UvA ECONOMETRICS

Discussion Paper: 2004/11

# On the solution of Stein's equation and Fisher information matrix of an ARMAX process

André Klein and Peter Spreij

**www.fee.uva.nl/ke/UvA-Econometrics**

Department of Quantitative Economics
Faculty of Economics and Econometrics
Universiteit van Amsterdam
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

# On the solution of Stein's equation and Fisher's information matrix of an ARMAX process

André Klein[a] and Peter Spreij[b]
[a]Department of Quantitative Economics
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam, The Netherlands.
E-mail: A.A.B.Klein@uva.nl.
[b] Korteweg-de Vries Institute for Mathematics
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam, The Netherlands.
E-mail: spreij@science.uva.nl.

**Abstract**

The main goal of this paper consists in expressing the solution of a Stein equation in terms of the Fisher information matrix (FIM) of a scalar ARMAX process. A condition for expressing the FIM in terms of a solution to a Stein equation is also set forth. Such interconnections can be derived when a companion matrix with eigenvalues equal to the roots of an appropriate polynomial associated with the ARMAX process is inserted in the Stein equation. The case of algebraic multiplicity greater than or equal to one is studied. The FIM and the corresponding solution to Stein's equation are presented as solutions to systems of linear equations. The interconnections are obtained by using the common particular solution of these systems. The kernels of the structured coefficient matrices are described as well as some right inverses. This enables us to find a solution to the newly obtained linear system of equations.

*AMS classification: 15*A*06, 15*A*09, 15*A*24, 62*F*12*

*Keywords*: Fisher information matrix; Stein equation; Linear systems; Kernel ; Coefficient matrix; ARMAX process.

## 1. Introduction

The purpose of this paper consists in deriving a solution to a Stein equation expressed in terms of the asymptotic Fisher information matrix of an ARMAX process. The condition for expressing the Fisher information matrix in terms of a Stein solution is also set forth. In [9] an alternative interconnection is established for the case of an ARMA process where the vectorized form of the Fisher information matrix is used.

The ARMAX processes are of common use in signal processing, control and system theory, statistics and econometrics, see e.g. [15], [1], [2]. The concept of the Fisher information plays a vital role in estimation theory and since more recently in physics, see e.g. [3], [4]. Various algorithms have been developed for computing the information matrix, e.g. [5], [6]. In [6] two algorithms have been proposed for a fast computation of the Fisher information matrix of a SISO process. The ARMAX process is a special case of the SISO process, the latter is discussed in [15].

A companion matrix with eigenvalues equal to the roots of an appropriate polynomial derived from the ARMAX representation is used as a coefficient in the Stein equation. The solution $S$ of this equation can be factorized as $Ax$, where $x$ is a solution of the equation $Ax = b$ for some $b$. A similar factorization is applied to the Fisher information matrix resulting in a system $A'x = b'$ and the coefficient matrices $A$ and $A'$ are $q \times q^2$ where $q$ is the degree of an appropriate polynomial associated with the ARMAX process. The Stein equation has been extensively studied in the mathematical literature, e.g. [13]. The use of a companion matrix in a Stein equation is also studied in [12].

By proving surjectivity of the coefficient matrices and using a common particular solution of both linear systems of equations leads to the following interconnections. A solution to Stein's equation is expressed in terms of the Fisher information matrix and vice versa. In [9] only a solution of a Stein equation expressed in terms of the vectorized form of the Fisher information matrix of an ARMA process is studied. The kernels of the newly obtained coefficient matrices are derived as well as a right inverse of the coefficient matrix associated with the Fisher information matrix. This makes it possible to find a solution to the newly obtained linear system of equations. The approach set forth in this paper is applied for one block of the Fisher information matrix.

The paper is organized as follows. First we present the definitions which are followed by interconnections between blocks space of the Fisher information matrix and a solution to a Stein equation. This is done for the algebraic multiplicity greater than or equal to one. In Section 3, algorithms describing the structure of the kernel of coefficient matrices associated with the linear systems of equations obtained in Section 2, are developed. In Section 4, an example is provided to illustrate the construction of a solution to Stein's equation in terms of the Fisher information matrix. In Section 5, the case of the Fisher information matrix containing all the parameter blocks and not decomposed is mentioned.

## 2. Link solution Stein's equation-Fisher's information

### 2.1. The ARMAX process

In this section a block of the Fisher information matrix of an ARMAX process is used to develop an interconnection with a solution to Stein's equation. For that purpose we first introduce the ARMAX process and we discuss Stein's equation in Section 2.2

$$\begin{aligned} a(z) &= z^p + a_1 z^{p-1} + \cdots + a_p, \\ b(z) &= z^q + b_1 z^{q-1} + \cdots + b_q, \\ c(z) &= z^r + c_1 z^{r-1} + \cdots + c_r. \end{aligned}$$

The reciprocal polynomials $a^*(z)$, $b^*(z)$ and $c^*(z)$ are $a^*(z) = z^p a(z^{-1})$, $b^*(z) = z^q b(z^{-1})$ and $c^*(z) = z^r c(z^{-1})$.

The ARMAX process $y(t)$ is specified as the stationary invertible (which exists under suitable conditions, see below) solution of

$$a^*(L)y(t) = b^*(L)x(t) + c^*(L)\varepsilon(t), \tag{2.1}$$

with $L$ the lag operator, $x(t)$ the input process which is independent of the white noise sequence $\varepsilon(t)$ that has variance $\sigma^2$. We make the assumptions that $a(z)$, $b(z)$ and $c(z)$ have zeros inside the unit disc. The input $x(t)$ is described by an AR process with spectral density $(2\pi)^{-1} R_x(z)$ where $R_x(z) = \sigma_\eta^2 \left(1/h(z)h(z^{-1})\right)$ and $1/h(z)$ is the transfer function. We assume $\sigma_\eta^2 = 1$, the latter represents the variance of the white noise sequence $\eta(t)$ which generates the AR process $x(t)$ and $\varepsilon(t)$ and $\eta(t)$ are independent.
Define the vectors

$$u_k(z) = (1, z, \ldots, z^{k-1})^\top, u_k^*(z) = (z^{k-1}, z^{k-2}, \ldots, 1)^\top$$

and

$$\theta = (a_1, a_2, \ldots, a_p, b_1, b_2, ..., b_q, c_1, c_2, \ldots, c_r)^\top.$$

We assume the polynomial $a(z)$ having $p_0$ distinct roots, $\alpha_1$, $\alpha_2$,..., $\alpha_{p_0}$ , with algebraic multiplicity $n_1 + 1, n_2 + 1, ..., n_{p_0} + 1$ respectively and $\sum_{i=1}^{p_0}(n_i + 1) = p$, $b(z)$ has $q_0$ distinct roots, $\beta_1$, $\beta_2$,..., $\beta_{q_0}$ , with algebraic multiplicity $m_1+1, m_2+1, ..., m_{q_0}+1$ respectively and $\sum_{i=1}^{q_0}(m_i+1) = q$ and polynomial $c(z)$ has $r_0$ distinct roots $\gamma_1$, $\gamma_2$,..., $\gamma_{r_0}$ with algebraic multiplicity $s_1+1, s_2+1, ..., s_{r_0}+1$ respectively and $\sum_{i=1}^{r_0}(s_i + 1) = r$. The function $h(z)$ has $v_0$ distinct zeros $\tau_1$, $\tau_2$,..., $\tau_{v_0}$ with algebraic multiplicity $\ell_1 + 1, \ell_2 + 2, ..., \ell_{v_0} + 1$ respectively and $\sum_{i=1}^{v_0}(\ell_i + 1) = v$.
It is known, see [7] and [6], that Fisher's information matrix of (2.1) is $F(\theta) = \left(1/\sigma^2\right) G(\theta)$ with the following block decomposition for $G(\theta)$,

$$G(\theta) = \begin{pmatrix} G_{aa}(\theta) & G_{ab}(\theta) & G_{ac}(\theta) \\ G_{ab}^\top(\theta) & G_{bb}(\theta) & G_{bc}(\theta) \\ G_{ac}^\top(\theta) & G_{bc}^\top(\theta) & G_{cc}(\theta) \end{pmatrix}. \tag{2.2}$$

The matrices appearing in (2.2) can be expressed as:

$$
\begin{aligned}
G_{aa}(\theta) &= \frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{b(z)b(z^{-1})R_x(z)u_p(z)u_p^\top(z^{-1})}{a(z)a(z^{-1})c(z)c(z^{-1})}\frac{dz}{z} && (2.3)\\
&\quad +\frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{u_p(z)u_p^\top(z^{-1})}{a(z)a(z^{-1})}\frac{dz}{z}\ , && (2.4)\\
G_{ab}(\theta) &= -\frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{b(z)R_x(z)u_p(z)u_q^\top(z^{-1})}{a(z)c(z)c(z^{-1})}\frac{dz}{z}\ , && (2.5)\\
G_{ac}(\theta) &= -\frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{u_p(z)u_r^\top(z^{-1})}{a(z)c(z^{-1})}\frac{dz}{z}\ , && (2.6)\\
G_{bb}(\theta) &= \frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{R_x(z)u_q(z)u_q^\top(z^{-1})}{c(z)c(z^{-1})}\frac{dz}{z}\ , && (2.7)\\
G_{bc}(\theta) &= 0, && (2.8)\\
G_{cc}(\theta) &= \frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{u_r(z)u_r^\top(z^{-1})}{c(z)c(z^{-1})}\frac{dz}{z}\ . && (2.9)
\end{aligned}
$$

As can be seen from the blocks (2.3)-(2.9) which constitute $G(\theta)$, the terms in block (2.4), (2.6) and (2.9) have representations which correspond to the ARMA part of $G(\theta)$, whereas the remaining blocks contain information of the input process $x(t)$. In [6] a detailed derivation of the representations (2.3)-(2.9) is provided. As mentioned earlier, in [9] interconnections are established using representations in vectorized form.

In this paper we abandon the idea of vectorizing matrices so that a different and more general approach is obtained. We derive linear systems of equations that lead to interconnections between a solution to Stein's equation and the Fisher information matrix. We consider the $(b,b)$-block extensively, the remaining blocks can be treated in a similar manner. Block $G_{bb}(\theta)$ given in (2.7) can alternatively be written as

$$G_{bb}(\theta) = \frac{1}{2\pi i}\oint\limits_{|z|=1} \frac{u_q(z)u_q^{*\top}(z)}{h(z)c(z)h^*(z)c^*(z)z^{l+1}}dz. \qquad (2.10)$$

For technical convenience we write $l+1 = q-v-r$ and the cases $l+1>0$, $l+1=0$ and $l+1<0$ shall be discussed. The polynomials $h(z)$, $c(z)$ and $z^{l+1}$ have their roots inside the unit circle. For typographical brevity we introduce the following notation. Given a polynomial $p(\cdot)$, assume that for some natural number $j$ $(z-\beta)^j$ is a factor of $p(\cdot)$, and $\beta$ has multiplicity $j \geq 1$, we define the polynomial $p_j(.\ ;\beta)$ by $p_j(z\ ;\beta) = \frac{p(z)}{(z-\beta)^j}$. Applying Cauchy's residue theorem to (2.10) for $l+1>0$, one obtains

$$G_{bb}(\theta) = g_1(\gamma_1) + g_2(\gamma_2) + \cdots + g_{r_0}(\gamma_{r_0}) + k_1(\tau_1) + k_2(\tau_2) + \cdots + k_{v_0}(\tau_{v_0}) + f(0),$$

where

$$
\begin{aligned}
g_i(\gamma_i) &= \frac{1}{s_i!}\left(\frac{\partial^{s_i}}{\partial z^{s_i}}\frac{u_q(z)u_q^{*\top}(z)}{c_{s_i+1}(z;\gamma_i)h(z)h^*(z)c^*(z)z^{l+1}}\right)_{z=\gamma_i} \quad, \ i=1,\ldots,r_0 \qquad (2.11)\\
k_j(\tau_j) &= \frac{1}{\ell_j!}\left(\frac{\partial^{\ell_j}}{\partial z^{\ell_j}}\frac{u_q(z)u_q^{*\top}(z)}{c(z)h_{\ell_j+1}(z;\tau_j)h^*(z)c^*(z)z^{l+1}}\right)_{z=\tau_j} \quad, \ j=1,\ldots,v_0 \qquad (2.12)\\
f(0) &= \frac{1}{l!}\left(\frac{\partial^{l}}{\partial z^{l}}\frac{u_q(z)u_q^{*\top}(z)}{c(z)h(z)h^*(z)c^*(z)}\right)_{z=0}. \qquad (2.13)
\end{aligned}
$$

A useful factorization of $G_{bb}(\theta)$ can be obtained by applying Leibnitz rule to $j$-fold differentiation of a ratio of two functions as in (2.11), (2.12) and (2.13). To that end we need to introduce a number of expressions, we define

$$
\mathcal{U}_{s_i}^{(k)}(z) = \frac{\partial^k}{\partial z^k}\left(u_q(z)u_q^{*\top}(z)\right) \qquad k=s_i,\ldots,0.
$$

The matrices $\mathcal{U}_{s_i}(z)$, for $i=1,\ldots,r_0$, have the structure

$$
\mathcal{U}_{s_i}(z) = \left(\ \mathcal{U}_{s_i}^{(s_i)}(z),\ \ \mathcal{U}_{s_i}^{(s_i-1)}(z),\ \ \ldots\ \ ,\mathcal{U}_{s_i}^{(0)}(z)\ \right).
$$

The following representations are now considered for all the eigenvalues

$$
\begin{aligned}
\mathcal{U}_r(\gamma) &= \left(\ \mathcal{U}_{s_1}(\gamma_1),\ \ \mathcal{U}_{s_2}(\gamma_2),\ \ \ldots,\ \ \mathcal{U}_{s_{r_0}}(\gamma_{r_0})\ \right),\\
\mathcal{U}_v(\tau) &= \left(\ \mathcal{U}_{\ell_1}(\tau_1),\ \ \mathcal{U}_{\ell_2}(\tau_2),\ \ \ldots,\ \ \mathcal{U}_{\ell_{v_0}}(\tau_{v_0})\ \right),\\
\mathcal{U}_l(0) &= \left(\ \mathcal{U}_l^{(l)}(0),\ \ \mathcal{U}_l^{(l-1)}(0),\ \ \ldots,\ \ \mathcal{U}_l^{(0)}(0)\ \right).
\end{aligned}
$$

Let

$$
\begin{aligned}
\mu_i(z) &= \frac{1}{c_{s_i+1}(z;\gamma_i)h(z)h^*(z)c^*(z)z^{l+1}},\\
\zeta_j(z) &= \frac{1}{c(z)h_{\ell_j+1}(z;\tau_j)h^*(z)c^*(z)z^{l+1}},\\
\xi(z) &= \frac{1}{c(z)h(z)h^*(z)c^*(z)},
\end{aligned}
$$

then we define

$$
\mu_{s_i}^{(k)}(z) = \begin{pmatrix} s_i \\ k \end{pmatrix}\frac{\partial^k}{\partial z^k}\mu_i(z) \quad k=0,\ldots,s_i,
$$

and

$$
\mu_{s_i}(\gamma_i) = \frac{1}{s_i!}\left(\mu_{s_i}^{(0)}(z),\mu_{s_i}^{(1)}(z),\ldots,\mu_{s_i}^{(s_i)}(z)\right)^{\top}_{z=\gamma_i} \quad i=1,\ldots,r_0.
$$

Analogously, we define

$$
\zeta_{\ell_j}^{(s)}(z) = \begin{pmatrix} \ell_j \\ m \end{pmatrix}\frac{\partial^m}{\partial z^m}\ \zeta_j(z) \qquad m=0,\ldots,\ell_j
$$

and introduce

$$
\zeta_{\ell_j}(\tau_j) = \frac{1}{\ell_j!}\left(\zeta_{\ell_j}^{(0)}(z),\zeta_{\ell_j}^{(1)}(z),\ldots,\zeta_{\ell_j}^{(\ell_j)}(z)\right)^{\top}_{z=\tau_j} \qquad j=1,\ldots,v_0.
$$

Similarly we define

$$\xi_0^{(n)}(z) = \left( \begin{array}{c} l \\ n \end{array} \right) \frac{\partial^n}{\partial z^n}\, \xi(z) \qquad n = 0, \ldots, l,$$

and

$$\xi_0(0) \quad = \frac{1}{l!} \left( \xi_0^{(0)}(z), \xi_0^{(1)}(z), \ldots, \xi_0^{(l)}(z) \right)^{\top}_{z=0}.$$

With the above notations we can now introduce the vector $\vartheta$ given by

$$\vartheta = \left( \mu_{s_1}^{\top}(\gamma_1), \mu_{s_2}^{\top}(\gamma_2), \ldots, \mu_{s_{r_0}}^{\top}(\gamma_{r_0}), \zeta_{\ell_1}^{\top}(\tau_1), \zeta_{\ell_2}^{\top}(\tau_2), \ldots, \zeta_{\ell_{v_0}}^{\top}(\tau_{v_0}), \xi_0^{\top}(0) \right)^{\top}. \tag{2.14}$$

With the aid of this notation we can factorize $G_{bb}(\theta)$ according to

$$G_{bb}(\theta) = (\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))\,(\vartheta \otimes I_q). \tag{2.15}$$

We illustate this notation with an example.

**Example 2.1.** Consider the ARMAX process with $p = q = 3$, $r = 2$ and $v = 1$, the polynomials involved are $c(z) = (z-\gamma)^2$ and $h(z) = (z-\tau)$. This case will be used for the remaining examples in this paper. Consequently, the Fisher information matrix block $G_{bb}(\theta)$ admits the form

$$G_{bb}(\theta) = \frac{1}{2\pi i} \oint_{|z|=1} \begin{pmatrix} z^2 & z & 1 \\ z^3 & z^2 & z \\ z^4 & z^3 & z^2 \end{pmatrix} \frac{dz}{(z-\tau)\,(z-\gamma)^2\,(1-z\,\tau)\ (1-z\,\gamma)^2}.$$

The components of the Toeplitz and symmetric matrix $G_{bb}(\theta)$ are obtained by means of Cauchy's residue theorem, to have

$$\frac{\tau^j}{(\tau-\gamma)^2\,(1-\ \tau^2)\ (1-\tau\,\gamma)^2} \quad + \quad \left( \frac{\partial}{\partial z} \frac{z^j}{(z-\tau)\,(1-z\,\tau)\ (1-z\,\gamma)^2} \right)_{z=\gamma} \quad \text{for } j = 0, 1, 2, 3, 4.$$

The matrix $G_{bb}(\theta)$ is then

$$G_{bb}(\theta) = \frac{1}{(\gamma^2-1)^3\ \ (\gamma\,\tau-1)^2\ \ (\tau^2-1)} \begin{pmatrix} G_{bb}^{11}(\theta) & G_{bb}^{12}(\theta) & G_{bb}^{13}(\theta) \\ G_{bb}^{21}(\theta) & G_{bb}^{22}(\theta) & G_{bb}^{23}(\theta) \\ G_{bb}^{31}(\theta) & G_{bb}^{32}(\theta) & G_{bb}^{33}(\theta) \end{pmatrix},$$

where

$$\begin{aligned} G_{bb}^{11}(\theta) &= G_{bb}^{22}(\theta) = G_{bb}^{33}(\theta) = 1 + 2\gamma\tau - 2\gamma^3\tau - \gamma^4\tau^2 - \gamma^2(\tau^2-1), \\ G_{bb}^{12}(\theta) &= G_{bb}^{23}(\theta) = G_{bb}^{21}(\theta) = G_{bb}^{32}(\theta) = 2\gamma + \tau - \gamma^4\tau - 2\gamma^3\tau^2, \\ G_{bb}^{13}(\theta) &= G_{bb}^{31}(\theta) = -\gamma^4 + 2\gamma\tau - 2\gamma^3\tau + \tau^2 - 3\gamma^2\left(\tau^2-1\right). \end{aligned}$$

### 2.2. The Stein equation

We now introduce the Stein equation and its solution. Let $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$ and $\Gamma \in \mathbb{C}^{n\times m}$ and consider the Stein equation

$$S - BSA^{\top} = \Gamma. \tag{2.16}$$

It has a unique solution iff $\lambda\mu \neq 1$ for any $\lambda \in \sigma(A)$ and $\mu \in \sigma(B)$. From [13] we take

**Theorem 2.2.** *Let $A$ and $B$ be such that there is a single closed contour $C$ with $\sigma(B)$ inside $C$ and for each nonzero $w \in \sigma(A)$, $w^{-1}$ is outside $C$. Then for an arbitrary $\Gamma$ the Stein equation (2.16) has a unique solution $S$*

$$S = \frac{1}{2\pi i} \oint_C (\lambda I - B)^{-1} \Gamma (I - \lambda A)^{-\top} d\lambda.$$

This theorem is used to interconnect the Fisher information matrix and a solution to a Stein equation. We are interested in the case $A = B = E$, where $E$ is an appropriate companion matrix. Representation (2.16) becomes

$$S_{bb} - E S_{bb} E^\top = \Gamma, \tag{2.17}$$

where the companion matrix $E$ is chosen to be

$$E = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & & 1 \\ -e_{r+v} & -e_{r+v-1} & \cdots & -e_1 \end{pmatrix}$$

and the entries $e_i$ are the coefficients of the polynomial $e(z) = c(z)h(z) = z^{r+v} + \sum_{i=1}^{r+v} e_i z^{r+v-i}$. The companion matrix $E$ has the property $\det(zI - E) = e(z)$, see e.g. [14]. The choice of the companion matrix $E$ yields a unique solution to the Stein equation (2.17) since all the eigenvalues of $E$ are within the unit disc. Using companion matrices in (2.16) for the coefficients $A$ and $B$ is also studied in [12].

In [8] the following result has been obtained. The Fisher information matrix of an ARMA process coincides with a corresponding solution to a Stein equation for a specific choice of $\Gamma$, namely $\Gamma = w_{p+r} w_{p+r}^\top$, where $w_{p+r}$ is the last standard basis vector in $\mathbb{R}^{p+r}$ and $p$ and $r$ are the degrees of the ARMA polynomials. In a similar way we can show that $G_{bb}(\theta)$ coincides with the solution to the Stein equation (2.17) for $l + 1 = 0$ and with $\Gamma = w_{r+v} w_{r+v}^\top$, where $w_{r+v}$ is the last standard basis vector in $\mathbb{R}^{r+v}$. See also Section 5. Consequently, $G_{bb}(\theta)$ satisfies the Stein equation

$$G_{bb}(\theta) - E G_{bb}(\theta) E^\top = w_{r+v} w_{r+v}^\top.$$

The general result of Theorem 2.2 applied to (2.17) gives

$$S_{bb} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{\operatorname{adj}(zI - E)\, \Gamma \operatorname{adj}(I - zE)^\top z^{l+1}}{h(z)c(z)h^*(z)c^*(z)z^{l+1}} dz. \tag{2.18}$$

Then, applying Cauchy's residue theorem to (2.18) yields,

$$S_{bb} = \mathcal{G}_1(\gamma_1) + \mathcal{G}_2(\gamma_2) + \cdots + \mathcal{G}_{r_0}(\gamma_{r_0}) + \mathcal{K}_1(\tau_1) + \mathcal{K}_2(\tau_2) + \cdots + \mathcal{K}_{v_0}(\tau_{v_0}) + \mathcal{F}(0),$$

where

$$\begin{aligned} \mathcal{G}_i(\gamma_i) &= \frac{1}{s_i!} \left( \frac{\partial^{s_i}}{\partial z^{s_i}} \frac{\operatorname{adj}(zI - E)\, \Gamma \operatorname{adj}(I - zE)^\top z^{l+1}}{c_{s_i+1}(z;\gamma_i) h(z) h^*(z) c^*(z) z^{l+1}} \right)_{z=\gamma_i}, \\ \mathcal{K}_j(\tau_j) &= \frac{1}{\ell_j!} \left( \frac{\partial^{\ell_j}}{\partial z^{\ell_j}} \frac{\operatorname{adj}(zI - E)\, \Gamma \operatorname{adj}(I - zE)^\top z^{l+1}}{c(z) h_{\ell_j+1}(z;\tau_j) h^*(z) c^*(z) z^{l+1}} \right)_{z=\tau_j}, \\ \mathcal{F}(0) &= \frac{1}{l!} \left( \frac{\partial^{l}}{\partial z^{l}} \frac{\operatorname{adj}(zI - E)\, \Gamma \operatorname{adj}(I - zE)^\top z^{l+1}}{c(z) h(z) h^*(z) c^*(z)} \right)_{z=0}. \end{aligned}$$

A similar factorization as in (2.15) can be applied. For that purpose we use

$$\widetilde{\mathcal{M}}_{s_i}^{(k)}(z) = \frac{\partial^k}{\partial z^k}\left(\mathrm{adj}\,(zI - E)\,\Gamma\mathrm{adj}\,(I - zE)^{\top} z^{l+1}\right) \qquad k = s_i, \ldots, 0,$$

to define

$$\widetilde{\mathcal{M}}_{s_i}(z) = \left(\ \widetilde{\mathcal{M}}_{s_i}^{(s_i)}(z), \quad \widetilde{\mathcal{M}}_{s_i}^{(s_i-1)}(z), \quad \ldots \quad , \widetilde{\mathcal{M}}_{s_i}^{(0)}(z)\ \right).$$

Representations that contain all the eigenvalues are then

$$\begin{aligned}
\widetilde{\mathcal{M}}_r(\gamma) &= \left(\ \widetilde{\mathcal{M}}_{s_1}(\gamma_1), \quad \widetilde{\mathcal{M}}_{s_2}(\gamma_2), \quad \ldots, \quad \widetilde{\mathcal{M}}_{s_{r_0}}(\gamma_{r_0})\ \right), \\
\widetilde{\mathcal{M}}_v(\tau) &= \left(\ \widetilde{\mathcal{M}}_{\ell_1}(\tau_1), \quad \widetilde{\mathcal{M}}_{\ell_2}(\tau_2), \quad \ldots, \quad \widetilde{\mathcal{M}}_{\ell_{v_0}}(\tau_{v_0})\ \right), \\
\widetilde{\mathcal{M}}_l(0) &= \left(\ \widetilde{\mathcal{M}}_l^{(l)}(0), \quad \widetilde{\mathcal{M}}_l^{(l-1)}(0), \quad \ldots, \quad \widetilde{\mathcal{M}}_l^{(0)}(0)\ \right).
\end{aligned}$$

With the same vector $\vartheta$ as given in (2.14) we then have the factorization

$$S_{bb} = \left(\widetilde{\mathcal{M}}_r(\gamma)\ \widetilde{\mathcal{M}}_v(\tau)\ \widetilde{\mathcal{M}}_l(0)\right)(\vartheta \otimes I_{r+v}). \tag{2.19}$$

### 2.3. Interconnections between a solution to Stein's equation and Fisher's information matrix

We now proceed constructing an interconnection between $G_{bb}(\theta)$ and $S_{bb}$ by solving $(\vartheta \otimes I_q)$ and $(\vartheta \otimes I_{r+v})$ from the linear equations (2.15) and (2.19) respectively. This will happen according to the solution of two linear systems of the form $AX = B$ where $A$, $B$ and $X$ are matrices of appropriate dimension. The matrix $A$ will be represented by the corresponding coefficient matrices $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))$ and $\left(\widetilde{\mathcal{M}}_r(\gamma)\ \widetilde{\mathcal{M}}_v(\tau)\ \widetilde{\mathcal{M}}_l(0)\right)$ in (2.15) and (2.19) respectively. The linear system $AX = B$ has a solution if and only if $B \in \mathrm{Im}\ (A)$, a solution of the linear system is given by $X = X_0 + \mathcal{A}$ where $X_0$ is a particular solution of the matrix equation $AX = B$ and $\mathcal{A} \in \mathrm{Ker}(A)$, the kernel of $A$. We take the matrix $X_0 = A^+B$ and where $A^+$ is the Moore-Penrose inverse of $A$, see e.g. [14]. In general, the solution set is a manifold of matrices obtained by a shift of $\mathrm{Ker}(A)$. This will be applied to the linear systems (2.15) and (2.19) in order to obtain an interconnection or equation involving the Fisher information matrix and a solution to Stein's equation. For that purpose the particular solutions of the linear systems (2.15) and (2.19) are considered. From (2.15) one obtains

$$(\vartheta \otimes I_q) = (\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))^+\ G_{bb}(\theta) + \mathcal{A}, \tag{2.20}$$

where $\mathcal{A} \in \mathrm{Ker}\ (\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))$.
Likewise, the solution of Stein's equation takes the form

$$(\vartheta \otimes I_{r+v}) = \left(\widetilde{\mathcal{M}}_r(\gamma)\ \widetilde{\mathcal{M}}_v(\tau)\ \widetilde{\mathcal{M}}_l(0)\right)^+ S_{bb} + \mathcal{B}, \tag{2.21}$$

where $\mathcal{B} \in \mathrm{Ker}\ \left(\widetilde{\mathcal{M}}_r(\gamma)\ \widetilde{\mathcal{M}}_v(\tau)\ \widetilde{\mathcal{M}}_l(0)\right)$.
Considering equations (2.10) and (2.18), three situations shall be considered, $l+1 > 0$, $l+1 = 0$ and $l+1 < 0$. The results will be presented as Proposition 2.6, Proposition 2.7 and Proposition 2.8.

First we consider the case $l+1 > 0$. Then we can write

$$I_q = \begin{pmatrix} I_{r+v} & 0 \\ 0 & I_{q-(r+v)} \end{pmatrix} \qquad \text{and} \qquad I_q \otimes \vartheta = \begin{pmatrix} I_{r+v} \otimes \vartheta & 0 \\ 0 & I_{q-(r+v)} \otimes \vartheta \end{pmatrix}.$$

In order to obtain the forms $(\vartheta \otimes I_{r+v})$ and $(\vartheta \otimes I_q)$ we use the following property of the Kronecker product of two matrices. Let $A$ be an $m\times n$ matrix and $B$ a $p\times q$ matrix. Then there exist $pm\times pm$ and $nq \times nq$ universal permutation matrices $\mathcal{R}_{pm}$ and $\mathcal{R}_{nq}$ such that $\forall\ A, B : \mathcal{R}_{pm}\,(A\otimes B)\,\mathcal{R}_{nq} = B\otimes A$, see e.g. [14]. Double application of this rule to $A = \vartheta$ and $B = I_q$, respectively $B = I_{r+v}$ results in an equation which involves the Fisher information matrix and a solution to a Stein equation. This is summarized in Proposition 2.6. First we need to show the surjectivity of the coefficient matrices in (2.15) and (2.19). For that purpose it remains to show that the rank of the corresponding coefficient matrices is $q$. This will be done with the help of the following results.

**Proposition 2.3.** *The matrix* $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ *has rank* $q$.

**Proof.** We shall show that the matrix $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ has a right inverse. For that purpose the $(q \times q)$ generalized Vandermonde matrix

$$\mathcal{W}_{r,v}(\gamma,\tau) = \Big(\ \mathcal{W}_{s_1}(\gamma_1),\quad \mathcal{W}_{s_2}(\gamma_2),\quad ...,\quad \mathcal{W}_{s_{r_0}}(\gamma_{r_0}),\quad \mathcal{V}_{\ell_1}(\tau_1),\quad \mathcal{V}_{\ell_2}(\tau_2),\quad ...,\quad \mathcal{V}_{\ell_{v_0}}(\tau_{v_0})\ \Big)$$

is introduced, where

$$\mathcal{W}_{s_i}(\gamma_i) = \Big(\ \mathcal{W}_{s_i}^{(s_i)}(z),\quad \mathcal{W}_{s_i}^{(s_i-1)}(z),\quad \ldots\quad ,\mathcal{W}_{s_i}^{(0)}(z)\ \Big)_{z=\gamma_i}$$

and

$$\mathcal{W}_{s_i}^{(s_i-k)}(\gamma_i) = \left(\frac{\partial^{s_i-k}}{\partial z^{s_i-k}}u_q(z)\right)_{z=\gamma_i} \qquad k = 0,1,\ldots,s_i.$$

The blocks that constitute the matrix

$$\mathcal{V}_{\ell_j}(\tau_j) = \Big(\ \mathcal{V}_{\ell_j}^{(\ell_j)}(z),\quad \mathcal{V}_{\ell_j}^{(\ell_j-1)}(z),\quad \ldots\quad ,\mathcal{V}_{\ell_j}^{(0)}(z)\ \Big)_{z=\tau_j}$$

have the following representation

$$\mathcal{V}_{\ell_j}^{(\ell_j-k)}(\tau_j) = \left(\frac{\partial^{\ell_j-k}}{\partial z^{\ell_j-k}}u_q(z)\right)_{z=\tau_j} \qquad k = 0,1,\ldots,\ell_j.$$

One may check that

$$(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))\,(I_q\otimes w_q) = \mathcal{W}_{r,\,v}(\gamma,\tau),$$

from which it follows that

$$(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))\left((\mathcal{W}_{r,\,v}(\gamma,\tau))^{-1}\otimes w_q\right) = I_q.$$

The Vandermonde matrix $\mathcal{W}_{r,\,v}(\gamma,\tau)$ is invertible, see e.g. [11]. Consequently, an appropriate right inverse of $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ is $\left((\mathcal{W}_{r,\,v}(\gamma,\tau))^{-1}\otimes w_q\right)$, where $w_q$ is the last standard basis vector in $\mathbb{R}^q$, from which we can conclude that the matrix $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ has full row rank. ■

For proving the surjectivity of the matrix $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))$, some additional general concepts are needed. To this end we introduce some notation. Consider a matrix $A \in \mathbb{R}^{n\times n}$ in the following companion form.

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 \\ -a_n & & & -a_2 & -a_1 \end{pmatrix}. \tag{2.22}$$

Let $a^\top = (a_1, \ldots, a_n)$, and we redefine $u(z)^\top = (1, z, \ldots, z^{n-1})$ and $u^*(z)^\top = (z^{n-1}, \ldots, 1)$. Define recursively the Hörner polynomials $a_k(z)$ by $a_0(z) = 1$ and $a_k(z) = za_{k-1}(z) + a_k$. Notice that $a_n(z)$ is the characteristic polynomial of $A$. We will denote it by $\pi(z)$.
Write $\widetilde{a}(z)$ for the $n$-vector $(a_0(z), \ldots, a_{n-1}(z))^\top$. Furthermore $S$ will denote the shift matrix, so $S_{ij} = \delta_{i,j+1}$ and $J$ the *backward* or *antidiagonal* identity matrix.

**Lemma 2.4.** *Let $A$ be an $n \times n$ companion matrix as in (2.22). Let $P_k(z) = (\mathrm{adj}(z - A), \frac{d}{dz}\mathrm{adj}(z - A), \ldots, \frac{d^{k-1}}{dz^{k-1}}\mathrm{adj}(z - A))$ and $P = (P_{k_1}(\lambda_1), \ldots, P_{k_s}(\lambda_s)) \in \mathbb{R}^{n \times n^2}$, where the $\lambda_j$ are all the different eigenvalues of $A$, with multiplicities $k_j$, so $\sum_{j=1}^{s} k_j = n$. Then $P$ has rank $n$.*

**Proof.** We will use Proposition 3.2 of [11], which says that the adjoint of $z - A$, with $A$ a companion matrix, is

$$\mathrm{adj}(z - A) = u(z)\widetilde{a}(z)^\top J - \pi(z) \sum_{j=0}^{n-1} z^j S^{j+1}. \tag{2.23}$$

If we evaluate this expression for $z$ equal to an eigenvalue, the second term at the RHS vanishes, and the same holds true if we consider multiple eigenvalues and compute the $(k-1)$-th derivative of $\mathrm{adj}(z - A)$ in an eigenvalue with multiplicity at least equal to $k$.
Let then $\lambda$ be an eigenvalue of multiplicity $k$. Clearly $\mathrm{Im}\,\mathrm{adj}(\lambda - A)$ is spanned by $u(\lambda)$, $\mathrm{Im}\,\frac{d}{dz}\mathrm{adj}(z - A)|_{z=\lambda}$ is spanned by $u(\lambda)$ and $u'(\lambda)$, etc. up to $\mathrm{Im}\,\frac{d^{k-1}}{dz^{k-1}}\mathrm{adj}(z - A)|_{z=\lambda}$ which is spanned by $u(\lambda)$ up to $u^{(k-1)}(\lambda)$. As a conclusion we get for such a $\lambda$ that $\mathrm{Im}\,(\mathrm{adj}(z - A), \frac{d}{dz}\mathrm{adj}(z - A), \ldots, \frac{d^{k-1}}{dz^{k-1}}\mathrm{adj}(z - A))|_{z=\lambda}$ is also spanned by $u(\lambda)$ up to $u^{(k-1)}(\lambda)$.
It now follows from the above that $\mathrm{Im}\, P$ is spanned by all the columns of a non-singular confluent Vandermonde matrix. Therefore $P$ has maximal (row) rank and is thus surjective. ■

In the next proposition we use a symmetrizer associated with a polynomial. For a given polynomial $p(z) = z^n + a_1 z^{n-1} + \cdots + a_n$ of degree $n$ we write $S(p)$ to denote the $n \times n$ matrix

$$S(p) = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_1 & 1 & 0 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & 1 & 0 \\ a_{n-1} & & & a_1 & 1 \end{pmatrix}. \tag{2.24}$$

**Proposition 2.5.** *Let $V$ be the confluent Vandermonde matrix associated with all the eigenvalues of $E$ and let $S(e)$ be the symmetrizer associated to the coefficients of the characteristic polynomial of $E$. Assume that $\Gamma$ is such that none of the rows of $V^\top S(e)\Gamma$ is the null vector. Then $R = (\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))$ has rank $q$.*

**Proof.** A sketchy proof, much in the spirit of the proof of Lemma 2.4, is given. We now have to consider all relevant derivatives of $\mathrm{adj}(zI - E)\Gamma\mathrm{adj}(I - zE)^\top$ evaluated at the different eigenvalues $\gamma_i$ and $\tau_i$, call them $\lambda_i$, with their multiplicities $k_i$. It is easy to see (by computing these derivatives and inserting the eigenvalues) that the range of $R$ is the same as the range of $R^0$ which is row block matrix with blocks $R^0_i$ defined by $R^0_i = (u(\lambda_i), u'(\lambda_i), \ldots, u^{(k_i-1)}(\lambda_i))\widetilde{a}(\lambda_i)^\top J\Gamma$.
Since the vectors $u(\lambda_i), u'(\lambda_i), \ldots, u^{(k_i-1)}(\lambda_i)$ with varying $i$ are independent, the only case in which $R^0$ has full row rank is obtained by having all $\widetilde{a}(\lambda_i)^\top J\Gamma$ not equal to the null vector. ■

**Remark.** The condition of this proposition can alternatively be described as follows. Let $e_k(\lambda)$ be the $k - th$ Hörner polynomial associated with the coefficients of $E$ evaluated at an eigenvalue $\lambda$. Put then $\overline{e}(\lambda) = (e_{q-1}(\lambda), \ldots, e_0(\lambda))$. Then none of the rows of $V^\top S(e)\Gamma$ is the null vector iff none of the vectors $\overline{e}(\lambda)$ belongs to the left kernel of $\Gamma$. This condition is satisfied if one chooses $\Gamma$ such that the resulting solution of the Stein equation is the Fisher information matrix. Indeed, with $\Gamma = w_q w_q^\top$, where $w_q$ is the last standard basis vector of $\mathbb{R}^q$ verifies this easily.
We can now formulate an equation that involves Fisher's information matrix and Stien's solution.

**Proposition 2.6.** *Let $l + 1 > 0$ and the matrix $\Gamma$ fulfills the condition given in Proposition 2.5. There exist matrices $\mathcal{A} \in$ Ker $(\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau)\, \mathcal{U}_l(0))$ and $\mathcal{B} \in$ Ker $\left(\widetilde{\mathcal{M}}_r(\gamma)\, \widetilde{\mathcal{M}}_v(\tau)\, \widetilde{\mathcal{M}}_l(0)\right)$ such that the corresponding equations (2.20) and (2.21) hold. The following equality then holds true*

$$R_{qq} \left\{ (\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau)\, \mathcal{U}_l(0))^+ \, G_{bb}(\theta) + \mathcal{A} \right\} \mathcal{R}_{qq} =$$

$$\begin{pmatrix} \mathcal{R}_{(r+v)q} \left\{ \left( \widetilde{\mathcal{M}}_r(\gamma)\, \widetilde{\mathcal{M}}_v(\tau)\, \widetilde{\mathcal{M}}_l(0) \right)^+ S_{bb} + \mathcal{B} \right\} \mathcal{R}_{(r+v)(r+v)} & 0 \\ 0 & I_{q-(r+v)} \otimes \vartheta \end{pmatrix}.$$

We now consider the case where $l + 1 = 0$ or equivalently $q = v + r$. The representations (2.15) and (2.19) become

$$G_{bb}(\theta) = (\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau))\, (\varphi \otimes I_q), \tag{2.25}$$

$$S_{bb} = (\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau))\, (\varphi \otimes I_{r+v}) = (\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau))\, (\varphi \otimes I_q). \tag{2.26}$$

The vector $\varphi$ has the same form as $\vartheta$ in (2.14) but without $\xi_0(0)$. The corresponding blocks composing $\mathcal{M}_r(\gamma)$ and $\mathcal{M}_v(\tau)$ are

$$\mathcal{M}_{s_i}^{(s_i - j)}(z) = \frac{\partial^{s_i - j}}{\partial z^{s_i - j}} \left( \text{adj } (zI - E)\, \Gamma \text{adj } (I - zE)^\top \right)$$

and

$$\mathcal{M}_{\ell_j}^{(\ell_j - k)}(z) = \frac{\partial^{\ell_j - k}}{\partial z^{\ell_j - k}} \left( \text{adj } (zI - E)\, \Gamma \text{adj } (I - zE)^\top \right)$$

respectively. It can be seen that by eliminating the common particular solution of the linear systems (2.25) and (2.26), which is $(\varphi \otimes I_q)$, leads to the equality

$$(\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau))^+ \, G_{bb}(\theta) + \mathcal{Q} = (\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau))^+ \, S_{bb} + \mathcal{T}, \tag{2.27}$$

where $\mathcal{Q} \in$ Ker $(\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau))$ and $\mathcal{T} \in$ Ker $(\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau))$.
In this case the interconnection between $G_{bb}(\theta)$ and $S_{bb}$ can therefore be represented in both directions. The interconnections between the Fisher information matrix and a solution to a Stein equation can now be summarized in the next proposition.

**Proposition 2.7.** *Let $l+1 = 0$ and the matrix $\Gamma$ fulfills the condition given in Proposition 2.5. There exist matrices $\mathcal{Q}\in$ Ker $(\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau))$ and $\mathcal{T}$ $\in$Ker $(\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau))$ such that equation (2.27) holds. The following interconnections then hold true*

$$S_{bb} = (\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau)) \left\{ (\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau))^+ \, G_{bb}(\theta) + \mathcal{Q} \right\},$$

$$G_{bb}(\theta) = (\mathcal{U}_r(\gamma)\, \mathcal{U}_v(\tau)) \left\{ (\mathcal{M}_r(\gamma)\, \mathcal{M}_v(\tau))^+ \, S_{bb} + \mathcal{T} \right\}.$$

In Section 3 a detailed description of Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ and Ker $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))$ will be given.It is clear that it is not necessary to impose any condition on $\Gamma$ when $S_{bb}$ is expressed in terms of $G_{bb}(\theta)$.

We now study the case $l+1<0$ or equivalently $q<v+r$. In this case we obtain from (2.15) and (2.19)

$$\begin{aligned} G_{bb}(\theta) &= \left(\widetilde{\mathcal{U}}_r(\gamma)\ \widetilde{\mathcal{U}}_v(\tau)\right)(\varphi\otimes I_q), & (2.28)\\ S_{bb} &= (\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))\,(\varphi\otimes I_{r+v}). & (2.29)\end{aligned}$$

The block components of $\widetilde{\mathcal{U}}_r(\gamma)$ and $\widetilde{\mathcal{U}}_v(\tau)$ are composed by $\widetilde{\mathcal{U}}_{s_i}^{(s_i-j)}(z)=\frac{\partial^{s_i-j}}{\partial z^{s_i-j}}\left(u_q(z)u_q^{*\top}(z)z^{l+1}\right)$ and $\widetilde{\mathcal{U}}_{\ell_j}^{(\ell_j-k)}(z)=\frac{\partial^{\ell_j-k}}{\partial z^{\ell_j-k}}\left(u_q(z)u_q^{*\top}(z)z^{l+1}\right)$ respectively and are evaluated for $z=\gamma_j$ and $z=\tau_j$. We extract the desired particular solution of the linear system of equations (2.28) and (2.29), to obtain

$$(\varphi\otimes I_q)=\left(\widetilde{\mathcal{U}}_r(\gamma)\ \widetilde{\mathcal{U}}_v(\tau)\right)^{+}G_{bb}(\theta)+\mathcal{D}, \tag{2.30}$$

where $\mathcal{D}\in$ Ker $\left(\widetilde{\mathcal{U}}_r(\gamma)\ \widetilde{\mathcal{U}}_v(\tau)\right)$ and

$$(\varphi\otimes I_{r+v})=(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))^{+}\,S_{bb}+\mathcal{E}, \tag{2.31}$$

where $\mathcal{E}\in$ Ker $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))$.
An equation involving the Fisher information matrix and a solution to Stein's equation, for the case considered, is given in the next proposition.

**Proposition 2.8.** *Let $l+1<0$ and the matrix $\Gamma$ fulfills the condition given in Proposition 2.5. There exist matrices $\mathcal{E}\in$ Ker $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))$ and $\mathcal{D}\in$ Ker $\left(\widetilde{\mathcal{U}}_r(\gamma)\ \widetilde{\mathcal{U}}_v(\tau)\right)$ such that the respective equations (2.31) and (2.30) hold and $\mathcal{R}$ is a permutation matrix. We then obtain the following equation involving $S_{bb}$ and $G_{bb}(\theta)$.*

$$R_{(r+v)(r+v)}\left\{(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))^{+}\,S_{bb}+\mathcal{E}\right\}\mathcal{R}_{(r+v)(r+v)}=$$

$$\begin{pmatrix} \mathcal{R}_{q(r+v)}\left\{\left(\widetilde{\mathcal{U}}_r(\gamma)\ \widetilde{\mathcal{U}}_v(\tau)\right)^{+}G_{bb}(\theta)+\mathcal{D}\right\}\mathcal{R}_{qq} & 0\\ 0 & I_{r+v-q}\otimes\varphi \end{pmatrix}.$$

**Example 2.9.** In this example a right inverse of the coefficient matrix $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ is set forth. Using the information of the ARMAX process given in Example 2.1 yields the following representation for the coefficient matrix $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$

$$\mathcal{U}_r(\gamma)=\left(\frac{\partial}{\partial z}\begin{pmatrix} z^2 & z & 1\\ z^3 & z^2 & z\\ z^4 & z^3 & z^2\end{pmatrix},\begin{pmatrix} z^2 & z & 1\\ z^3 & z^2 & z\\ z^4 & z^3 & z^2\end{pmatrix}\right)_{z=\gamma}\quad\text{and}\quad \mathcal{U}_v(\tau)=\begin{pmatrix} z^2 & z & 1\\ z^3 & z^2 & z\\ z^4 & z^3 & z^2\end{pmatrix}_{z=\tau}.$$

The right inverse of $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ in the proof of Proposition 2.3 is

$$(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))_R^{-}=\left((\mathcal{W}_{r,\,v}(\gamma,\tau))^{-1}\otimes w_3\right),$$

where

$$w_3 = (0,0,1)^\top \quad \text{and} \quad (\mathcal{W}_{r,\,v}(\gamma,\tau)) = \left\{ \left(\frac{\partial}{\partial z} u_3(z)\ u_3(z)\right)_{z=\gamma} (u_3(z))_{z=\tau} \right\} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & \gamma & \tau \\ 2\,\gamma & \gamma^2 & \tau^2 \end{pmatrix}.$$

We eventually obtain

$$(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))_R^- = \frac{1}{(\gamma-\tau)^2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \gamma\ \tau\ (\gamma-\tau) & -\left(\gamma^2-\tau^2\right) & (\gamma-\tau) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ (\gamma-\tau)^2 - \gamma^2 & 2\,\gamma & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \gamma^2 & -2\,\gamma & 1 \end{pmatrix}.$$

## 3. Kernel description

In this section algorithms for the kernels of the coefficient matrices in the linear system of equations (2.15) and (2.19) are described.

### 3.1. General case

We first focus on the null space appearing in Proposition 2.6, namely Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))$. Since the matrix blocks which constitute $\mathcal{U}_r(\gamma)$, $\mathcal{U}_v(\tau)$ and $\mathcal{U}_l(0)$ are evaluated at distinct roots, we then have the property Im $(\mathcal{U}_\nu(\sigma)) \cap$ Im $(\mathcal{U}_\mu(\rho)) = \{0\}$ for all the distinct eigenvalues $\sigma$ and $\rho$ (with corresponding algebraic multiplicity $\nu$ and $\mu$). Consequently, the subspace Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))$ can be decomposed into a direct sum Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0)) =$ Ker $(\mathcal{U}_r(\gamma)) \oplus$ Ker $(\mathcal{U}_v(\tau)) \oplus$ Ker $(\mathcal{U}_l(0))$. A similar decomposition can also be applied to the subspaces on the right-hand side, to obtain Ker $(\mathcal{U}_r(\gamma)) = \bigoplus_{i=1}^{r_0}$ Ker $(\mathcal{U}_{s_i}(\gamma_i))$, Ker $(\mathcal{U}_v(\tau)) = \bigoplus_{j=1}^{v_0} \mathrm{Ker}\big(\mathcal{U}_{\ell_j}(\tau_j)\big)$. This property follows from the next lemma.

**Lemma 3.1.** *Consider two matrices $A$ and $B$ with appropriate dimensions, then*

$$\textit{Im } A \cap \textit{Im } B = \{0\} \textit{ iff Ker } (A\ B) = \begin{pmatrix} \textit{Ker } A \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ \textit{Ker } B \end{pmatrix}.$$

**Proof.** When moving from right to left it follows from the dimension rules for $A$, $B$ and $(A,B)$ that

$$\text{dim Im } A + \text{dim Im } B - \text{dim Im } (AB) + \text{dim Ker } A + \text{dim Ker } B - \text{dim Ker } (A,B) = 0.$$

By assumption we have

$$\text{dim Ker } A + \text{dim Ker } B - \text{dim Ker}(A,B) = 0$$

so that

$$\text{dim Im } A + \text{dim Im } B = \text{dim Im } (AB).$$

Because Im $A+$ Im $B=$ Im $(AB)$, we must therefore have Im $A\cap$ Im $B=\{0\}$. From left to right, we assume $\begin{pmatrix} x \\ y \end{pmatrix} \in$ Ker $(AB)$, this implies $Ax+By=0$ and since Im $A\cap$ Im $B=\{0\}$ we have $Ax=0$ and $By=0$. ∎

Since the individual null spaces have the same representation, it therefore suffices to specify the null space evaluated at one single root. For that purpose we represent a root by $\sigma$ with algebraic multiplicity $\nu+1$. In the next sections an algorithm for Ker $(\mathcal{U}_\nu(\sigma))$ is described and is followed by properties of Ker $(\mathcal{M}_\nu(\sigma))$.

### 3.1.1. An algorithm for computing Ker $(\mathcal{U}_\nu(\sigma))$

In this section we shall adapt the notations used in the previous section accordingly. Consider $u_q(z) = (1, z, \ldots, z^{q-1})^\top$ and $v_p(z) = z^{p-1}u_p(z^{-1})^\top$. Define the $q\times p(n+1)$ matrix $U_{nqp}(z) = (U^n_{qp}, \cdots, U^0_{qp})$ (this is equivalent with $\mathcal{U}_\nu(z)$) by

$$U^k_{qp}(z) = \left(\frac{d}{dz}\right)^k u_q(z)v_p(z).$$

We will give an expression for Ker $U_{nqp}(z)$. Let $x$ be vector belonging to this kernel and decompose $x$ as $x^\top = (x_0^\top, \ldots, x_n^\top)$, with the $x_k \in \mathbb{R}^p$. Then

$$\begin{aligned}
U_{nqp}(z)x &= \sum_{k=0}^{n} (\frac{d}{dz})^k u_q(z)v_p(z)x_{n-k} \\
&= \sum_{k=0}^{n}\sum_{j=0}^{k} \binom{k}{j} u_q^{(j)}(z)v_p^{(k-j)}(z)x_{n-k} \\
&= \sum_{j=0}^{n} u_q^{(j)}(z) \sum_{k=j}^{n} \binom{k}{j} v_p^{(k-j)}(z)x_{n-k} \\
&= \sum_{j=0}^{n\wedge(q-1)} u_q^{(j)}(z) \sum_{k=j}^{n} \binom{k}{j} v_p^{(k-j)}(z)x_{n-k}.
\end{aligned}$$

Since the vectors $u_q^{(j)}(z)$ are independent as long as $j\le q-1$, we see that $U_{nqp}(z)x=0$ iff for all $j\le (q-1)\wedge n$ we have

$$\sum_{k=j}^{n} \binom{k}{j} v_p^{(k-j)}(z)x_{n-k} = 0. \tag{3.1}$$

(Notice that in this summation we only have non-zero contributions for $k\le (j+p-1)\wedge n$.)
Thus we consider a system of $(q-1)\wedge n+1$ equations of type (3.1). Clearly, this system is triangular, which leads to a recursive solution procedure.
We introduce some more notation. Let $K_p(z)$ be a $p\times(p-1)$ matrix whose columns span Ker $v_p(z)$ (later on we will specify a certain choice for $K_p(z)$). We proceed in steps.

First we consider the case in which $n<q$, so we have a system of $n+1$ equations.
Set $j=n$. Then the corresponding equation becomes $v_p(z)x_0 = 0$. Hence $x_0 = K_p(z)\gamma_0$ for an arbitrary vector $\gamma_0\in\mathbb{R}^{p-1}$.
Consider now (with $x_0$ given above) the equation for $j=n-1$:

$$v_p(z)x_1 + nv_p'(z)x_0 = 0.$$

A particular solution of this equation is $x_1 = -nl_pv_p'(z)x_0$, with $l_p$ the last standard basis vector of $\mathbb{R}^p$ and hence the general solution is given by $x_1 = -nl_pv_p'(z)x_0 + K_p(z)\gamma_1$ with arbitrary $\gamma_1$, so

$x_1 = K_p(z)\gamma_1 - nl_p v_p'(z)K_p(z)\gamma_0$.
Continuing this way, we look at the equation for $j = n-2$. It is

$$v_p(z)x_2 + (n-1)v_p'(z)x_1 + \frac{1}{2}n(n-1)v_p''(z)x_0 = 0.$$

A particular solution is given by

$$\begin{aligned} x_2 &= -l_p\left((n-1)v_p'(z)x_1 + \frac{1}{2}n(n-1)v_p''(z)x_0\right) \\ &= -l_p\left((n-1)v_p'(z)(-nl_p v_p'(z)K_p(z)\gamma_0 + K_p(z)\gamma_1) + \frac{1}{2}n(n-1)v_p''(z)x_0\right) \\ &= -l_p\left((n-1)v_p'(z)K_p(z)\gamma_1 + \frac{1}{2}n(n-1)v_p''(z)K_p(z)\gamma_0\right), \end{aligned}$$

where we used in the last equality that $v_p'(z)l_p = 0$. The general solution now becomes

$$x_2 = K_p(z)\gamma_2 - l_q\left((n-1)v_q'(z)K_p(z)\gamma_1 + \frac{1}{2}n(n-1)v_q''(z)K_p(z)\gamma_0\right).$$

Proceeding in this way, one obtains the following recursion for the $x_k$ and then its explicit form.

$$x_{k+1} = K_p(z)\gamma_{k+1} - \sum_{j=1}^{k}\binom{n-k+j}{j}l_p v_p^{(j)}(z)x_{k+1-j} \tag{3.2}$$

$$x_k = K_p(z)\gamma_k - \sum_{j=1}^{k}\binom{n-k+j}{j}l_p v_p^{(j)}(z)K_p(z)\gamma_{k-j}. \tag{3.3}$$

If we put all the $x_k$ underneath each other, we get

$$x = L_n(z)(I_{n+1}\otimes K_p(z))\gamma, \tag{3.4}$$

with $L_n(z) \in \mathbb{R}^{(n+1)p\times(n+1)p}$ the lower triangular matrix

$$\begin{pmatrix} I_p & 0 & & & & \\ -\binom{n}{1}l_p v_p^{(1)}(z) & I_p & 0 & & & \\ -\binom{n}{2}l_p v_p^{(2)}(z) & -\binom{n-1}{1}l_p v_p^{(1)}(z) & I_p & 0 & & \\ \vdots & & & & & \\ -l_p v_p^{(n)}(z) & & & & -l_p v_p^{(1)}(z) & I_p \end{pmatrix}. \tag{3.5}$$

Clearly dim Ker $U_{nqp}(z) = (n+1)(p-1)$.
Since obviously, the image space of $U_{nqp}(z)$ is spanned by the vectors $u_q^{(j)}(z)$, for $j = 0,\ldots,n$ (recall that $n < q$), it has dimension $n+1$. This is in agreement with the dimension rule: $(n+1)(p-1)+n+1$ is the number of coulmns of $U_{nqp}(z)$.

A convenient choice of $K_p(z)$ is

$$\begin{pmatrix} -1 & 0 & & 0 \\ z & -1 & & \\ 0 & z & \ddots & 0 \\ & \ddots & \ddots & -1 \\ 0 & & 0 & z \end{pmatrix}.$$

In particular the computation of the products $v_p^{(j)}(z)K_p(z)$ now becomes easy. Differentiate $v_p(z)K_p(z)$ $j$ times. Since $K$ has zero derivatives of order greater than 1 and since $v_p(z)K_p(z) = 0$, we get $v_p^{(j)}(z)K_p(z) = -jv_p^{(j-1)}(z)K'(z)$. But this is nothing else than the vector $-jv_p^{(j-1)}(z)$ without its first element.

For the case in which $n \geq q$, a similar procedure as above has to be followed. The prime difference is that we now consider the set of $q$ equations (3.1), for $j = 0, \ldots, q-1$. Consider first the equation for $j = q-1$:

$$\sum_{k=q-1}^{n} \binom{k}{q-1} v_p^{(k-q+1)}(z)x_{n-k} = 0.$$

To get a solution we choose the $x_0, \ldots, x_{n-q}$ completely free, say $x_k = \beta_k$ with $\beta_k \in \mathbb{R}^p$. Then we get for $x_{n-q+1}$ the general solution

$$x_{n-q+1} = -l_p \sum_{k=q}^{n} \binom{k}{q-1} v_p^{(k-q+1)}(z)\beta_{n-k} + K_p(z)\gamma_{n-q+1},$$

with $\gamma_{n-q+1}$ an arbitrary vector in $\mathbb{R}^{p-1}$. Continuing this way as in the case with $n < q$ we now get the solution $x$ given by

$$x = \begin{pmatrix} I_{(n-q+1)p} & 0_{(n-q+1)p \times qp} \\ M(z) & L_q(z)(I_q \otimes K_p(z)) \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix}, \tag{3.6}$$

with $L_q(z) \in \mathbb{R}^{qp \times q(p-1)}$ like the matrix $L_n(z)$ above, $M(z) \in \mathbb{R}^{qp \times (n-q+1)p}$ defined by

$$M(z) = \begin{pmatrix} -\binom{n}{q-1} l_p v_p^{(n-q+1)}(z) & \cdots & \binom{q}{q-1} l_p v_p^{(1)}(z) \\ \vdots & & \vdots \\ -\binom{n}{0} l_p v_p^{(n)})(z) & \cdots & \binom{q}{0} l_p v_p^{(q)}(z) \end{pmatrix}$$

and $\beta = (\beta_0^\top, \ldots, \beta_{n-q}^\top)^\top$, $\gamma = (\gamma_{n-q+1}^\top, \ldots, \gamma_n^\top)^\top$.

Since the image of $U_{nqp}(z)$ is now spanned by the vectors $u_q^{(j)}(z)$, for $j = 0, \ldots, q-1$ (recall that $n \geq q$), it has dimension $q$. For the kernel we now have that its dimension is $(n-q+1)p$ (from the first components) plus $q(p-1)$ (from the other other components), $np + p - q$ in total. Notice again that this is in agreement with the dimension rule.

For constructing the subspace Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)\ \mathcal{U}_l(0))$ one considers the direct sum of the kernels of the $\mathcal{U}_{\nu_i}(\sigma_i)$ for all the distinct eigenvalues $\sigma_i$ and hence it's dimension is the sum of the dimensions of the summands.

**Example 3.2.** In this example the implementation of the algorithm just developed will be illustrated. Consider the coefficient matrix $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ used in Example 2.9. A vector in the subspace Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)) = \text{Ker}\ \mathcal{U}_r(\gamma) \oplus \text{Ker}\ \mathcal{U}_v(\tau)$ is derived. The multiplicity equal to one yields according to (3.9)

$$\text{Ker}\ \mathcal{U}_v(\tau) = \text{span} \begin{pmatrix} -z^{-2}\ u_2^\top(z) \\ J_2 \end{pmatrix}_{z=\tau} = \text{span} \begin{pmatrix} -z^{-2} & -z^{-1} \\ 0 & 1 \\ 1 & 0 \end{pmatrix}_{z=\tau}.$$

The parameters necessary for constructing the null space Ker $\mathcal{U}_r(\gamma)$ are, $n = k = 1$ and $p = q = 3$. This results in the equations

$$\begin{aligned} x_0 &= K_3(z)\gamma_0 \\ x_1 &= -l_3 v_3'(z) K_3(z)\gamma_0 + K_3(z)\gamma_1 \end{aligned}$$

that belong to the subspace Ker $\mathcal{U}_r(\gamma)$, compactly written

$$x = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} I_3 & 0 \\ -l_3 v_3'(z) & I_3 \end{pmatrix} (I_2 \otimes K_3(z)) \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix}.$$

An explicit representation yields

$$K_3(z) = \begin{pmatrix} -1 & 0 \\ z & -1 \\ 0 & z \end{pmatrix}, v_3(z) = (z^2, z, 1), l_3 = (0, 0, 1)^\top$$

and $\gamma_0$ and $\gamma_1$ are arbitrary. Let us denote the components of $\gamma_0$ and $\gamma_1$ by $\left(\gamma_0^1\ \gamma_0^2\right)^\top$ and $\left(\gamma_1^1\ \gamma_1^2\right)^\top$ respectively, so that the vector belonging to the subspace Ker $\mathcal{U}_r(\gamma)$ can be expressed as

$$x = \begin{pmatrix} -1 & 0 & 0 & 0 \\ z & -1 & 0 & 0 \\ 0 & z & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & z & -1 \\ z & 1 & 0 & z \end{pmatrix}_{z=\gamma} \begin{pmatrix} \gamma_0^1 \\ \gamma_0^2 \\ \gamma_1^1 \\ \gamma_1^2 \end{pmatrix} = \begin{pmatrix} -\gamma_0^1 \\ \gamma_0^1\ z - \gamma_0^2 \\ \gamma_0^2\ z \\ -\gamma_1^1 \\ \gamma_1^1\ z - \gamma_1^2 \\ \left(\gamma_0^1 + \gamma_1^2\right)\ z + \gamma_0^2 \end{pmatrix}_{z=\gamma}.$$

According to the results obtained in Section 3.1.1. it can be concluded that dim Ker $\mathcal{U}_r(\gamma) = 4$ and dim Ker $\mathcal{U}_v(\tau) = 2$, consequently dim Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)) = 6$. It is then clear that the matrix $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ is surjective since dim Im $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau)) = 3$, a confirmation of Proposition 2.3.
Since $\gamma_0$ and $\gamma_1$ are arbitrary, we choose $\gamma_0 = (1, 1)^\top$ and $\gamma_1 = (2, 3)^\top$ so that a choice for a $9 \times 3$ matrix $\mathcal{Q}$ such that $\mathcal{Q} \in$ Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$, can be expressed as

$$\mathcal{Q} = \begin{pmatrix} -1 & 0 & 0 \\ \gamma - 1 & 0 & 0 \\ \gamma & 0 & 0 \\ -2 & 0 & 0 \\ 2\,\gamma - 3 & 0 & 0 \\ 4\,\gamma + 1 & 0 & 0 \\ 0 & -\tau^{-2} & -\tau^{-1} \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

### 3.1.2. An algorithm for computing Ker ( $\mathcal{M}_\nu(\sigma)$)

In this subsection we study the null space Ker ( $\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau)\ \mathcal{M}_l(0)$). The subspaces Im $(\mathcal{M}_r(\gamma))$, Im $(\mathcal{M}_v(\tau))$ and Im $(\mathcal{M}_l(0))$ have the property formulated in Lemma 3.1 and this can be justified since the matrix blocks which form $\mathcal{M}_r(\gamma)$, $\mathcal{M}_v(\tau)$ and $\mathcal{M}_l(0)$ are evaluated at dinstinct roots. Therefore, we have Ker $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau)\ \mathcal{M}_l(0))$ = Ker $(\mathcal{M}_r(\gamma)) \oplus$ Ker $(\mathcal{M}_v(\tau)) \oplus$ Ker $(\mathcal{M}_l(0))$ with Ker $(\mathcal{M}_r(\gamma)) = \bigoplus_{i=1}^{r_0}$ Ker $(\mathcal{M}_{s_i}(\gamma_i))$, Ker $(\mathcal{M}_v(\tau)) = \bigoplus_{j=1}^{v_0} \mathrm{Ker}\big(\mathcal{M}_{\ell_j}(\tau_j)\big)$ and since $\mathcal{M}_l(0) = 0$, we have Ker $\mathcal{M}_l(0) = \mathbb{C}^{r+v} \oplus \mathbb{C}^{r+v} \oplus \cdots \oplus \mathbb{C}^{r+v}$. Considering that the individual null spaces have the same structure, it is therefore sufficient to describe a null space evaluated at one single root.

Let $C$ be a $m \times m$ companion matrix of the form (2.22). Let $c(z)$ be the $m$-vector of polynomials $c_i(z)$ defined by $c_0(z) = 1$ and $c_i(z) = zc_{i-1}(z) + c_i$, so $c(z) = (c_0(z), \ldots, c_{m-1}(z))^\top$. Let $\sigma$ be one of its eigenvalues and assume that it has algebraic multiplicity equal to $\nu + 1$. We have of course $\nu + 1 \leq m$. Let $\Gamma$ be an arbitrary $m \times m$ matrix. We will also use $f(\sigma) = c(\sigma)^\top J\Gamma$ and $A(\sigma)$ =adj$(I - \sigma C)^\top$. We study the $m \times m(\nu + 1)$ matrix

$$\mathcal{M}(\sigma) = \Big( \ \mathcal{M}^{(\nu)}(z), \quad \mathcal{M}^{(\nu-1)}(z), \quad \ldots \quad , \mathcal{M}^{(0)}(z) \ \Big)_{z=\sigma},$$

where each $m \times m$ block $\mathcal{M}^{(j)}(z)$ is given by

$$\mathcal{M}^{(j)}(z) = \frac{\partial^j}{\partial z^j}\Big(\mathrm{adj}(zI - C)\Gamma\mathrm{adj}(I - zC)^\top\Big).$$

In the next proposition an explicit representation for $\mathcal{M}^{(j)}(z)$ is given.

**Proposition 3.3.** *For $n \leq \nu$ we have*

$$\mathcal{M}^{(n)}(\sigma) = \sum_{k=0}^{n} \binom{n}{k} u^{(n-k)}(\sigma) \sum_{j=0}^{k} \binom{k}{j} f^{(j)}(\sigma) A^{(k-j)}(\sigma). \tag{3.7}$$

**Proof.** First we compute the derivatives of adj$(zI - C)$. Using (2.23), we see that the terms that involves the characteristic polynomial of $C$ vanish for $z = \sigma$ and $k \leq \nu$. So, we get from the Leibniz rule

$$\frac{\partial^k}{\partial z^k}\mathrm{adj}(zI - C)|_{z=\sigma} = \sum_{j=0}^{k} \binom{k}{j} u^{(j)}(\sigma) c^{(k-j)}(\sigma)^\top J.$$

Recall that $\mathcal{M}(z)$ =adj$(zI - C)\Gamma A(z)$. Applying the Leibniz rule once more, we obtain

$$\mathcal{M}^{(n)}(z) = \sum_{k=0}^{n} \binom{n}{k} \frac{\partial^k}{\partial z^k}\mathrm{adj}(zI - C)\Gamma \frac{\partial^{n-k}}{\partial z^{n-k}} A(z).$$

Insertion of the previously found derivative for adj$(zI - C)$ yields

$$\mathcal{M}^{(n)}(\sigma) = \sum_{k=0}^{n} \binom{n}{k} \sum_{j=0}^{k} \binom{k}{j} u^{(j)}(\sigma) f^{(k-j)}(\sigma) A^{(n-k)}(\sigma),$$

which is equivalent (rearrange the summation) to (3.7). ■

An appropriate factorization is applied and is summarized in the following proposition.

**Proposition 3.4.** *The matrix $\mathcal{M}(\sigma)$ can be factored as the product*

$$\mathcal{M}(\sigma) = UHA, \tag{3.8}$$

*where the matrices $U$, $H$ and $A$ are as follows. $U = (u(\sigma), u'(\sigma), \ldots, u^{(\nu)}(\sigma))$. $H$ has a block-triangular structure in which its $i$-th column ($i = 1, \ldots, \nu+1$) has $j$-th element given by the row vector $\binom{\nu+1-j}{i-1} f^{(\nu+1-j)}(\sigma)$. The elements become zero for $i + j > \nu + 2$. The matrix $A$ is invertible, block-triangular and has as $ij$-block element the matrix $\binom{\nu+1-j}{i-j} A^{(i-j)}(\sigma)$. The $ij$-th element of this matrix becomes zero if $j > i$. The matrix $U$ has full column rank.*

**Proof.** To each of the entries of $\mathcal{M}(\sigma)$ we apply the preceding proposition. Then the next step is to compute the factors by which have to postmultiply the $u^{(l)}(\sigma)$. Consider first $u(\sigma)$. It is postmultiplied by the row vectors $\sum_{j=0}^{k} f^{(j)}(\sigma)A^{(k-j)}(\sigma)$, for $k = \nu$ down to 0 to get its contribution to each of the $\mathcal{M}^{(k)}(\sigma)$. Notice that $\sum_{j=0}^{k} f^{(j)}(\sigma)A^{(k-j)}(\sigma)$ is the product of the first row of $H$ and the $k$-th column of $A$. The contributions of the other $u^{(l)}(\sigma)$ can be treated similarly. On the diagonal of the matrix $A$ we find the invertible matrices $A(\sigma)$. Because of its triangular structure the matrix $A$ is invertible. That $U$ has full column rank is obvious. ■

The sizes of the above matrices are as follows. $U$ is a $m \times m(\nu+1)$ matrix, $H$ is of size $(\nu+1) \times (\nu+1)m$ and $A$ has dimensions $(\nu+1)m \times (\nu+1)m$.

We are interested in Ker $\mathcal{M}(\sigma)$. Let $x \in$ Ker $\mathcal{M}(\sigma)$, so $\mathcal{M}(\sigma)x = 0$. Since $A$ is invertible, we can write $x = Ay$ and $y = A^{-1}x$. So we look at $UHy = 0$. But since $U$ has full column rank, this is equivalent to $Hy = 0$. Below, we will investigate in some detail the structure of Ker $H$. As a side remark we mention that for an explicit expression for Ker $H$, we also need the inverse of $A$. Because of the block triangular structure the block-elements of this inverse are products of the derivatives of $A(z)$ and $A(z)^{-1}$. But $A(z)^{-1} = \det(I - zC)^{-1} \times (I - zC^\top)$ and so this causes no computational problems. To compute $\text{adj}(I - zC)$ and its derivatives we can use an expression similar to (2.23).

A condition for specifying the dimension of the kernel of $\mathcal{M}(\sigma)$ is given and can be seen as an alternative to Proposition 2.5.

**Theorem 3.5.** *The rank of the matrix $\mathcal{M}(\sigma)$ is equal to $\nu+1-k(\sigma)$, where $k(\sigma) = \min\{j : f^{(j)}(\sigma) \neq 0\}$, with the understanding that $k(\sigma) = \nu+1$ if all $f^{(j)}(\sigma)$ are zero. Furthermore,* dim *Ker$\mathcal{M}(\sigma) = (m-1)(\nu+1) + k(\sigma)$.*

**Proof.** From the above discussion it is clear that the rank of $\mathcal{M}(\sigma)$ is equal to the rank of $H$. From the triangular structure of $H$, which would be block-Hankel if we ignore the binomial coefficients, the result on the rank is obvious. The dimension of the kernel follows form the addition rule for the dimensions of kernel and image space. ■

It is hard to give an explicit description of a basis of the kernel of $H$. However, for the special case of $k(\sigma) = 1$, there is a neat expression available. This case is motivated by the case where we deal with the Fisher information matrix as mentioned earlier in this paper. In this case we use $\Gamma = w_m w_m^\top$, where $w_m$ is the last basis vector of $\mathbb{R}^m$. Since now $J\Gamma = w_1 w_m^\top$, we get $f(\sigma) = w_m^\top$, whatever $\sigma$ and hence $k(\sigma) = 1$ for all $\sigma$. So we assume now that $k(\sigma) = 1$. Let $H_0$ be a matrix of size $m \times (m-1)$ whose columns span Ker $f(\sigma)$. Let $p_0^* = p_0^*(\sigma) \in \mathbb{R}^m$ be a column vector such that $f(\sigma)p_0^* = 1$. Such a vector obviously exists, when $k(\sigma) = 1$.

Let $y \in$ Ker $H$ and write $y = (y_0^\top, \ldots, y_\nu^\top)^\top$, with $y_i \in \mathbb{R}^m$. Then we have a recursive set of equations for the $y_i$. We briefly outline the procedure how to compute these $y_i$. The first equation we solve is

$f(\sigma)y_0 = 0$. We can express $y_0$ as $y_0 = H_0\eta_0$, where $\eta_0$ is a free vector in $\mathbb{R}^{m-1}$. The next equation we solve is $f'(\sigma)y_0 + f(\sigma)y_1 = 0$, whose general solution can be written as $y_1 = -p_0^* f'(\sigma)y_0 + H_0\eta_1$, where $\eta_1$ is another free vector in $\mathbb{R}^{m-1}$. Continuing this way, we can express the whole vector $y$ as a certain matrix times the vector that is obtained by stacking the $\eta_i$ into one vector of dimension $(\nu+1)(m-1)$. This matrix doesn't look very nice though, but there is still something to say.
Let $\Delta_0 = I_{\nu+1} \otimes H_0$. We need the matrix $K = K(\sigma)$ (of block-triangular structure) whose elements are $m \times m$ matrices and where the $ij$-th element is specified as $I$ if $i = j$, zero for $j > i$ and for $i > j$ we have $K_{ij} = \binom{\nu+2-j}{i-j} p_0^* f^{(i-j)}(\sigma)$. We observe that $K$ is invertible and that $JK$ has a structure similar to the one of $H$. Then the columns of the matrix $K^{-1}\Delta_0$ span Ker $H$. Notice that this matrix has $(\nu+1)(m-1)$ independent columns.
If $f(\sigma) = 0$, but $f'(\sigma) \neq 0$, the procedure is similar. The recursive set of equations doesn't contain $f(\sigma)$ anymore. But we can find $p_1^*$ such that $f'(\sigma)p_1^* = 1$ and we proceed along the same lines as in the previous case, upon noticing that we now need a matrix whose columns span Ker $f'(\sigma)$. The vector $y_\nu$ is entirely free in the present case. The other cases can be treated similarly.

### 3.2. Special case

In this section we compute the kernels of the coefficient matrices in (2.15) and (2.19) for the case when the zeros of the polynomials $a(z)$, $b(z)$, $c(z)$ and $h(z)$ all have multiplicity equal to one. First, we consider the subspace

$$\text{Ker } (\mathcal{U}_r(\gamma)\; \mathcal{U}_v(\tau)\; \mathcal{U}_l(0)) = \text{Ker } (\mathcal{U}_r(\gamma)) \oplus \text{Ker } (\mathcal{U}_v(\tau)) \oplus \text{Ker } (\mathcal{U}_l(0)),$$

with

$$\text{Ker } (\mathcal{U}_r(\gamma)) = \bigoplus_{i=1}^{r} \text{Ker } (\mathcal{U}_i(\gamma_i)),\ \text{Ker } (\mathcal{U}_v(\tau)) = \bigoplus_{j=1}^{v} \text{Ker } (\mathcal{U}_j(\tau_j)).$$

It is sufficient to represent one case, to obtain

$$\text{Ker } (\mathcal{U}_i(\gamma_i)) = \text{span } \begin{pmatrix} -z^{-(q-1)} u_{q-1}^\top(z) \\ J_{q-1} \end{pmatrix}_{z=\gamma_i}, \tag{3.9}$$

where $J_{q-1}$ is the $(q-1)$ backward or antidiagonal identity matrix and dim Ker $(\mathcal{U}_i(\gamma_i)) = q-1$.
A similar representation holds for Ker $(\mathcal{U}_j(\tau_j))$ when $z = \tau_j$. Observe the properties

$$\text{Ker } (\mathcal{U}_\delta(0)) = \text{Ker } \left(\frac{\partial^\delta}{\partial z^\delta}\left(u_q(z)u_q^{*\top}(z)\right)\right)_{z=0} = \text{span}\begin{cases} \begin{pmatrix} J_{q-1-\delta} \\ 0_{1+\delta} \end{pmatrix} & \text{for } \delta = 0,\ 1, \ldots, q-1 \\ \begin{pmatrix} 0_{2q-1-\delta} \\ J_{\delta-(q-1)} \end{pmatrix} & \text{for } \delta = q,\ q+1, \ldots, 2q-2 \end{cases}$$

and

$$\text{dim Ker } \left(\frac{\partial^\delta}{\partial z^\delta}\left(u_q(z)u_q^{*\top}(z)\right)\right)_{z=0} = \begin{cases} (q-1)-\delta & \text{for } \delta = 0, 1, \ldots, q-1 \\ \delta-(q-1) & \text{for } \delta = q, q+1, \ldots, 2q-2. \end{cases}$$

The null spaces which compose the subspace Ker $(\mathcal{M}_r(\gamma)\; \mathcal{M}_v(\tau)\; \mathcal{M}_l(0))$ are obtained according to

$$\text{Ker } (\mathcal{M}_r(\gamma)\; \mathcal{M}_v(\tau)\; \mathcal{M}_l(0)) = \text{Ker } (\mathcal{M}_r(\gamma)) \oplus \text{Ker } (\mathcal{M}_v(\tau)) \oplus \text{Ker } (\mathcal{M}_l(0)),$$

with

$$\text{Ker } (\mathcal{M}_r(\gamma)) = \bigoplus_{i=1}^{r} \text{Ker } (\mathcal{M}_i(\gamma_i)),\ \text{Ker } (\mathcal{M}_v(\tau)) = \bigoplus_{j=1}^{v} \text{Ker}(\mathcal{M}_j(\tau_j))$$

and as in the general case Ker $\mathcal{M}_l(0) = \mathbb{C}^{r+v} \oplus \mathbb{C}^{r+v} \oplus \cdots \oplus \mathbb{C}^{r+v}$. Since there is an equivalent functional form for all the subspaces, it suffices to consider the case

$$\text{Ker }(\mathcal{M}_r(\gamma)) = \bigoplus_{i=1}^{r} \text{Ker }(\mathcal{M}_i(\gamma_i)).$$

The factorization

$$\mathcal{M}_r(\gamma) = \mathcal{M}_r^{(1)}(\gamma)\mathcal{M}_r^{(2)}(\gamma)$$

is applied, where

$$\mathcal{M}_r^{(1)}(\gamma) = \Big(\text{adj}\,(zI-E)_{z=\gamma_1}, \text{adj}\,(zI-E)_{z=\gamma_2}, \dots, \text{adj}\,(zI-E)_{z=\gamma_r}\Big)$$

and

$$\mathcal{M}_r^{(2)}(\gamma) = \begin{pmatrix} \Gamma\Big(\text{adj}\,(I-zE)^\top z^{l+1}\Big)_{z=\gamma_1} & 0 & \dots & 0 \\ 0 & \Gamma\Big(\text{adj}\,(I-zE)^\top z^{l+1}\Big)_{z=\gamma_2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \Gamma\Big(\text{adj}\,(I-zE)^\top z^{l+1}\Big)_{z=\gamma_r} \end{pmatrix}.$$

Since the blocks composing $\mathcal{M}_r^{(2)}(\gamma)$ are square invertible matrices (for simplicity we assume that the matrix $\Gamma$ is invertible), we then have

$$\text{Ker }(\mathcal{M}_r(\gamma)) = \Big(\mathcal{M}_r^{(2)}(\gamma)\Big)^{-1} \text{ Ker }\Big(\mathcal{M}_r^{(1)}(\gamma)\Big).$$

Using a similar argument as in Lemma 3.1, we have the following direct sum

$$\text{Ker }\Big(\mathcal{M}_r^{(1)}(\gamma)\Big) = \text{Ker }\Big(\text{adj }\ (zI-E)_{z=\gamma_1}\Big) \oplus \text{Ker }\Big(\text{adj }\ (zI-E)_{z=\gamma_2}\Big) \oplus \dots \oplus \text{Ker }\Big(\text{adj }\ (zI-E)_{z=\gamma_r}\Big).$$

A representation of Ker (adj $(zI-E)_{z=\sigma}$) is now given. We therefore consider equation (2.23). Observe that $\widetilde{e}(z)^\top J = u^*(z)^\top S(e)$. The vector $\widetilde{e}(z)$ consists of Hörner polynomials associated with the companion matrix $E$ and $S(e)$ is the symmetrizer associated with the coefficients of the characteristic polynomial of the companion matrix $E$. An equivalent representation to (2.23) is then

$$\text{adj}(zI-E) = u(z)u^*(z)^\top S(e) - \pi(z)\sum z^j S^{j+1}.$$

Let $y \in$ Ker (adj $(zI-E)_{z=\sigma}$) and let $x = S(e)y$, then we have $y = S^{-1}(e)x$ and $x$ is a column in subspace (3.9). This will be illustrated in Example 3.6.

It can be seen that dim Ker $\Big(\mathcal{M}_r^{(1)}(\gamma)\Big) = r\ (r+v-1)$. When an interconnection takes place, or when $l+1=0$, we have dim Ker $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau)) = (r+v)\ (r+v-1)$ and dim Im $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau)) = (r+v)$.

**Example 3.6.** Consider the case of 4 distinct eigenvalues $\alpha, \beta, \gamma,$ and $\tau$. The $4\times 4$ adj $(zI-E)$ matrix, where $E$ is the companion matrix introduced in equation (2.17), is

$$\text{adj }\ (zI-E) = \begin{pmatrix} e_3 + e_2\ z + e_1\ z^2 + z^3 & e_2 + e_1\ z + z^2 & e_1 +\ z & 1 \\ -\ e_4 & e_2\ z + e_1\ z^2 + z^3 & e_1\ z + z^2 & z \\ -\ e_4\ z & -\ e_4 - e_3\ z & e_1\ z^2 + z^3 & z^2 \\ -\ e_4\ z^2 & -\ e_4\ z - e_3\ z^2 & -e_4 - e_3\ z - e_2\ z^2 & z^3 \end{pmatrix}.$$

The entries of the companion matrix $E$, when expressed in terms of the eigenvalues, are after identification with the corresponding coefficients of the characteristic equation

$e_1 = -(\alpha+\beta+\gamma+\tau),\ e_2 = \gamma\tau+\beta\tau+\alpha\tau+\beta\gamma+\alpha\gamma+\alpha\beta,\ e_3 = -(\beta\gamma\tau+\alpha\gamma\tau+\alpha\beta\tau+\alpha\beta\gamma)$ and $e_4 = \alpha\beta\gamma\tau$.

An explicit expression for adj $(zI - E)$ is then for $z{=}\alpha$

$$\text{adj } (\alpha I - E) = \begin{pmatrix} -\beta\gamma\tau & \beta\gamma+\beta\tau+\gamma\tau & -(\beta+\gamma+\tau) & 1 \\ -\alpha\beta\gamma\tau & \alpha\beta\gamma+\alpha\beta\tau+\alpha\gamma\tau & -(\alpha\beta+\alpha\gamma+\alpha\tau) & \alpha \\ -\alpha^2\beta\gamma\tau & \alpha^2\beta\gamma+\alpha^2\beta\tau+\alpha^2\gamma\tau & -(\alpha^2\beta+\alpha^2\gamma+\alpha^2\tau) & \alpha^2 \\ -\alpha^3\beta\gamma\tau & \alpha^3\beta\gamma+\alpha^3\beta\tau+\alpha^3\gamma\tau & -(\alpha^3\beta+\alpha^3\gamma+\alpha^3\tau) & \alpha^3 \end{pmatrix}.$$

The matrix $\mathcal{B}$ with columns in the subspace Ker $\{u(z)u^*(z)^\top\}_{z=\alpha}$, is according to (3.9)

$$\mathcal{B} = \begin{pmatrix} -\frac{1}{\alpha^3} & -\frac{1}{\alpha^2} & -\frac{1}{\alpha} \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The symmetrizer is given by

$$S(e) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -(\alpha+\beta+\gamma+\tau) & 1 & 0 & 0 \\ \gamma\tau+\beta\tau+\alpha\tau+\beta\gamma+\alpha\gamma+\alpha\beta & -(\alpha+\beta+\gamma+\tau) & 1 & 0 \\ -(\beta\gamma\tau+\alpha\gamma\tau+\alpha\beta\tau+\alpha\beta\gamma) & \gamma\tau+\beta\tau+\alpha\tau+\beta\gamma+\alpha\gamma+\alpha\beta & -(\alpha+\beta+\gamma+\tau) & 1 \end{pmatrix}.$$

The subspace Ker(adj $(zI - E)_{z=\alpha}$) is spanned by columns of the matrix $S^{-1}(e)\mathcal{B}$. For example, $y \in$ Ker (adj $(zI - E)_{z=\alpha}$) can have the following form when $\alpha \neq 0$

$$\begin{pmatrix} -\frac{1}{\alpha^3} \\ -\frac{\alpha+\beta+\gamma+\tau}{\alpha^3} \\ -\frac{\alpha^2+\beta^2+\gamma^2+\tau^2+\gamma\tau+\beta(\gamma+\tau)+\alpha(\beta+\gamma+\tau)}{\alpha^3} \\ -\frac{\beta^3+\gamma^3+\tau^3+\gamma^2\tau+\gamma\tau^2+\beta^2(\gamma+\tau)+\alpha^2(\beta+\gamma+\tau)+\beta(\gamma^2+\tau^2+\gamma\tau)+\alpha(\beta^2+\gamma^2+\tau^2+\gamma\tau+\beta(\gamma+\tau))}{\alpha^3} \end{pmatrix}.$$

## 4. Example

In this section an interconnection between $G_{bb}(\theta)$ and a corresponding solution to Stein's equation is illustrated for $p = q = 3$, $r = 2$ and $v = 1$. The same parametrization as in Examples 2.1 and 2.9 is used.

Note that in Example 3.2, a matrix $\mathcal{Q}$, which is in the kernel of $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$, is set forth so that a general solution to the linear system of equations (2.25) can be deduced. The particular solution $(\varphi \otimes I_q)$ is just one of the many solutions of the appropriate linear system of equations.

However, for establishing an interconnection between the Fisher information matrix and a solution to Stein's equation, the particular solution $(\varphi \otimes I_q)$, common to both linear systems (2.25) and (2.26) is considered. Consequently, the choice of the matrix, denoted by $\mathcal{A}$, contained in the subspace Ker $(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))$ and associated with the particular solution $(\varphi \otimes I_q)$ is then evaluated accordingly, to obtain

$$\mathcal{A} = (\varphi \otimes I_q) - (\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))^+\, G_{bb}(\theta).$$

For that purpose $(\varphi \otimes I_3)$ is first considered and $\varphi$ is constructed according to a variant of (2.14) given in (2.25) and (2.26). We obtain

$$(\varphi \otimes I_3) =$$

$$\begin{pmatrix} \frac{1}{(1-\gamma^2)^2(\gamma-\tau)(1-\gamma\tau)} & 0 & 0 \\ 0 & \frac{1}{(1-\gamma^2)^2(\gamma-\tau)(1-\gamma\tau)} & 0 \\ 0 & 0 & \frac{1}{(1-\gamma^2)^2(\gamma-\tau)(1-\gamma\tau)} \\ \frac{1+4\gamma^3\tau+\tau^2-3\gamma^2(1+\tau^2)}{(-1+\gamma^2)^3(-\gamma+\tau)^2(-1+\gamma\tau)^2} & 0 & 0 \\ 0 & \frac{1+4\gamma^3\tau+\tau^2-3\gamma^2(1+\tau^2)}{(-1+\gamma^2)^3(-\gamma+\tau)^2(-1+\gamma\tau)^2} & 0 \\ 0 & 0 & \frac{1+4\gamma^3\tau+\tau^2-3\gamma^2(1+\tau^2)}{(-1+\gamma^2)^3(-\gamma+\tau)^2(-1+\gamma\tau)^2} \\ \frac{1}{(-\gamma+\tau)^2(1-\tau^2)(1-\gamma\tau)^2} & 0 & 0 \\ 0 & \frac{1}{(-\gamma+\tau)^2(1-\tau^2)(1-\gamma\tau)^2} & 0 \\ 0 & 0 & \frac{1}{(-\gamma+\tau)^2(1-\tau^2)(1-\gamma\tau)^2} \end{pmatrix}$$

and

$$(\mathcal{U}_r(\gamma)\,\mathcal{U}_v(\tau))^+\,G_{bb}(\theta) =$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\frac{\gamma^2}{(-1+\gamma^2)^2(\gamma-\tau)(-1+\gamma\tau)} & -\frac{\gamma}{(-1+\gamma^2)^2(\gamma-\tau)(-1+\gamma\tau)} & -\frac{1}{(-1+\gamma^2)^2(\gamma-\tau)(-1+\gamma\tau)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\gamma(2\tau+2\gamma^4\tau-\gamma(1+\tau^2)-\gamma^3(1+\tau^2))}{(-1+\gamma^2)^3(\gamma-\tau)^2(-1+\gamma\tau)^2} & \frac{\tau+3\gamma^4\tau-2\gamma^3(1+\tau^2)}{(-1+\gamma^2)^3(\gamma-\tau)^2(-1+\gamma\tau)^2} & \frac{1+4\gamma^3\tau+\tau^2-3\gamma^2(1+\tau^2)}{(-1+\gamma^2)^3(\gamma-\tau)^2(-1+\gamma\tau)^2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\frac{\tau^2}{(\gamma-\tau)^2(-1+\gamma\tau)^2(-1+\tau^2)} & -\frac{\tau}{(\gamma-\tau)^2(-1+\gamma\tau)^2(-1+\tau^2)} & -\frac{1}{(\gamma-\tau)^2(-1+\gamma\tau)^2(-1+\tau^2)} \end{pmatrix},$$

with $(\mathcal{U}_r(\gamma)\,\mathcal{U}_v(\tau))^-_R$ given in Example 2.9 as an appropriate choice for $(\mathcal{U}_r(\gamma)\,\mathcal{U}_v(\tau))^+$ and $G_{bb}(\theta)$ is computed in Example 2.1.

Computation of matrix $\mathcal{A}$ shows that

$$\mathcal{A} = \frac{1}{(-1+\gamma^2)^3(\gamma-\tau)^2(-1+\gamma\tau)^2(-1+\tau^2)} \times$$

$$\begin{pmatrix} -(-1+\gamma^2)(\gamma-\tau)(-1+\gamma\tau)(-1+\tau^2) & 0 & 0 \\ 0 & -(-1+\gamma^2)(\gamma-\tau)(-1+\gamma\tau)(-1+\tau^2) & 0 \\ \gamma^2(-1+\gamma^2)(\gamma-\tau)(-1+\gamma\tau)(-1+\tau^2) & \gamma(-1+\gamma^2)(\gamma-\tau)(-1+\gamma\tau)(-1+\tau^2) & 0 \\ (-1+\tau^2)(1+4\gamma^3\tau+\tau^2-3\gamma^2(1+\tau^2)) & 0 & 0 \\ 0 & (-1+\tau^2)(1+4\gamma^3\tau+\tau^2-3\gamma^2(1+\tau^2)) & 0 \\ \gamma(-1+\tau^2)(-2\tau-2\gamma^4\tau+\gamma(1+\tau^2)+\gamma^3(1+\tau^2)) & -(-1+\tau^2)(\tau+3\gamma^4\tau-2\gamma^3(1+\tau^2)) & 0 \\ -(-1+\gamma^2)^3 & 0 & 0 \\ 0 & -(-1+\gamma^2)^3 & 0 \\ \tau^2(-1+\gamma^2)^3 & \tau(-1+\gamma^2)^3 & 0 \end{pmatrix},$$

it can be verified that the property $\mathcal{A} \in \mathrm{Ker}\,(\mathcal{U}_r(\gamma)\,\mathcal{U}_v(\tau))$ holds.

The factorization

$$(\mathcal{M}_2(\gamma)\ \mathcal{M}_1(\tau)) = \mathcal{M}^{(1)}_{2,\ 1}(\gamma,\tau)\ \mathcal{M}^{\Gamma}_{2,\ 1}\ \mathcal{M}^{(2)}_{2,\ 1}(\gamma,\tau)$$

is used for $r = 2$ and $v = 1$.

The $12 \times 12$ matrix $\mathcal{M}^{\Gamma}_{2,\ 1}$ has the form $\mathcal{M}^{\Gamma}_{2,\ 1} = \mathrm{diag}\ \{\ \Gamma, \Gamma, \Gamma, \Gamma\}$.

The block $\mathcal{M}^{(1)}_{2,\ 1}(\gamma,\tau)$ is given by the $3\times12$ matrix $\mathcal{M}^{(1)}_{2,\ 1}(\gamma,\tau) = \left(\mathcal{M}^{(1)}_1(\gamma)\ \ \mathcal{M}^{(1)}_0(\tau)\right)$, with $\mathcal{M}^{(1)}_1(\gamma) = \left(\mathcal{M}^{(1)\ (1)}_1(z)\ \ \mathcal{M}^{(0)\ (1)}_1(z)\right)_{z=\gamma}$. The blocks constituting $\mathcal{M}^{(1)}_1(\gamma)$ are $\mathcal{M}^{(1)\ (1)}_1(\gamma) = \left(\frac{\partial}{\partial z}\mathrm{adj}\ (zI-E)\ \mathrm{adj}\ (zI-E)\right)_{z=\gamma}$, $\mathcal{M}^{(0)\ (1)}_1(\gamma) = (\mathrm{adj}\ (zI-E))_{z=\gamma}$ respectively and $\mathcal{M}\ ^{(1)}_0(\tau) = (\mathrm{adj}\ (zI-E))_{z=\tau}$.

The desired adjoint matrices are

$$\mathrm{adj}\ (zI-E) = \begin{pmatrix} z^2+e_1\ z+e_2 & z+e_1 & 1 \\ -e_3 & z^2+e_1\ z & z \\ -e_3\ z & -e_2\ z-e_3 & z^2 \end{pmatrix},$$

$$\mathrm{adj}\ (I-zE) = \begin{pmatrix} 1+e_1\ z+e_2\ z^2 & z+e_1\ z^2 & z^2 \\ -e_3\ z^2 & 1+e_1\ z & z \\ -e_3\ z & -e_2\ z-e_3\ z^2 & 1 \end{pmatrix}.$$

The entries $e_1$, $e_2$ and $e_3$ of the companion matrix $E$ will be subsequently expressed in terms of the roots-eigenvalues $\gamma$ and $\tau$. This results in the following representation of the appropriate matrices

$$\mathcal{M}^{(1)}_1(\gamma) = \begin{pmatrix} -\tau & 1 & 0 & \tau\gamma & -\gamma-\tau & 1 & \tau\gamma & -\gamma-\tau & 1 \\ 0 & -\tau & 1 & \tau\gamma^2 & -\gamma^2-\tau\gamma & \gamma & \tau\gamma^2 & -\gamma^2-\tau\gamma & \gamma \\ \tau\gamma^2 & -\gamma^2-2\tau\gamma & 2\gamma & \tau\gamma^3 & -\gamma^3-\tau\gamma^2 & \gamma^2 & \tau\gamma^3 & -\gamma^3-\tau\gamma^2 & \gamma^2 \end{pmatrix}$$

and

$$\mathcal{M}^{(1)}_0(\tau) = \begin{pmatrix} \gamma^2 & -2\gamma & 1 \\ \tau\gamma^2 & -2\tau\gamma & \tau \\ \tau^2\gamma^2 & -2\tau^2\gamma & \tau^2 \end{pmatrix}.$$

Whereas the $12\times9$ matrix $\mathcal{M}^{(2)}_{2,\ 1}(\gamma,\tau)$ has the representation, $\mathcal{M}^{(2)}_{2,\ 1}(\gamma,\tau) =$diag $\left\{\mathcal{M}^{(2)}_1(\gamma)\ \mathcal{M}^{(2)}_0(\tau)\right\}$, with

$$\mathcal{M}^{(2)}_1(\gamma) = \begin{pmatrix} \mathcal{M}^{(1)\ (2)}(z) & 0 \\ 0 & \mathcal{M}^{(0)\ (2)}(z) \end{pmatrix}_{z=\gamma},\ \mathcal{M}^{(0)\ (2)}(\tau) = \mathcal{M}\ ^{(2)}_0(\tau) = \left(\ \mathrm{adj}\ \ (I-zE)^\top\right)_{z=\tau},$$

$$\mathcal{M}^{(1)\ (2)}(\gamma) = \begin{pmatrix} \mathrm{adj}\ \ (I-zE)^\top \\ \frac{\partial}{\partial z}\mathrm{adj}\ \ (I-zE)^\top \end{pmatrix}_{z=\gamma} \quad \text{and} \quad \mathcal{M}^{(0)\ (2)}(\gamma) = \left(\ \mathrm{adj}\ \ (I-zE)^\top\right)_{z=\gamma}.$$

An explicit form is

$$\mathcal{M}^{(1)\ (2)}(\gamma) = \begin{pmatrix} 1-2\gamma^2+\gamma^4-\gamma\tau+2\gamma^3\tau & \gamma^4\tau & \gamma^3\tau \\ \gamma-2\gamma^3-\gamma^2\tau & 1-2\gamma^2-\gamma\tau & -\gamma^3-2\gamma^2\tau+\gamma^4\tau \\ \gamma^2 & \gamma & 1 \\ -2\gamma+2\gamma^3-\tau+4\gamma^2\tau & 2\gamma^3\tau & \gamma^2\tau \\ 1-4\gamma^2-2\gamma\tau & -2\gamma-\tau & -\gamma^2-2\gamma\tau+2\gamma^3\tau \\ 2\gamma & 1 & 0 \end{pmatrix}$$

and

$$\mathcal{M}^{(0)\ (2)}(\gamma) = \begin{pmatrix} 1-2\gamma^2+\gamma^4-\gamma\tau+2\gamma^3\tau & \gamma^4\tau & \gamma^3\tau \\ \gamma-2\gamma^3-\gamma^2\tau & 1-2\gamma^2-\gamma\tau & -\gamma^3-2\gamma^2\tau+\gamma^4\tau \\ \gamma^2 & \gamma & 1 \end{pmatrix},$$

$$\mathcal{M}_0^{(2)}(\tau) = \begin{pmatrix} 1-2\gamma\tau-\tau^2+\gamma^2\tau^2+2\gamma\tau^3 & \gamma^2\tau^3 & \gamma^2\tau^2 \\ \tau-2\gamma\tau^2-\tau^3 & 1-2\gamma\tau-\tau^2 & -\gamma^2\tau-2\gamma\tau^2+\gamma^2\tau^3 \\ \tau^2 & \tau & 1 \end{pmatrix}.$$

For this example we choose $\Gamma = I_3$ or the identity matrix.
The matrices $(\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau))$,$(\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))^+ = (\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))_R^-$ , $G_{bb}(\theta)$ and $\mathcal{A}$ are now inserted in the equation

$$S_{bb} = (\mathcal{M}_r(\gamma)\ \mathcal{M}_v(\tau)) \Big\{ (\mathcal{U}_r(\gamma)\ \mathcal{U}_v(\tau))^+ G_{bb}(\theta) + \mathcal{A} \Big\}.$$

A solution to the Stein equation when expressed in terms of the Fisher information matrix is derived, to obtain

$$S_{bb} = \frac{1}{(\gamma^2-1)^3(\gamma-\tau)^2(\tau^2-1)(-1+\gamma\tau)^2} \begin{pmatrix} S_{bb}^{11} & S_{bb}^{12} & S_{bb}^{13} \\ S_{bb}^{21} & S_{bb}^{22} & S_{bb}^{23} \\ S_{bb}^{31} & S_{bb}^{32} & S_{bb}^{33} \end{pmatrix},$$

where

$S_{bb}^{11} = -(\gamma-\tau)^2(-3+5\gamma^2-7\gamma^4+\gamma^6+2\gamma\tau-14\gamma^3\tau+10\gamma^5\tau-2\gamma^7\tau+2\tau^2-11\gamma^2\tau^2+9\gamma^4\tau^2-\gamma^6\tau^2+\gamma^8\tau^2-4\gamma\tau^3+8\gamma^3\tau^3-4\gamma^5\tau^3+4\gamma^7\tau^3+2\gamma^2\tau^4+4\gamma^6\tau^4)$,

$S_{bb}^{12} = -(\gamma-\tau)^2(-2\gamma-2\gamma^5-\tau-7\gamma^4\tau+4\gamma^6\tau-8\gamma^3\tau^2+8\gamma^5\tau^2-4\gamma^2\tau^3+7\gamma^4\tau^3+\gamma^8\tau^3+2\gamma^3\tau^4+2\gamma^7\tau^4)$,

$S_{bb}^{13} = -(-2\gamma^4-2\gamma^6+4\gamma^3\tau-2\gamma^5\tau+2\gamma^9\tau-3\gamma^2\tau^2+6\gamma^4\tau^2+6\gamma^6\tau^2-2\gamma^8\tau^2+\gamma^{10}\tau^2+2\gamma\tau^3-4\gamma^3\tau^3-4\gamma^7\tau^3-2\gamma^9\tau^3-\tau^4+2\gamma^2\tau^4-8\gamma^6\tau^4+3\gamma^8\tau^4-2\gamma\tau^5+2\gamma^5\tau^5+4\gamma^7\tau^5+4\gamma^4\tau^6-4\gamma^6\tau^6)$,

$S_{bb}^{21} = -(\gamma-\tau)^2(-2\gamma-2\gamma^5-\tau-7\gamma^4\tau+4\gamma^6\tau-8\gamma^3\tau^2+8\gamma^5\tau^2-4\gamma^2\tau^3+7\gamma^4\tau^3+\gamma^8\tau^3+2\gamma^3\tau^4+2\gamma^7\tau^4)$,

$S_{bb}^{22} = -(\gamma-\tau)^2(-2+2\gamma^2-4\gamma^4-8\gamma^3\tau+4\gamma^5\tau+\tau^2-7\gamma^2\tau^2+3\gamma^4\tau^2+3\gamma^6\tau^2-2\gamma\tau^3+2\gamma^3\tau^3+2\gamma^5\tau^3+2\gamma^7\tau^3+\gamma^2\tau^4+\gamma^4\tau^4+\gamma^6\tau^4+\gamma^8\tau^4)$,

$S_{bb}^{23} = -(-2\gamma^3-2\gamma^7+3\gamma^2\tau-3\gamma^6\tau+4\gamma^8\tau+2\gamma^5\tau^2+6\gamma^7\tau^2-\tau^3+5\gamma^4\tau^3-5\gamma^6\tau^3-8\gamma^8\tau^3+\gamma^{10}\tau^3-4\gamma^5\tau^4-4\gamma^2\tau^5+3\gamma^4\tau^5+5\gamma^8\tau^5+2\gamma^3\tau^6+2\gamma^5\tau^6-4\gamma^7\tau^6)$,

$S_{bb}^{31} = -(\gamma-\tau)^2(-2\gamma^2-2\gamma^4-6\gamma^3\tau+2\gamma^7\tau-\tau^2-2\gamma^2\tau^2+2\gamma^6\tau^2+\gamma^8\tau^2-2\gamma\tau^3+6\gamma^5\tau^3+2\gamma^4\tau^4+2\gamma^6\tau^4)$,

$S_{bb}^{32} = S_{bb}^{12}$,

$S_{bb}^{33} = -(-\gamma^2-\gamma^4-\gamma^6-\gamma^8+2\gamma\tau+2\gamma^9\tau-\tau^2+3\gamma^2\tau^2-2\gamma^6\tau^2+9\gamma^8\tau^2-\gamma^{10}\tau^2-2\gamma\tau^3+4\gamma^3\tau^3-4\gamma^7\tau^3-6\gamma^9\tau^3-3\gamma^2\tau^4+5\gamma^4\tau^4-5\gamma^6\tau^4-3\gamma^8\tau^4+2\gamma^{10}\tau^4-4\gamma^3\tau^5+4\gamma^7\tau^5+4\gamma^9\tau^5+4\gamma^4\tau^6-4\gamma^8\tau^6)$.

It can be verified that when $\Gamma = w_3 w_3^\top$, where $w_3$ is the last standard basis vector in $\mathbb{R}^3$, the solution to Stein's equation $S_{bb}$ indeed coincides with the Fisher information matrix $G_{bb}(\theta)$.

## 5. The global Fisher information matrix

In this section the entire Fisher information matrix, not decomposed, is considered. In Klein and Spreij [7], the representation of the global Fisher information matrix is given by

$$G(\theta) = \begin{pmatrix} -S_p(b) \\ S_q(a) \\ 0 \end{pmatrix} Q(\theta) \begin{pmatrix} -S_p(b) \\ S_q(a) \\ 0 \end{pmatrix}^\top + \begin{pmatrix} -S_p(c) \\ 0 \\ S_r(a) \end{pmatrix} P(\theta) \begin{pmatrix} -S_p(c) \\ 0 \\ S_r(a) \end{pmatrix}^\top, \tag{5.1}$$

where

$$Q(\theta) = \frac{1}{2\pi i} \oint_{|z|=1} \frac{R_x(z) u_{p+q}(z) u_{p+q}^\top(z^{-1})}{a(z)a(z^{-1})c(z)c(z^{-1})} \frac{dz}{z},$$

$$P(\theta) = \frac{1}{2\pi i} \oint\limits_{|z|=1} \frac{u_{p+r}(z)u_{p+r}^{\top}(z^{-1})}{a(z)a(z^{-1})c(z)c(z^{-1})} \frac{dz}{z}$$

and $S_p(b)$ and $S_q(a)$ are blocks of the Sylvester resultant matrices $S(-b, a)$

$$S(-b, a) = \begin{pmatrix} -S_p(b) \\ S_q(a) \end{pmatrix} \text{ and } S(-c, a) = \begin{pmatrix} -S_p(c) \\ S_r(a) \end{pmatrix}.$$

Where $S_p(b)$ is formed by the top $p$ rows of $S(-b, a)$ and similarly for the remaining blocks. The Sylvester resultant $S(c, -a)$ is the $(p+r) \times (p+r)$ matrix defined as

$$S(a, c) = \begin{pmatrix} 1 & a_1 & a_2 & \cdots & a_p & \cdots & 0 \\ & \ddots & \ddots & \ddots & & \ddots & \\ 0 & & 1 & a_1 & a_2 & \cdots & a_p \\ 1 & c_1 & c_2 & \cdots & c_r & \cdots & 0 \\ & \ddots & \ddots & \ddots & & \ddots & \\ 0 & & 1 & c_1 & c_2 & \cdots & c_r \end{pmatrix}.$$

We shall show that both terms of (5.1) can be expressed by solutions of corresponding Stein equations. In [8] it is shown that the matrix $P(\theta)$ fulfills a Stein equation, it is given by

$$P(\theta) - E_{ac} P(\theta) E_{ac}^{\top} = w_{p+r} w_{p+r}^{\top}, \tag{5.2}$$

where $w_{p+r}$ is the last standard basis vector in $\mathbb{R}^{p+r}$. The entries of the companion matrix $E_{ac}$ are associated with the coefficients of the polynomial $a(z)c(z)$.We consider the case of an interconnection, this implies $q = r + v$. We therefore rewrite $Q(\theta)$ as

$$Q(\theta) = \frac{1}{2\pi i} \oint\limits_{|z|=1} \frac{u_{p+q}(z)u_{p+q}^{\top}(z^{-1})}{h(z)h(z^{-1})a(z)a(z^{-1})c(z)c(z^{-1})} \frac{dz}{z}.$$

We now construct a companion matrix of degree $p + r + v = p + q$, denoted by $E_{ach}$, with entries that are associated with the coefficients of the polynomial $a(z)c(z)h(z)$. Consequently, the matrix $Q(\theta)$ verifies the following Stein equation

$$Q(\theta) - E_{ach} Q(\theta) E_{ach}^{\top} = w_{p+q} w_{p+q}^{\top}, \tag{5.3}$$

where $w_{p+q}$ is the last standard basis vector in $\mathbb{R}^{p+q}$. When appropriate choices for $\Gamma$ in (2.16) are considered, we can express the Fisher information matrix $G(\theta)$ in terms of solutions to the Stein equations (5.2) and (5.3). Observe that this result can also be used to express each block of $G(\theta)$ in terms of solutions to the Stein equations (5.2) and (5.3). For example, the Fisher information matrix $G_{aa}(\theta)$ is then explained by the $(p \times p)$ matrices in the right-hand side of (5.1), to obtain

$$G_{aa}(\theta) = S_p(b) Q(\theta) S_p^{\top}(b) + S_p(c) P(\theta) S_p^{\top}(c).$$

That can be generalized for different values of $\Gamma$, with the condition formulated in Proposition 2.5. In this case $P(\theta)$ and $Q(\theta)$ in (5.1) can be replaced by elements that are expressed by corresponding solutions to Stein equations. These solutions are obtained by solving systems of linear equations. This

is done in a similar manner as in the block case when appropriate companion matrices are inserted in the Stein equations. Consequently, the Fisher information matrix $G(\theta)$ is then explained by these solutions as well as by Sylvester resultants (in [7] it is shown through equation (5.1) that the Fisher information matrix has the resultant property). Algorithms for the kernels of the appropriate coefficient matrices can be constructed according to the development described in Section 3.
An algorithm of the Fisher information matrix of an ARMAX process is developed in [10]. Consequently, when a solution to a Stein equation coincides with the Fisher information matrix (the condition is mentioned in this paper), the value of this Stein solution can then be straightforwardly computed by this algorithm. More generally, by using the algorithm developed in [10], combined with the results obtained in this paper, allows us to develop numerical computations of a solution to a Stein equation by means of Fisher's information matrix. This can be a subject for further study.

**Example 5.1.** We shall illustate some results outlined in Section 5. Consider the ARMAX process with $p = 1$, $r = 1$, $v = 1$ and $q = 2$. The following polynomials are given, $a(z) = z + a$, $c(z) = z + c$, $b(z) = z^2 + b_1 z + b_2$ and $h(z) = z + \tau$. The matrix $P(\theta)$ is then,

$$P(\theta) = \frac{1}{(1-a^2)(1-ac)(1-c^2)} \begin{pmatrix} 1+ac & -(a+c) \\ -(a+c) & 1+ac \end{pmatrix}.$$

Consider the companion matrix in (5.2),

$$E_{ac} = \begin{pmatrix} 0 & 1 \\ -ac & -(a+c) \end{pmatrix}.$$

It can be verified that the following Stein equation holds true

$$P(\theta) - E_{ac} P(\theta) E_{ac}^{\top} = w_2 w_2^{\top},$$

where $w_2$ is the last standard basis vector in $\mathbb{R}^2$. The Sylvester matrix is

$$\begin{pmatrix} -S_p(c) \\ 0 \\ S_r(a) \end{pmatrix} = \begin{pmatrix} -1 & -c \\ 0 & 0 \\ 0 & 0 \\ 1 & a \end{pmatrix}.$$

We have the symmetric and Toeplitz matrix

$$Q(\theta) = \frac{1}{(1-a^2)(1-ac)(1-c^2)(1-a\tau)(1-c\tau)(1-\tau^2)} \begin{pmatrix} Q_{11}(\theta) & Q_{12}(\theta) & Q_{13}(\theta) \\ Q_{21}(\theta) & Q_{22}(\theta) & Q_{23}(\theta) \\ Q_{31}(\theta) & Q_{32}(\theta) & Q_{33}(\theta) \end{pmatrix},$$

with

$$\begin{aligned} Q_{11}(\theta) &= Q_{22}(\theta) = Q_{33}(\theta) = 1 + c\tau - a^2 c\tau(1 + c\tau) + a(c + \tau - c^2\tau - c\tau^2), \\ Q_{12}(\theta) &= Q_{23}(\theta) = Q_{21}(\theta) = Q_{32}(\theta) = -c - \tau + a^2 c\tau(c + \tau) + a(-1 + c^2\tau^2), \\ Q_{13}(\theta) &= Q_{31}(\theta) = c\tau + \tau^2 - c^2(-1 + \tau^2) - a^2(-1 + c^2 + c\tau + \tau^2) + a(c + \tau - c^2\tau - c\tau^2). \end{aligned}$$

The companion matrix associated with (5.3) is

$$E_{ach} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -ac\tau & -ac - a\tau - c\tau & -a - c - \tau \end{pmatrix}.$$

It can be verified that the following Stein equation holds true

$$Q(\theta) - E_{ach}Q(\theta)E_{ach}^{\top} = w_3 w_3^{\top},$$

where $w_3$ is the last standard basis vector in $\mathbb{R}^3$.The matrix containing Sylvester matrices is given by,

$$\begin{pmatrix} -S_p(b) \\ S_q(a) \\ 0 \end{pmatrix} = \begin{pmatrix} -1 & -b_1 & -b_2 \\ 1 & a & 0 \\ 0 & 1 & a \\ 0 & 0 & 0 \end{pmatrix}.$$

We can now compute the Fisher information matrix $G(\theta)$ by means of solutions of Stein equations.

**Acknowledgements**

*The authors wish to thank the associate editor Leiba Rodman, Andre Ran and an anonymous referee for valuable comments and suggestions.*

## References

[1] P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*, Springer Verlag, New York, 1991.

[2] E.J. Hannan and M. Deistler, *The Statistical Theory of Linear Systems*, John Wiley & Sons, Inc.New York, 1988.

[3] B.R. Frieden, *Physics from Fisher Information: A Unification,* Cambridge University Press, New York, 1998.

[4] B.R. Frieden, *Science from Fisher Information: A Unification,* Cambridge University Press, New York, 2004.

[5] B. Friedlander, On the computation of the Cramér-Rao bound for ARMA parameter estimation, *IEEE Trans. Acoust., Speech, Signal Processing* **32** (1984) no. 4, 721-727.

[6] A. Klein and G. Mélard, Computation of the Fisher information matrix for SISO models, *IEEE Transactions on Signal Processing* **42** (1994), 684-688.

[7] A. Klein and P. Spreij, On Fisher's Information Matrix of an ARMAX Process and Sylvester's Resultant Matrices, *Linear Algebra and its Applications* **237/238** (1996), 579-590.

[8] A. Klein and P. Spreij, On Fisher's information matrix of an ARMA process. *Stochastic Differential and Difference Equations. Progress in Systems and Control Theory.* I. Csiszar and Gy. Michaletzky, Editors. Birkhauser Boston **23** (1997), 273-284.

[9] A. Klein and P. Spreij, On Stein's equation, Vandermonde matrices and Fisher's information matrix of time series processes. Part I: The autoregressive moving average process, *Linear Algebra and its Applications* **329** (2001), 9-47.

[10] A.Klein and G. Mélard, An algorithm for computing the asymptotic Fisher information matrix for seasonal SISO models, *Journal of Time Series Analysis*, Vol.**25**, No.5 (2004), 627-648.

[11] A. Klein and P. Spreij, Some results on Vandermonde matrices with an application to time series analysis, *SIAM Journal on Matrix Analysis and Applications*, Vol.**25**, No.1 (2003), 213-223.

[12] P. Lancaster, L. Lerer and M.Tismenetsky, Factored Forms for Solutions of AX−XB=C and X−AXB=C in Companion Matrices, *Linear Algebra and its Applications* **62** (1984), 19-49.

[13] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.

[14] P. Lancaster and M. Tismenetsky, *The Theory of Matrices with Applications* (second edition), Academic Press, Orlando, 1985.

[15] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, M.I.T. Press, Cambridge, Mass., 1983.