

A SCALARIZATION TECHNIQUE FOR COMPUTING THE POWER AND EXPONENTIAL MOMENTS OF GAUSSIAN RANDOM MATRICES

IGOR VLADIMIROV AND BEVAN THOMPSON

Received 28 December 2004; Revised 13 May 2005; Accepted 13 May 2005

We consider the problems of computing the power and exponential moments EX^s and Ee^{tX} of square Gaussian random matrices $X = A + BWC$ for positive integer s and real t , where W is a standard normal random vector and A, B, C are appropriately dimensioned constant matrices. We solve the problems by a matrix product scalarization technique and interpret the solutions in system-theoretic terms. The results of the paper are applicable to Bayesian prediction in multivariate autoregressive time series and mean-reverting diffusion processes.

Copyright © 2006 I. Vladimirov and B. Thompson. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

We consider the problems of computing the power moments EX^s and the exponential moments Ee^{tX} of square Gaussian random matrices

$$X = A + BWC \tag{1.1}$$

for positive integer s and real t . Here W is a standard normal random column vector, further referred to as the *source*, and A, B, C are appropriately dimensioned constant matrices.

The power moments problem is relevant to Bayesian prediction of autoregressive time series wherein the normal-gamma distribution [1, page 140] is used as a conjugate prior for the unknown parameters; see, also [5, 6, 12].

Being a continuous-time counterpart of the above, the exponential moments arise from Bayesian forecasting in stochastic diffusion models with affine drift terms whose unknown parameters are ascribed Gaussian prior distributions; see, for example, [7, pages 279–280] and [9]. In these applications, s and t are interpreted as the time ahead for which the prediction is to be made.

2 Moments of Gaussian random matrices

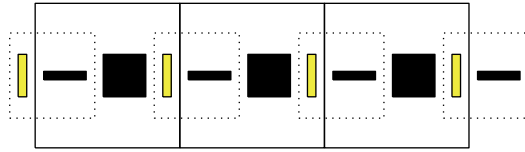


Figure 1.1. A product of constant matrices (black squares) alternating with rank-one random matrices (dotted-frame boxes). These last are obtained by multiplying jointly Gaussian random column vectors (vertical shaded bars) by constant row vectors (horizontal black bars). Regrouping the terms reduces the matrix product to a Gaussian random column vector (the leftmost bar) multiplied by *scalar* Gaussian random variables (solid-frame boxes) and by a constant row vector (the rightmost bar).

If X is *scalar*, the solutions to the power and exponential moments problems are provided by the well-known moment-generating function of a univariate Gaussian distribution. However as soon as the order of X is greater than one, the problems become complicated by the noncommutativity of the matrix algebra. Nevertheless, if X is generated by a *vector* source W as assumed above, the noncommutativity is essentially surmountable with the aid of a *scalarization* technique introduced below which appears to be as efficacious as it is simple.

The technique applies to products of constant matrices alternating with jointly Gaussian random matrices of rank one as elucidated informally by Figure 1.1. Regrouping the terms reduces such a product to a rank-one random matrix multiplied by *scalar* random variables, which together have a Gaussian joint distribution. The essentially “scalarized” representation makes it easy to relate the expectation of the matrix product with the covariance structure of the multipliers. This concluding ingredient of the technique is accomplished by an extension of the well-known formula for the expectation of the product of *four* jointly Gaussian zero-mean random variables [3, Example 4.7.1, page 84] to an arbitrary even number of such variables.

The matrix product scalarization technique makes it possible to express the power moments in terms of recursively computed state-space realization matrices of an iterated linear discrete-time invariant dynamical system. This system-theoretic interpretation enables us to give a tractable closed form solution of the problem and provides a machinery for computing this solution. Moreover, a similar, though different, representation via continuous-time systems turns out to be valid for the exponential moments. These are reduced to a *univariate* integral of another matrix exponential along a pencil of matrices that, in principle, admits numerical computation by the methods of [10].

The scalarization technique outlined above is inapplicable to the more general situation where the source W in (1.1) is a standard normal random *matrix*. In the matrix source case, however, the power moments problem is amenable to a more sophisticated graph-theoretic analysis [11].

The paper is organized as follows. Section 2 develops the matrix product scalarization technique and provides a solution to the power moments problem. Section 3 gives a system-theoretic interpretation to the solution as well as the necessary background

material on linear time-invariant systems. In Section 4, a similar system-theoretic formalism is developed for the exponential moments problem. Some proofs are relegated to appendices.

2. Matrix product scalarization

For given matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and a given row vector $C \in \mathbb{R}^{1 \times n}$, we consider a $\mathbb{R}^{n \times n}$ -valued Gaussian random matrix X of the form (1.1), where

$$W \triangleright \triangleleft \mathcal{N}_m. \quad (2.1)$$

Here \mathcal{N}_m denotes the standard normal distribution on \mathbb{R}^m , so that the source W is an m -dimensional Gaussian random vector with zero mean and identity covariance matrix. Without loss of generality, it may be assumed that $m \leq n$. Unless otherwise indicated, vectors are organized as columns.

A solution to the problem of computing the s th power moment $\mathbb{E}X^s$ is provided by the theorem below. Its proof develops the matrix product scalarization technique outlined in the introduction. To this end, we associate with the matrix A a map $\mathcal{A} : U \rightarrow \mathbb{R}^{n \times n}$ defined by

$$\mathcal{A}(u) = \sum_{k \in \mathbb{Z}_+} (uA)^k = (I_n - uA)^{-1} \quad (2.2)$$

on the set

$$U = \{u \in \mathbb{R} : |u| < 1/\mathbf{r}(A)\}. \quad (2.3)$$

Here \mathbb{Z}_+ is the set of nonnegative integers, I_n denotes the identity matrix of order n , and $\mathbf{r}(\cdot)$ is the spectral radius of a square matrix. Furthermore, we introduce a map $\mathcal{E} : U \rightarrow \mathbb{R}^{n \times n}$ by

$$\mathcal{E}(u) = \mathcal{A}(u)BB^T(\mathcal{A}(u))^T C^T C, \quad (2.4)$$

where $(\cdot)^T$ is the matrix transpose. One verifies that $(\mathcal{E}(u))^2 = \mathcal{E}(u) \text{Tr} \mathcal{E}(u)$, with $\text{Tr} \mathcal{E}(u) = |C\mathcal{A}(u)B|^2$, and hence the matrix $\mathcal{E}(u)/\text{Tr} \mathcal{E}(u)$ is idempotent. Finally, for every $r \in \mathbb{N}$, we define a map $\mathcal{H}_r : U \rightarrow \mathbb{R}^{n \times n}$ by

$$\mathcal{H}_r(u) = (\mathcal{E}(u))^r \mathcal{A}(u) = (\text{Tr} \mathcal{E}(u))^{r-1} \mathcal{E}(u) \mathcal{A}(u) = \sum_{k \in \mathbb{Z}_+} u^k H_{r,k}, \quad (2.5)$$

where the matrices $H_{r,k} \in \mathbb{R}^{n \times n}$ are given by

$$H_{r,k} = \frac{1}{k!} \left. \frac{d^k \mathcal{H}_r(u)}{du^k} \right|_{u=0}. \quad (2.6)$$

4 Moments of Gaussian random matrices

THEOREM 2.1. *For every $s \in \mathbb{N}$, the s th power moment of the Gaussian random matrix X defined by (1.1)–(2.1) can be expressed in terms of (2.6) as*

$$\mathbf{E}X^s = A^s + \sum_{r=1}^{\lfloor s/2 \rfloor} (2r-1)!! H_{r,s-2r}, \quad (2.7)$$

where $\lfloor \cdot \rfloor$ is the floor function which sends $v \in \mathbb{R}$ to the largest $z \in \mathbb{Z}$ not exceeding v .

Proof. Since the m -dimensional standard normal distribution \mathcal{N}_m is symmetric about the origin,

$$\mathbf{E}X^s = \sum_{r=0}^{\lfloor s/2 \rfloor} \sum_{(\mu_1, \dots, \mu_{2r+1}) \in \mathcal{M}_{s-2r}^{2r+1}} A^{\mu_1} \mathbf{E} \prod_{k=2}^{2r+1} BWCA^{\mu_k}. \quad (2.8)$$

Here, for any $d \in \mathbb{N}$ and $\sigma \in \mathbb{Z}_+$,

$$\mathcal{M}_\sigma^d = \{(\mu_1, \dots, \mu_d) \in \mathbb{Z}_+^d : \mu_1 + \dots + \mu_d = \sigma\} \quad (2.9)$$

is the set of those d -indices whose entries add up to σ . Since matrices do not commute, in general, we explicitly indicate the order in which they are multiplied as follows:

$$\prod_{k=a}^b M_k = M_a \times \dots \times M_b = \left(\prod_{k=a}^b (M_k^T) \right)^T. \quad (2.10)$$

For any $r \in \mathbb{N}$, the matrices which are multiplied on the right-hand side of (2.8) can be regrouped so as to partially “scalarize” the product,

$$\begin{aligned} \prod_{k=2}^{2r+1} BWCA^{\mu_k} &= \underbrace{BWCA^{\mu_2}} \times \dots \times \underbrace{BWCA^{\mu_{2r+1}}} \\ &= BW \underbrace{CA^{\mu_2} BW} \times \dots \times \underbrace{CA^{\mu_{2r}} BW} CA^{\mu_{2r+1}} \\ &= BW \left(\prod_{k=2}^{2r} (CA^{\mu_k} BW) \right) CA^{\mu_{2r+1}}. \end{aligned} \quad (2.11)$$

Thus, the original product of random matrices is reduced to the random vector W multiplied by the *scalar* random variables

$$\eta_k = CA^{\mu_k} BW, \quad k \in \{2, \dots, 2r\}, \quad (2.12)$$

where $\{a, \dots, b\} = [a, b] \cap \mathbb{Z}$ is the discrete interval with endpoints $a, b \in \mathbb{Z}$ satisfying $a \leq b$. Note that the joint distribution of W and η_2, \dots, η_{2r} is Gaussian with zero mean. To proceed, we need the following definition.

Definition 2.2. Let S be a linearly ordered set with $\#S = 2r$, where $\#$ is the counting measure and $r \in \mathbb{N}$. A permutation $(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r)$ of S is called *regular* if it satisfies $\alpha_k < \beta_k$ for every $k \in \{1, \dots, r\}$, and $\alpha_1 < \dots < \alpha_r$.

The class of regular permutations of S will be referred to as \mathcal{P}_S . Clearly $\alpha_1 = \min S$ and $\alpha_r < \max S$ for any $(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_S$. Note that \mathcal{P}_S can be identified with all possible partitions of S into pairs; hence, $\#\mathcal{P}_S = (2r - 1)!!$.

The class of all possible permutations of S will be referred to as Π_S . Slightly abusing notation, $\Pi_{\{1, \dots, r\}}$ will be abbreviated as Π_r , and we will write \mathcal{P}_r for the class of regular permutations associated with the set

$$\mathcal{L}_r = \{1, \dots, 2r\}. \quad (2.13)$$

For instance, \mathcal{P}_2 consists of three permutations $(1, 2, 3, 4)$, $(1, 3, 2, 4)$, and $(1, 4, 2, 3)$, with the number of all possible permutations of \mathcal{L}_2 being $\#\Pi_4 = 4! = 24$.

LEMMA 2.3. *Let $\xi = (\xi_k)_{k \in \mathcal{L}_r}$ be an \mathbb{R}^{2r} -valued Gaussian random vector with zero mean and covariance matrix $\Sigma = (\sigma_{jk})_{j, k \in \mathcal{L}_r}$. Then*

$$\mathbf{E} \prod_{k=1}^{2r} \xi_k = \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r} \prod_{k=1}^r \sigma_{\alpha_k \beta_k}. \quad (2.14)$$

The assertion of the lemma is well known in the case $r = 2$; see, for instance, [3, Example 4.7.1, page 84]. To make the exposition self-contained, Lemma 2.3 is proved for arbitrary r in Appendix A.

We will need the following extension of the lemma. Let ξ be an \mathbb{R}^m -valued random vector and let η_2, \dots, η_{2r} be scalar random variables, which together have a zero-mean Gaussian joint distribution. Then application of (2.14) yields

$$\begin{aligned} \mathbf{E} \left(\xi \prod_{k=2}^{2r} \eta_k \right) &= \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r} \mathbf{E}(\xi \eta_{\beta_1}) \prod_{k=2}^r \mathbf{E}(\eta_{\alpha_k} \eta_{\beta_k}) \\ &= \sum_{j=2}^{2r} \mathbf{E}(\xi \eta_j) \mathbf{E} \prod_{k \in \{2, \dots, 2r\} \setminus \{j\}} \eta_k. \end{aligned} \quad (2.15)$$

Here we made use of the properties that $\alpha_1 = 1$ and $(\alpha_2, \beta_2, \dots, \alpha_r, \beta_r) \in \mathcal{P}_{\{2, \dots, 2r\} \setminus \{1\}}$ for any $(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r$. More precisely, for every $j \in \{2, \dots, 2r\}$, the inclusions $(1, j, \alpha_2, \beta_2, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r$ and $(\alpha_2, \beta_2, \dots, \alpha_r, \beta_r) \in \mathcal{P}_{\{2, \dots, 2r\} \setminus \{j\}}$ are equivalent. The equivalence provides a basis for the recursive generation of the classes \mathcal{P}_r . Now applying (2.15) to the random vector $\xi = W \triangleright \triangleleft \mathcal{N}_m$ and to the random variables (2.12), one verifies that

$$\begin{aligned} &\mathbf{E} \left(W \prod_{k=2}^{2r} (CA^{\mu_k} BW) \right) \\ &= \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r} \mathbf{E}(WCA^{\mu_{\beta_1}} BW) \prod_{k=2}^r \mathbf{E}(CA^{\mu_{\alpha_k}} BWCA^{\mu_{\beta_k}} BW) \\ &= \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r} B^T (A^{\mu_{\beta_1}})^T C^T \prod_{k=2}^r (CA^{\mu_{\alpha_k}} BB^T (A^{\mu_{\beta_k}})^T C^T). \end{aligned} \quad (2.16)$$

6 Moments of Gaussian random matrices

Using the last representation and bearing (2.8) and (2.11) in mind, it follows that for any $r \in \mathbb{N}$ and $\sigma \in \mathbb{Z}_+$,

$$\begin{aligned} & \sum_{(\mu_1, \dots, \mu_{2r+1}) \in \mathcal{M}_\sigma^{2r+1}} A^{\mu_1} \mathbf{E} \prod_{k=2}^{2r+1} B W C A^{\mu_k} \\ &= \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r} \sum_{(\mu_1, \dots, \mu_{2r+1}) \in \mathcal{M}_\sigma^{2r+1}} A^{\mu_1} B B^T (A^{\mu_{\beta_1}})^T C^T C A^{\mu_{2r+1}} \prod_{k=2}^r (C A^{\mu_{\alpha_k}} B B^T (A^{\mu_{\beta_k}})^T C^T). \end{aligned} \quad (2.17)$$

Now note that each of the sets (2.9) is invariant under permutations of the entries of the multi-indices. Therefore, the rightmost sum in (2.17), taken over the set $\mathcal{M}_\sigma^{2r+1}$ for a given permutation $(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r$, does not depend on the latter. Moreover, the sum is equal to the matrix $H_{r,\sigma}$ defined by (2.6), that is,

$$\Sigma_{r,\sigma} = \sum_{(\mu_1, \dots, \mu_{2r+1}) \in \mathcal{M}_\sigma^{2r+1}} A^{\mu_1} B B^T (A^{\mu_2})^T C^T C A^{\mu_{2r+1}} \prod_{k=2}^r (C A^{\mu_{2k-1}} B B^T (A^{\mu_{2k}})^T C^T) = H_{r,\sigma}. \quad (2.18)$$

Indeed, for any u from the set U in (2.3),

$$\begin{aligned} \mathcal{P}_r(u) &= \sum_{\sigma \in \mathbb{Z}_+} u^\sigma \Sigma_{r,\sigma} \\ &= \sum_{(\mu_1, \dots, \mu_{2r+1}) \in \mathbb{Z}_+^{2r+1}} (uA)^{\mu_1} B B^T (uA^T)^{\mu_2} C^T C (uA)^{\mu_{2r+1}} \prod_{k=2}^r (C (uA)^{\mu_{2k-1}} B B^T (uA^T)^{\mu_{2k}} C^T) \\ &= \mathcal{A}(u) B B^T (\mathcal{A}(u))^T C^T C \mathcal{A}(u) |B^T (\mathcal{A}(u))^T C^T|^{2(r-1)} = \mathcal{H}_r(u), \end{aligned} \quad (2.19)$$

where we have used (2.2), (2.4), and (2.5). That the functions \mathcal{P}_r and \mathcal{H}_r coincide on their common domain of analyticity implies the validity of (2.18) for every $\sigma \in \mathbb{Z}_+$. Consequently, since $\#\mathcal{P}_r = (2r-1)!!$, the relationship (2.17) reads

$$\sum_{(\mu_1, \dots, \mu_{2r+1}) \in \mathcal{M}_\sigma^{2r+1}} A^{\mu_1} \mathbf{E} \prod_{k=2}^{2r+1} B W C A^{\mu_k} = (2r-1)!! H_{r,\sigma}. \quad (2.20)$$

Substituting the last representation back to (2.8), we arrive at (2.7), completing the proof of Theorem 2.1. \square

Theorem 2.1 makes it easy to compute the power and exponential moments of the matrix (1.1) if its expectation is zero. To formulate the corollary below, note that $\mathcal{A}(0) = I_n$ by (2.2). Hence, by (2.4)–(2.6),

$$H_{r,0} = D^r = (\text{Tr} D)^{r-1} D, \quad (2.21)$$

where

$$D = \mathcal{E}(0) = B B^T C^T C. \quad (2.22)$$

In (2.21), we have also used the idempotency of the matrix $D/\text{Tr}D$ inherited from $\mathcal{E}(u)$, with $\text{Tr}D = |CB|^2$.

COROLLARY 2.4. *Let X be given by (1.1)–(2.1), where A is zero. Then for any $r \in \mathbb{N}$ and $t \in \mathbb{R}$,*

$$\mathbf{E}X^{2r} = (2r - 1)!!(\text{Tr}D)^{r-1}D, \quad (2.23)$$

$$\mathbf{E}e^{tX} = I_n + \frac{\exp(\text{Tr}Dt^2/2) - 1}{\text{Tr}D}D, \quad (2.24)$$

where the matrix D is defined by (2.22).

Proof. Since $A = 0_{n \times n}$, with $0_{a \times b}$ denoting the $(a \times b)$ -matrix of zeros, the map \mathcal{A} in (2.2) is identically constant. Hence, the matrices (2.6) satisfy $H_{r,k} = 0_{n \times n}$ for all $r, k \in \mathbb{N}$. Therefore (2.7) yields

$$\mathbf{E}X^{2r} = (2r - 1)!!H_{r,0} \quad (2.25)$$

which, by (2.21)–(2.22), implies (2.23). To prove (2.24), note that in the case considered, all the odd power moments of X are zero since its distribution inherits from \mathcal{N}_m the symmetry about the origin. Consequently, by (2.23) and by the identity $(2r)! = (2r - 1)!!r!2^r$,

$$\begin{aligned} \mathbf{E}e^{tX} &= I_n + \sum_{r \in \mathbb{N}} \frac{t^{2r}}{(2r)!} \mathbf{E}X^{2r} \\ &= I_n + \frac{D}{\text{Tr}D} \sum_{r \in \mathbb{N}} \frac{(2r - 1)!!}{(2r)!} (\text{Tr}Dt^2)^r \\ &= I_n + \frac{D}{\text{Tr}D} \sum_{r \in \mathbb{N}} \frac{1}{r!} (\text{Tr}Dt^2/2)^r. \end{aligned} \quad (2.26)$$

The last representation implies (2.24), and the proof of the corollary is complete. \square

In the scalar case $n = 1$, the right-hand side of (2.24) indeed yields the moment-generating function $\exp(Dt^2/2)$ of the univariate Gaussian distribution with zero mean and variance D .

3. System theoretic interpretation

Three matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{q \times n}$ generate a state-space realization of a linear operator which maps a sequence $i = (i_k)_{k \in \mathbb{Z}_+} \in (\mathbb{R}^m)^{\mathbb{Z}_+}$ to $o = (o_k)_{k \in \mathbb{Z}_+} \in (\mathbb{R}^q)^{\mathbb{Z}_+}$ through the equations

$$\sigma_k = A\sigma_{k-1} + Bi_k, \quad o_k = C\sigma_k, \quad (3.1)$$

8 Moments of Gaussian random matrices

where $\sigma = (\sigma_k)_{k \in \mathbb{Z}_+} \in (\mathbb{R}^n)^{\mathbb{Z}_+}$ satisfies the initial condition $\sigma_{-1} = 0_{n \times 1}$. The linear operator is denoted by the ordered triplet (A, B, C) or, interchangeably, by

$$\left[\begin{array}{c|c} A & B \\ \hline C & 0_{q \times m} \end{array} \right]. \quad (3.2)$$

It is interpreted as a linear *discrete*-time invariant (LDTI) causal system with input i , output o , and the internal state σ ; see, for example, [4, pages 90–93] or [8, page 35]. Recall that o is the convolution of i with the impulse response $(d_k)_{k \in \mathbb{Z}_+} \in (\mathbb{R}^{q \times m})^{\mathbb{Z}_+}$ given by

$$d_k = CA^k B = \frac{1}{k!} \left. \frac{d^k \mathcal{D}(u)}{du^k} \right|_{u=0}, \quad (3.3)$$

where $\mathcal{D} : U \rightarrow \mathbb{R}^{q \times m}$ is the transfer function of the system defined on the set (2.3) by

$$\mathcal{D}(u) = \sum_{k \in \mathbb{Z}_+} u^k d_k = C \mathcal{A}(u) B, \quad (3.4)$$

and $\mathcal{A}(u)$ is given by (2.2). On the other hand, (3.2) can be regarded as a state-space realization of a linear *continuous*-time invariant (LCTI) causal system which maps a locally integrable \mathbb{R}^m -valued input $i = (i_t)_{t \in \mathbb{R}_+}$ to an absolutely continuous \mathbb{R}^q -valued output $o = (o_t)_{t \in \mathbb{R}_+}$ by

$$\frac{d}{dt} \sigma_t = A \sigma_t + B i_t, \quad o_t = C \sigma_t, \quad (3.5)$$

with $\sigma_0 = 0_{n \times 1}$. The corresponding map $i \mapsto o$ is the Volterra integral operator

$$o_t = \int_0^t c_{t-s} i_s ds. \quad (3.6)$$

Its kernel, $c_t = C e^{tA} B$, may be interpreted as the impulse response of the LCTI system. Accordingly, the transfer function of the system is defined as the Laplace transform of the kernel so that

$$\mathcal{C}(v) = \int_{\mathbb{R}_+} e^{-vt} c_t dt = C(vI_n - A)^{-1} B, \quad v > \ln \mathbf{r}(e^A). \quad (3.7)$$

For simplicity, we consider the transfer functions only for real values of their arguments. The transfer functions (3.4) and (3.7) of the LDTI and LCTI systems generated by the same state-space realization triplet are related by

$$\mathcal{C}(v) = v^{-1} \mathcal{D}(v^{-1}). \quad (3.8)$$

Here v^{-1} is the transfer function of the LCTI system $\mathbf{I} = (0, 1, 1)$, which we will refer to as the *integrator*. If applied to \mathbb{R}^m -valued functions, the integrator acts coordinate-wise so as to correspond to the LCTI system $\mathbf{I}_m = (0_{m \times m}, I_m, I_m)$.

As it is well known, each of the classes of LDTI and LCTI systems with rational transfer functions is closed under the associated composition which corresponds to the convolution of their impulse responses or, equivalently, to the multiplication of their transfer functions. In state space, the operation is described immediately below.

LEMMA 3.1. *Let $\Sigma_1 = (A_1, B_1, C_1)$ and $\Sigma_2 = (A_2, B_2, C_2)$ be conformable state-space realization triplets, so that the output of Σ_1 is the input to Σ_2 . Then the compositions of the corresponding LDTI and LCTI systems can be computed, respectively, as*

$$\Sigma_2 \textcircled{d} \Sigma_1 = \left[\begin{array}{c|c|c} A_1 & 0 & B_1 \\ \hline B_2 C_1 A_1 & A_2 & B_2 C_1 B_1 \\ \hline 0 & C_2 & 0 \end{array} \right], \quad (3.9)$$

$$\Sigma_2 \textcircled{c} \Sigma_1 = \left[\begin{array}{c|c|c} A_1 & 0 & B_1 \\ \hline B_2 C_1 & A_2 & 0 \\ \hline 0 & C_2 & 0 \end{array} \right]. \quad (3.10)$$

If the state-space realization triplets on the right-hand side of (3.9) and (3.10) are both considered as LCTI systems, then these latter satisfy the identity

$$\Sigma_2 \textcircled{c} \Sigma_1 = \mathbf{I} \textcircled{c} (\Sigma_2 \textcircled{d} \Sigma_1). \quad (3.11)$$

To make the exposition self-contained, we prove Lemma 3.1 in Appendix B. Note that the relations (3.9)–(3.11) are all understood as equalities between linear operators, not between their state-space realization matrices. In this connection, recall that (TAT^{-1}, TB, CT^{-1}) and (A, B, C) represent the same system for any conformable nonsingular matrix T .

In (3.11), slightly sacrificing notational consistency, \mathbf{I} stands for the class of integrators \mathbf{I}_m of the appropriate dimension m . With this interpretation, \mathbf{I} commutes with any LCTI system Σ in the sense that

$$\mathbf{I}_q \textcircled{c} \Sigma = \Sigma \textcircled{c} \mathbf{I}_m, \quad (3.12)$$

where m and q are the dimensions of the input and output of Σ , respectively.

Let

$$\textcircled{d}^r \Sigma = \underbrace{\Sigma \textcircled{d} \cdots \textcircled{d} \Sigma}_r, \quad (3.13)$$

respectively,

$$\textcircled{c}^r \Sigma = \underbrace{\Sigma \textcircled{c} \cdots \textcircled{c} \Sigma}_r, \quad (3.14)$$

denote the r -fold iterate of the LDTI, respectively LCTI, system Σ , so that $\textcircled{d}^1 \Sigma$ and $\textcircled{c}^1 \Sigma$ coincide with Σ . Using the constructs defined above, it is possible to reformulate Theorem 2.1 in system-theoretic terms.

PROPOSITION 3.2. For any $r \in \mathbb{N}$, the $\mathbb{R}^{n \times n}$ -valued sequence $(H_{r,k})_{k \in \mathbb{Z}_+}$ defined by (2.6) is the impulse response of the LDTI system

$$H_r = (\textcircled{D}^r E) \textcircled{D} F. \quad (3.15)$$

Here,

$$E = F \textcircled{D} G = (\mathbf{A}, \mathbf{B}, \mathbf{C}), \quad (3.16)$$

$$F = (A, I_n, I_n), \quad (3.17)$$

$$G = (A^T, C^T C, BB^T), \quad (3.18)$$

and the state-space realization matrices $\mathbf{A} \in \mathbb{R}^{2n \times 2n}$, $\mathbf{B} \in \mathbb{R}^{2n \times n}$, and $\mathbf{C} \in \mathbb{R}^{n \times 2n}$ are given by

$$\mathbf{A} = \left[\begin{array}{c|c} A^T & 0_{n \times n} \\ \hline BB^T A^T & A \end{array} \right], \quad \mathbf{B} = \left[\begin{array}{c} C^T C \\ \hline D \end{array} \right], \quad \mathbf{C} = \left[0_{n \times n} \mid I_n \right], \quad (3.19)$$

where the matrix D is defined by (2.22).

Proof. By (2.2), \mathcal{A} and $BB^T \mathcal{A}^T C^T C$ are the transfer functions of the LDTI systems F and G defined by (3.17) and (3.18). Hence, their product \mathcal{E} in (2.4) is the transfer function of the LDTI system E given by (3.16). Its state-space realization matrices (3.19) are obtained by applying (3.9) of Lemma 3.1. It now remains to note that \mathcal{H}_r in (2.5) is the transfer function of the LDTI system H_r in (3.15), and the proposition is proved. \square

Besides the system-theoretic interpretation of Theorem 2.1, Proposition 3.2 encapsulates a useful formalism for computing the power moments of the matrix X in (1.1). More precisely,

$$H_{r,k} = \mathbf{C}_r \mathbf{A}_r^k \mathbf{B}_r, \quad (3.20)$$

where $\mathbf{A}_r \in \mathbb{R}^{(2r+1)n \times (2r+1)n}$, $\mathbf{B}_r \in \mathbb{R}^{(2r+1)n \times n}$, and $\mathbf{C}_r \in \mathbb{R}^{n \times (2r+1)n}$ are the state-space realization matrices of the LDTI system (3.15), that is,

$$H_r = (\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r). \quad (3.21)$$

The three matrices can be calculated recursively in r . Indeed, applying (3.9) of Lemma 3.1 to the recurrence relation $H_{r+1} = E \textcircled{D} H_r$ and using (3.19), we arrive at the equations

$$\mathbf{A}_{r+1} = \left[\begin{array}{c|c} \mathbf{A}_r & 0_{(2r+1)n \times 2n} \\ \hline \mathbf{B} \mathbf{C}_r \mathbf{A}_r & \mathbf{A} \end{array} \right], \quad \mathbf{B}_{r+1} = \left[\begin{array}{c} \mathbf{B}_r \\ \hline \mathbf{B} \mathbf{C}_r \mathbf{B}_r \end{array} \right], \quad (3.22)$$

$$\mathbf{C}_r = \left[0_{n \times 2rn} \mid I_n \right], \quad (3.23)$$

with initial conditions

$$\mathbf{A}_0 = A, \quad \mathbf{B}_0 = \mathbf{C}_0 = I_n. \quad (3.24)$$

Clearly, both \mathbf{A}_r and \mathbf{B}_r in (3.22) are submatrices of \mathbf{A}_{r+1} and \mathbf{B}_{r+1} , respectively, with \mathbf{A}_r being block lower triangular. From (3.19) and from the special structure of the matrices \mathbf{C}_r in (3.23), it follows that

$$\mathbf{C}_{r+1}\mathbf{A}_{r+1} = \left[\mathbf{DC}_r\mathbf{A}_r \mid \mathbf{CA} \right], \quad \mathbf{C}_{r+1}\mathbf{B}_{r+1} = \mathbf{DC}_r\mathbf{B}_r, \quad (3.25)$$

where $\mathbf{CA} = [\mathbf{BB}^T\mathbf{A}^T \mid \mathbf{A}]$. In particular, by a straightforward induction, the rightmost relations in (3.24) and in (3.25) imply that $\mathbf{C}_r\mathbf{B}_r = \mathbf{D}^r$ for all $r \in \mathbb{Z}_+$.

Note that the proposed system-theoretic implementation (3.20)–(3.25) of the solution (2.7) to the problem of computing the power moment $\mathbf{E}X^s$ of (1.1) is expensive for large s , because the order of the matrix \mathbf{A}_r grows linearly in r .

4. Exponential moments

The theorem below shows that the solution to the exponential moments problem also admits a system-theoretic representation, intrinsically related to that in Proposition 3.2.

THEOREM 4.1. *The exponential moments of the Gaussian random matrix X defined by (1.1)–(2.1) are representable as*

$$\mathbf{E}e^{tX} = e^{tA} + \sum_{r \in \mathbb{N}} (2r-1)!! M_r(t), \quad t \in \mathbb{R}_+, \quad (4.1)$$

where $M_r : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$ is the impulse response of the LCTI system

$$\mathbf{H}_r = (\odot^r \mathbf{E}) \odot F. \quad (4.2)$$

Here

$$\mathbf{E} = F \odot G = (\mathbf{A}, \mathbf{B}, \mathbf{C}), \quad (4.3)$$

where F and G are defined by (3.17)–(3.18), and the matrix \mathbf{C} is given in (3.19), while $\mathbf{A} \in \mathbb{R}^{2n \times 2n}$ and $\mathbf{B} \in \mathbb{R}^{2n \times n}$ are defined by

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}^T & \mathbf{0}_{n \times n} \\ \hline \mathbf{BB}^T & \mathbf{A} \end{array} \right], \quad \mathbf{B} = \left[\begin{array}{c} \mathbf{C}^T \mathbf{C} \\ \hline \mathbf{0}_{n \times n} \end{array} \right]. \quad (4.4)$$

The proof of Theorem 4.1 is given in Appendix C. Applying (3.10) of Lemma 3.1 to the recurrence relation $\mathbf{H}_{r+1} = \mathbf{E} \odot \mathbf{H}_r$ which follows from (4.2), we obtain that

$$\mathbf{H}_r = (\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r), \quad (4.5)$$

where the matrix \mathbf{C}_r is given by (3.23) as before, while \mathbf{A}_r and \mathbf{B}_r satisfy the equations

$$\mathbf{A}_{r+1} = \left[\begin{array}{c|c} \mathbf{A}_r & \mathbf{0}_{(2r+1)n \times 2n} \\ \hline \mathbf{B}_r \mathbf{C}_r & \mathbf{A} \end{array} \right], \quad \mathbf{B}_{r+1} = \left[\begin{array}{c} \mathbf{B}_r \\ \hline \mathbf{0}_{2n \times n} \end{array} \right], \quad (4.6)$$

with

$$\mathbf{A}_0 = \mathbf{A}, \quad \mathbf{B}_0 = \mathbf{I}_n. \quad (4.7)$$

In view of (4.5), the impulse response M_r of the LCTI system (4.2) takes the form

$$M_r(t) = \mathbf{C}_r e^{t\mathbf{A}_r} \mathbf{B}_r. \quad (4.8)$$

By a straightforward induction, (4.6)–(4.7) and (4.4) imply that for every $r \in \mathbb{N}$,

$$\mathbf{A}_r = \begin{bmatrix} \begin{array}{c|c|c|c|c} A & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ \hline C^T C & A^T & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ \hline \mathbf{0}_{n \times n} & BB^T & A & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ \hline \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & C^T C & A^T & \mathbf{0}_{n \times n} \\ \hline \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & BB^T & A \end{array} \\ \vdots \end{bmatrix}, \quad \mathbf{B}_r = \begin{bmatrix} I_n \\ \mathbf{0}_{2rn \times n} \end{bmatrix}. \quad (4.9)$$

In particular, \mathbf{A}_r is the principal submatrix, of order $(2r+1)n$, of an infinite block two-diagonal matrix where the blocks A alternate with A^T along the main diagonal and $C^T C$ alternate with BB^T on the subdiagonal.

Let \otimes denote the Kronecker product of matrices, and χ_1^2 stands for the χ^2 -distribution [13, pages 183–184] with one degree of freedom.

THEOREM 4.2. *The exponential moments (4.1) of the Gaussian random matrix X defined by (1.1)–(2.1) are given by*

$$\mathbf{E} e^{tX} = P_n \mathbf{E} \exp \left(t \left[\begin{array}{c|c} A & \omega_1 BB^T \\ \hline \omega_2 C^T C & A^T \end{array} \right] \right) P_n^T, \quad t \in \mathbb{R}_+. \quad (4.10)$$

Here ω_1 and ω_2 are arbitrary scalar random variables satisfying $\omega_1 \omega_2 \triangleright \triangleleft \chi_1^2$, and

$$P_n = [1 | 0] \otimes I_n = [I_n | \mathbf{0}_{n \times n}]. \quad (4.11)$$

We prove Theorem 4.2 in Appendix D. Note that (4.10) extends from \mathbb{R}_+ to the complex plane by a standard argument using the uniqueness theorem for analytic functions.

The significance of Theorem 4.2 is as follows. Despite the presence of two random variables on the right-hand side of (4.10), the m -variate integral of the matrix exponential in $\mathbf{E} e^{tX}$ is reduced to a *univariate* integral of another matrix exponential along a pencil of matrices. Indeed, if ω is a standard normal random variable, then $\omega^2 \triangleright \triangleleft \chi_1^2$, and hence, application of (4.10) with $\omega_1 = \omega_2 = \omega$ gives

$$\mathbf{E} e^{tX} = P_n \mathbf{E} \exp \left(t \left[\begin{array}{c|c} A & \mathbf{0}_{n \times n} \\ \hline \mathbf{0}_{n \times n} & A^T \end{array} \right] + \omega t \left[\begin{array}{c|c} \mathbf{0}_{n \times n} & BB^T \\ \hline C^T C & \mathbf{0}_{n \times n} \end{array} \right] \right) P_n^T. \quad (4.12)$$

The matrix exponential on the right-hand side of (4.12) is applied to a Gaussian random matrix whose order is twice the order of X . However, in contrast to the m -dimensional source (2.1) in the original matrix, the random perturbation ω in the new matrix is one-dimensional. By the discussion above, (4.12) remains valid if we replace ω with $|\omega|$, yielding

$$\mathbf{E}e^{tX} = \sqrt{\frac{2}{\pi}} P_n \int_{\mathbb{R}_+} \exp\left(t \left[\frac{A}{wC^T C} \middle| \frac{wBB^T}{A^T} \right] - \frac{w^2}{2} I_{2n}\right) dw P_n^T. \quad (4.13)$$

Alternatively, putting $\omega_1 \equiv 1$ and $\omega_2 \gg \chi_1^2$ in (4.10) and recalling the PDF of the χ_1^2 -law $e^{-z/2}/\sqrt{2\pi z}$, we obtain

$$\mathbf{E}e^{tX} = \frac{1}{\sqrt{2\pi}} P_n \int_{\mathbb{R}_+} \exp\left(t \left[\frac{A}{zC^T C} \middle| \frac{BB^T}{A^T} \right] - \frac{z}{2} I_{2n}\right) \frac{dz}{\sqrt{z}} P_n^T. \quad (4.14)$$

Further reduction of the integrals on the right-hand side of (4.13) and (4.14) constitutes, to our best knowledge, an open problem. From the practical point of view, these univariate integrals are, in principle, amenable to numerical computation using the methods of [10].

Appendices

A. Proof of Lemma 2.3

Consider the moment-generating function of the random vector ξ which maps $u = (u_k)_{k \in \mathcal{L}_r} \in \mathbb{R}^{2r}$ to

$$\begin{aligned} \psi(u) &= \mathbf{E}e^{u^T \xi} = \exp\left(\frac{1}{2} u^T \Sigma u\right) = \sum_{j \in \mathbb{Z}_+} \frac{1}{j!} \left(\frac{1}{2} \sum_{\alpha, \beta \in \mathcal{L}_r} \sigma_{\alpha\beta} u_\alpha u_\beta\right)^j \\ &= \sum_{j \in \mathbb{Z}_+} \frac{1}{(2j)!!} \sum_{\alpha_1, \beta_1, \dots, \alpha_j, \beta_j \in \mathcal{L}_r} \prod_{k=1}^j \sigma_{\alpha_k \beta_k} u_{\alpha_k} u_{\beta_k}, \end{aligned} \quad (A.1)$$

where we have used the identity $j!2^j = (2j)!!$. Only those terms on the right-hand side of (A.1) for which $j = r$ and $(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r)$ is a permutation of the set \mathcal{L}_r in (2.13), contribute to the expectation on the left-hand side of (2.14). Thus

$$\mathbf{E} \prod_{k=1}^{2r} \xi_k = \frac{\partial^{2r} \psi(u)}{\partial u_1 \cdots \partial u_{2r}} \Big|_{u=0_{2r \times 1}} = \frac{1}{(2r)!!} \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \Pi_{2r}} \prod_{k=1}^r \sigma_{\alpha_k \beta_k}, \quad (A.2)$$

where Π_{2r} is the class of all possible permutations of \mathcal{L}_r . For any $\gamma = (\gamma_1, \dots, \gamma_r) \in \Pi_r$, let E_γ be the operator which synchronously permutes the odd and even entries of a $(2r)$ -dimensional vector so that

$$E_\gamma(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) = (\alpha_{\gamma_1}, \beta_{\gamma_1}, \dots, \alpha_{\gamma_r}, \beta_{\gamma_r}). \quad (A.3)$$

Let T_k be the transposition which exchanges the $(2k - 1)$ th and $(2k)$ th entries, leaving all others fixed. Clearly, the maps E_γ and T_1, \dots, T_r commute with each other. For any binary word $\delta = (\delta_1, \dots, \delta_r) \in \{0, 1\}^r$ of length r , let

$$T^\delta = T_1^{\delta_1} \circ \dots \circ T_r^{\delta_r}. \quad (\text{A.4})$$

As a function of $p = (\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \Pi_{2r}$, the quantity $\prod_{k=1}^r \sigma_{\alpha_k \beta_k}$ in (A.2) is invariant with respect to the group of transformations $p \mapsto (E_\gamma \circ T^\delta)(p)$; the invariance under T_k follows from the symmetry of the covariance matrix Σ . The action of the group partitions the set Π_{2r} into equivalence classes, invariant under the group. Each of these classes is representable as the orbit $C_{r,p}$ of the group on a unique regular permutation $p \in \mathcal{P}_r$, where

$$C_{r,p} = \{(E_\gamma \circ T^\delta)(p) : \gamma \in \Pi_r, \delta \in \{0, 1\}^r\}. \quad (\text{A.5})$$

Clearly, $\#C_{r,p} = r!2^r = (2r)!!$. Therefore, the sum on the right-hand side of (A.2) can be reduced to a sum over regular permutations as follows:

$$\mathbf{E} \prod_{k=1}^{2r} \xi_k = \frac{1}{(2r)!!} \sum_{p \in \mathcal{P}_r} \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in C_{r,p}} \prod_{k=1}^r \sigma_{\alpha_k \beta_k} = \sum_{(\alpha_1, \beta_1, \dots, \alpha_r, \beta_r) \in \mathcal{P}_r} \prod_{k=1}^r \sigma_{\alpha_k \beta_k}. \quad (\text{A.6})$$

The last representation coincides with (2.14), thereby completing the proof of the lemma.

B. Proof of Lemma 3.1

Let σ , i , and ω be, respectively, the internal state, input, and output of the LDTI system Σ_1 described by

$$\sigma_k = A_1 \sigma_{k-1} + B_1 i_k, \quad \omega_k = C_1 \sigma_k. \quad (\text{B.1})$$

The sequence ω is also the input to the LDTI system Σ_2 whose internal state s and output o are governed by

$$s_k = A_2 s_{k-1} + B_2 \omega_k, \quad o_k = C_2 s_k. \quad (\text{B.2})$$

Substitution of (B.1) to the leftmost equation in (B.2) transforms the latter to $s_k = B_2 C_1 A_1 \sigma_{k-1} + A_2 s_{k-1} + B_2 C_1 B_1 i_k$. Combining the last equation with (B.1) and with (B.2), and assembling the internal state sequences σ and s to

$$S = \begin{bmatrix} \sigma \\ s \end{bmatrix}, \quad (\text{B.3})$$

one verifies that the operator $\Sigma_2 \textcircled{D} \Sigma_1$ relating o with i is described by the equations

$$S_k = \left[\begin{array}{c|c} A_1 & 0 \\ \hline B_2 C_1 A_1 & A_2 \end{array} \right] S_{k-1} + \left[\begin{array}{c} B_1 \\ B_2 C_1 B_1 \end{array} \right] i_k, \quad (\text{B.4})$$

$$o_k = \left[0 \mid C_2 \right] S_k,$$

which yield (3.9). The proof of (3.10) below almost verbatim replicates the treatment of the discrete-time case above. The LCTI systems Σ_1 and Σ_2 are governed by

$$\frac{d}{dt}\sigma_t = A_1\sigma_t + B_1i_t, \quad \omega_t = C_1\sigma_t, \quad (\text{B.5})$$

$$\frac{d}{dt}s_t = A_2s_t + B_2\omega_t, \quad o_t = C_2s_t. \quad (\text{B.6})$$

Substitution of the rightmost equation in (B.5) to the leftmost equation in (B.6) yields $(d/dt)s_t = B_2C_1\sigma_t + A_2s_t$. Combining this last equation with the remaining ones in (B.5) and in (B.6), and assembling σ and s as in (B.3), we obtain

$$\begin{aligned} \frac{d}{dt}S_t &= \left[\begin{array}{c|c} A_1 & 0 \\ \hline B_2C_1 & A_2 \end{array} \right] S_t + \left[\begin{array}{c} B_1 \\ 0 \end{array} \right] i_t, \\ o_t &= \left[0 \mid C_2 \right] S_t, \end{aligned} \quad (\text{B.7})$$

from whence (3.10) follows. To prove (3.11), we write the transfer function of the LCTI system $\Sigma_2 \circledcirc \Sigma_1$ computed on the right-hand side of (3.10) as

$$\begin{aligned} \mathcal{C}_{\Sigma_2 \circledcirc \Sigma_1}(v) &= \mathcal{C}_{\Sigma_2}(v)\mathcal{C}_{\Sigma_1}(v) \\ &= v^{-2}\mathcal{D}_{\Sigma_2}(v^{-1})\mathcal{D}_{\Sigma_1}(v^{-1}) \\ &= v^{-2}\mathcal{D}_{\Sigma_2 \circledast \Sigma_1}(v^{-1}) = v^{-1}\mathcal{C}_{\Sigma_2 \circledast \Sigma_1}(v), \end{aligned} \quad (\text{B.8})$$

where we have used (3.8) twice. The right-hand side of (B.8) is the transfer function of the LCTI system $\mathbf{I} \circledcirc (\Sigma_2 \circledast \Sigma_1)$, establishing (3.10) and completing the proof of the lemma.

C. Proof of Theorem 4.1

From (2.7) of Theorem 2.1, it follows that

$$\begin{aligned} \mathbf{E}e^{tX} &= e^{tA} + \sum_{s=2}^{+\infty} \frac{t^s}{s!} \sum_{r=1}^{\lfloor s/2 \rfloor} (2r-1)!! H_{r,s-2r} \\ &= e^{tA} + \sum_{r \in \mathbb{N}} (2r-1)!! M_r(t), \end{aligned} \quad (\text{C.1})$$

where

$$M_r(t) = \sum_{k \in \mathbb{Z}_+} \frac{t^{2r+k}}{(2r+k)!} H_{r,k}. \quad (\text{C.2})$$

The interchangeability of sums in (C.1), as well as the absolute summability of the series, follows from the bound

$$\sup_{r \in \mathbb{N}, k \in \mathbb{Z}_+} \frac{\ln \|H_{r,k}\|}{r+k} < +\infty, \quad (\text{C.3})$$

where $\|\cdot\|$ is a matrix norm. For every $r \in \mathbb{N}$, the function M_r in (C.2) coincides with the $(2r)$ -fold integral of the function $J_r : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$ defined by

$$J_r(t) = \sum_{k \in \mathbb{Z}_+} \frac{t^k}{k!} H_{r,k}. \quad (\text{C.4})$$

More precisely,

$$M_r(t) = \underbrace{\int_0^t \int_0^{t_2} \cdots \int_0^{t_3} \int_0^{t_2}}_{2r \text{ times}} J_r(t_1) dt_1 dt_2 \times \cdots \times dt_{2r-1} dt_{2r}. \quad (\text{C.5})$$

Substituting (3.20) into (C.4) yields

$$J_r(t) = \mathbf{C}_r \sum_{k \in \mathbb{Z}_+} \frac{(t\mathbf{A}_r)^k}{k!} \mathbf{B}_r = \mathbf{C}_r e^{t\mathbf{A}_r} \mathbf{B}_r. \quad (\text{C.6})$$

Thus, J_r is the impulse response associated with the state-space realization triplet (3.21), this time interpreted as an LCTI system. Therefore, by (C.5) and by (3.15)–(3.16), the function M_r is the impulse response of the LCTI system

$$\begin{aligned} (\mathbb{C}^{2r} \mathbf{I}) \mathbb{C} H_r &= (\mathbb{C}^{2r} \mathbf{I}) \mathbb{C} (F \mathbb{D} G \mathbb{D} H_{r-1}) \\ &= (\mathbb{C}^{2r-1} \mathbf{I}) \mathbb{C} F \mathbb{C} (G \mathbb{D} H_{r-1}) \\ &= (\mathbb{C}^{2(r-1)} \mathbf{I}) \mathbb{C} F \mathbb{C} G \mathbb{C} H_{r-1} \\ &= \mathbf{E} \mathbb{C} (\mathbb{C}^{2(r-1)} \mathbf{I}) \mathbb{C} H_{r-1} \\ &= (\mathbb{C}^r \mathbf{E}) \mathbb{C} F = H_r. \end{aligned} \quad (\text{C.7})$$

Here, \mathbf{I} is the integrator (see Section 3), and we have used the identity (3.11), the commutative property of the integrator (3.12), and the definitions (4.2) and (4.3). It follows from (3.10) of Lemma 3.1 that the state-space realization matrices of the LCTI system \mathbf{E} can be computed by (4.4) and by the rightmost relation in (3.19), completing the proof of the theorem.

D. Proof of Theorem 4.2

To prove the theorem, we will need the two subsidiary results immediately below.

LEMMA D.1. *Let $\Sigma = (A, B, C)$ be an LCTI system with equally dimensioned input and output. Then*

$$\sum_{r \in \mathbb{N}} \mathbb{C}^r \Sigma = (\mathcal{F} - \Sigma)^{-1} - \mathcal{F} = (A + BC, B, C), \quad (\text{D.1})$$

where \mathcal{F} is the identity operator.

The convergence of the series on the left-hand side of (D.1) is ensured by general results on the Neumann series associated with linear Volterra integral equations of the second kind. The rest of Lemma D.1 is well known in system theory; see, for example, [4, page 656]. To make the exposition self-contained, we prove it below following the line of [2, Proposition 6.4.1, page 81].

Proof of Lemma D.1. Let i and σ be the input and the internal state of Σ , respectively. Then the operator $\mathcal{F} - \Sigma : i \mapsto o$ is governed by

$$\frac{d}{dt}\sigma_t = A\sigma_t + Bi_t, \quad o_t = i_t - C\sigma_t, \quad (\text{D.2})$$

and is invertible on the space of locally integrable functions. Expressing i in terms of σ and o from the rightmost equation in (D.2) and substituting the result into the leftmost equation, we obtain that the inverse operator $(\mathcal{F} - \Sigma)^{-1} : o \mapsto i$ is described by

$$\frac{d}{dt}\sigma_t = (A + BC)\sigma_t + Bo_t, \quad i_t = C\sigma_t + o_t. \quad (\text{D.3})$$

The state-space representation of $(\mathcal{F} - \Sigma)^{-1} - \mathcal{F}$ is now obtained by subtracting o_t from both parts of the rightmost equation in (D.3), that establishes (D.1) and completes the proof of the lemma. \square

LEMMA D.2. Let Q be a square matrix partitioned into four blocks as follows:

$$Q = \left[\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline Q_{21} & Q_{22} \end{array} \right], \quad (\text{D.4})$$

where $Q_{11} \in \mathbb{R}^{a \times a}$ and $Q_{22} \in \mathbb{R}^{b \times b}$. Let f be a function, analytic in a neighbourhood of the origin, and such that the radius of convergence of its Taylor series expansion is greater than $r(Q)$. Then for every $\lambda \in \mathbb{R} \setminus \{0\}$,

$$Pf(Q_\lambda)P^T = Pf(Q)P^T, \quad (\text{D.5})$$

where

$$Q_\lambda = \left[\begin{array}{c|c} Q_{11} & \lambda Q_{12} \\ \hline \lambda^{-1} Q_{21} & Q_{22} \end{array} \right], \quad P = \left[I_a \mid 0_{a \times b} \right]. \quad (\text{D.6})$$

The assertion of the last lemma means the invariance of the appropriate principal submatrices of $f(Q)$ under the balanced rescaling of the off-diagonal blocks of Q controlled by the parameter λ .

Proof of Lemma D.2. For any $\lambda \neq 0$, the matrix Q_λ is obtained from $Q = Q_1$ by the similarity transformation

$$Q_\lambda = T_\lambda^{-1}QT_\lambda, \quad T_\lambda = \left[\begin{array}{c|c} I_a & 0_{a \times b} \\ \hline 0_{b \times a} & \lambda I_b \end{array} \right]. \quad (\text{D.7})$$

Hence,

$$f(Q_\lambda) = T_\lambda^{-1} f(Q) T_\lambda. \quad (\text{D.8})$$

Since $PT_\lambda^{-1} = P$ and $T_\lambda P^T = P^T$, the left multiplication by P and the right multiplication by P^T of both parts of (D.8) yield (D.5), completing the proof of the lemma. \square

Now turn to the proof of Theorem 4.2. Let ζ be a χ_1^2 -distributed random variable,

$$\zeta \triangleright \triangleleft \chi_1^2. \quad (\text{D.9})$$

For any $t \in \mathbb{R}_+$, we introduce the $\mathbb{R}^{n \times n}$ -valued random matrix

$$K(\zeta, t) = \sum_{r \in \mathbb{N}} \zeta^r M_r(t), \quad (\text{D.10})$$

where M_r are the impulse responses of the LCTI systems (4.2) defined in Theorem 4.1. With probability one, the series on the right-hand side of (D.10) is absolutely summable and its convergence is uniform in t over any bounded subset of \mathbb{R}_+ . Recalling that for any $r \in \mathbb{N}$, the r th moment of (D.9) is $\mathbf{E}\zeta^r = (2r - 1)!!$, and applying Fubini's theorem, it follows from (4.1) and (D.10) that

$$\mathbf{E}e^{tX} = e^{tA} + \mathbf{E}K(\zeta, t). \quad (\text{D.11})$$

From (4.2), it follows that $K(\zeta, \cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}^{n \times n}$ is the impulse response of the LCTI system

$$L_\zeta = \left(\sum_{r \in \mathbb{N}} \textcircled{\zeta}^r (\zeta E) \right) \textcircled{\zeta} F, \quad (\text{D.12})$$

where ζE is an LCTI system associated with (4.3) as follows:

$$\zeta E = (A, \zeta B, C). \quad (\text{D.13})$$

Both L_ζ and ζE are random linear operators acting on the space of locally integrable \mathbb{R}^n -valued functions on \mathbb{R}_+ . Applying Lemma D.1 to the system $\Sigma = \zeta E$ in (D.13) and using (3.17), one verifies that the operator L_ζ in (D.12) admits a finite-dimensional state-space representation,

$$L_\zeta = (A + \zeta B C, \zeta B, C) \textcircled{\zeta} (A, I_n, I_n). \quad (\text{D.14})$$

Calculating the right-hand side of (D.14) using (3.10) of Lemma 3.1 and recalling (3.19) and (4.4), we obtain

$$\begin{aligned}
 L_\zeta &= \left[\begin{array}{c|c|c|c} A & 0_{n \times 2n} & I_n & \\ \hline \zeta B & A + \zeta BC & 0_{2n \times n} & \\ \hline 0_{n \times n} & C & 0_{n \times n} & \end{array} \right] = \left[\begin{array}{c|c|c|c} A & 0_{n \times n} & 0_{n \times n} & I_n \\ \hline \zeta C^T C & A^T & \zeta C^T C & 0_{n \times n} \\ \hline 0_{n \times n} & BB^T & A & 0_{n \times n} \\ \hline 0_{n \times n} & 0_{n \times n} & I_n & 0_{n \times n} \end{array} \right] \\
 &= \left[\begin{array}{c|c|c|c} A & 0_{n \times n} & 0_{n \times n} & I_n \\ \hline 0_{n \times n} & A & BB^T & I_n \\ \hline 0_{n \times n} & \zeta C^T C & A^T & 0_{n \times n} \\ \hline -I_n & I_n & 0_{n \times n} & 0_{n \times n} \end{array} \right].
 \end{aligned} \tag{D.15}$$

The reduction to the rightmost state-space representation in (D.15) is achieved by the similarity transformation

$$T \left[\begin{array}{c|c|c} A & 0_{n \times n} & 0_{n \times n} \\ \hline \zeta C^T C & A^T & \zeta C^T C \\ \hline 0_{n \times n} & BB^T & A \end{array} \right] T^{-1} = \left[\begin{array}{c|c|c} A & 0_{n \times n} & 0_{n \times n} \\ \hline 0_{n \times n} & A & BB^T \\ \hline 0_{n \times n} & \zeta C^T C & A^T \end{array} \right], \tag{D.16}$$

where

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \otimes I_n, \quad T^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \otimes I_n. \tag{D.17}$$

By the block-diagonal structure of the matrix on the right-hand side of (D.16), the system L_ζ in (D.15) splits into two autonomous LCTI subsystems so that its impulse response (D.10) takes the form

$$K(\zeta, t) = -e^{At} + P_n \exp \left(t \left[\begin{array}{c|c} A & BB^T \\ \hline \zeta C^T C & A^T \end{array} \right] \right) P_n^T, \tag{D.18}$$

where the definition (4.11) has been used. Substituting (D.18) into (D.11) yields

$$\mathbf{E}e^{tX} = P_n \mathbf{E} \exp \left(t \left[\begin{array}{c|c} A & BB^T \\ \hline \zeta C^T C & A^T \end{array} \right] \right) P_n^T. \tag{D.19}$$

Applying Lemma D.2 to the matrix exponential on the right-hand side of (D.19), one verifies that

$$P_n \exp \left(t \left[\begin{array}{c|c} A & BB^T \\ \hline \zeta C^T C & A^T \end{array} \right] \right) P_n^T = P_n \exp \left(t \left[\begin{array}{c|c} A & \omega_1 BB^T \\ \hline \omega_2 C^T C & A^T \end{array} \right] \right) P_n^T \tag{D.20}$$

for any ω_1 and ω_2 satisfying $\omega_1 \omega_2 = \zeta$. Therefore, combining (D.20) with (D.19), we arrive at (4.10), completing the proof of the theorem.

Acknowledgment

The work is supported by the Australian Research Council and Tarong Energy Ltd. Linkage Grant C 0010 6980.

References

- [1] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Chichester, 1994.
- [2] M. J. Corless and A. E. Frazho, *Linear Systems and Control: An Operator Perspective*, Marcel Dekker, New York, 2003.
- [3] N. C. Giri, *Multivariate Statistical Analysis*, Marcel Dekker, New York, 1996.
- [4] T. Kailath, *Linear Systems*, Prentice-Hall Information and System Sciences Series, Prentice-Hall, New Jersey, 1980.
- [5] J. F. Monahan, *Fully Bayesian analysis of ARMA time series models*, J. Econometrics **21** (1983), no. 3, 307–331.
- [6] J. C. Naylor and J. M. Marriott, *A Bayesian analysis of non-stationary AR series*, Bayesian Statistics, 5 (Alicante, 1994) (J. M. Bernardo, J. O. Berger, and A. P. Dawid, eds.), Oxford Sci. Publ., Oxford University Press, New York, 1996, pp. 705–712.
- [7] A. N. Shiryaev, *Essentials of Stochastic Finance: Facts, Models, Theory*, Advanced Series on Statistical Science & Applied Probability, vol. 3, World Scientific, New Jersey, 1999.
- [8] R. E. Skelton, T. Iwasaki, and K. M. Grigoriadis, *A Unified Algebraic Approach to Linear Control Design*, The Taylor & Francis Systems and Control Book Series, Taylor & Francis, London, 1998.
- [9] H. B. Thompson and I. Vladimirov, *Bayesian parameter estimation and prediction in mean reverting stochastic diffusion models*, Nonlinear Analysis. Theory, Methods and Applications **63** (2005), no. 5-7, e2367–e2375.
- [10] C. F. Van Loan, *Computing integrals involving the matrix exponential*, IEEE Trans. Automat. Control **23** (1978), no. 3, 395–404.
- [11] I. Vladimirov and H. B. Thompson, *Algebraic moments of Gaussian random matrices via graphs generated by translated bipartitions*, 2004, submitted.
- [12] ———, *Bayesian forecasting in univariate autoregressive models with normal-gamma prior distribution of unknown parameters*, 2004, submitted.
- [13] S. S. Wilks, *Mathematical Statistics*, A Wiley Publication in Mathematical Statistics, John Wiley & Sons, New York, 1962.

Igor Vladimirov: Department of Mathematics, School of Physical Sciences, Faculty of Engineering, Physical Sciences, and Architecture, the University of Queensland, Brisbane, QLD 4072, Australia
E-mail address: igv@maths.uq.edu.au

Bevan Thompson: Department of Mathematics, School of Physical Sciences, Faculty of Engineering, Physical Sciences, and Architecture, the University of Queensland, Brisbane, QLD 4072, Australia
E-mail address: hbt@maths.uq.edu.au