, as was shown in [1], theorem 1 implies the following FII:

$$J_X \geq A^T J_Y A \tag{3}$$

$$J_Y \leq \left( A J_X^{-1} A^T \right)^{-1} \tag{4}$$

The simplified proof we propose here follows from the positivity of block matrices $U_1$ and $U_2$:

$$\begin{cases} U_1 = E \begin{bmatrix} \phi_X(X) \\ \phi_Y(Y) \end{bmatrix} \begin{bmatrix} \phi_X^T(X) & \phi_Y^T(Y) \end{bmatrix} \\ U_2 = E \begin{bmatrix} \phi_Y(Y) \\ \phi_X(X) \end{bmatrix} \begin{bmatrix} \phi_Y^T(Y) & \phi_X^T(X) \end{bmatrix}. \end{cases}$$

Moreover, we remark that this result can be retrieved as a consequence of a more general theorem by Papathanasiou [2].

## III. Case of equality in Zamir's F.I.I.

In this section, the components $X_i$ of $X$ are supposed mutually independent. The case of equality in inequality (3) is characterized by the following theorem.

**Theorem 2** *Equality holds in (3) if and only of matrix $A$ possesses $(n-m)$ null columns or, equivalently, if $A$ writes, up to a permutation of its column vectors*

$$A = [A_0 | 0],$$

*where $A_0$ is a $(m \times m)$ non-singular matrix.*

The case of equality in (4) is characterized as follows:

**Theorem 3** *Equality holds in (4) if and only if each component $X_i$ of $X$ verifies at least one of the three properties*

- *a- $X_i$ is gaussian*
- *b- $X_i$ coincides with the i-th component of the minimum norm solution $X_0$ of system $Y = AX$*
- *c- $X_i$ correspond to a null column of $A$.*

We show that this result can be easily proved by considering the set of solutions of system $Y = AX$ expressed as $X = X_0 + (I - A^{\#}A)Z$ where $X_0 = A^{\#}Y$ is the minimum norm solution, $A^{\#}$ denotes the pseudoinverse of $A$ while $Z$ is any vector.

As a conclusion, equality in (4) requires that all components of $X$ that differ from the components of the minimum norm solution $X_0$ and that influence the output $Y$ – the so-called non-extractable and non-irrelevant [1] components of $X$ – should have a gaussian distribution.

## Acknowledgments

The authors wish to thank Pr. A. Hero for helpful remarks about this problem.

## References

[1] R. Zamir, "A Proof of the Fisher Information Inequality via a Data Processing Argument", IEEE Trans. on Information Theory, IT 44, 3, pp. 1246-1250, 1998.

[2] V. Papathanasiou, "Some Characteristic Properties of the Fisher Information Matrix via Cacoullos-type Inequalities", Journal of Multivariate Analysis, 14, pp. 256-265, 1993.

[3] R. Zamir, "A Necessary and Sufficient Condition for Equality in the Matrix Fisher Information Inequality", technical report, Tel Aviv University, Dept. Elec. Eng. Syst., 1997.