On the unique representation of non-Gaussian multivariate linear processes

(SHORT RUNNING TITLE: UNIQUE MOVING-AVERAGE REPRESENTATION) *†

Kung-Sik Chan ‡& Lop-Hing Ho
Department of Statistics & Actuarial Science,
University of Iowa, Iowa City, IA 52242
and Department of Mathematics & Statistics,
Wichita State University, Wichita, KS 67260

In contrast to the fact that Gaussian linear processes generally have non-unique moving-average representations, non-Gaussian univariate linear processes have been shown to admit essentially unique moving-average representation, under various regularity conditions. We extend the one-dimensional result to multivariate processes. Under various conditions on the inter-component dependence structure of the error process, we prove that for non-Gaussian multivariate linear processes the moving-average representation is essentially unique under the condition that the transfer function is of full-rank, plus other mild conditions.

*AMS 2000 subject classifications. Primary 62M10; Secondary 60G18

†Key words and phrases. Abelian group, Fourier transform, Random fields, Stationarity, Time series

‡Partially supported by the National Science Foundation (DMS-0405267) and the University of Iowa (Career Development Award).

1 Introduction

Let $X(t)$ be a p-dimensional stationary linear process defined on a countable Abelian group H, i.e.

$$X(t) = W(t) * U(t) = \sum_{k \in H} W(k) U(t-k), \tag{1}$$

where for all $t, k \in H, t-k = t + (-k)$, where $+$ denotes the commutative group operation and $-k$ the inverse of k, $U(t) = (U_1(t), \cdots, U_q(t))^T$ (the superscript T denotes transposition) is an iid q-dimensional random series that is of zero mean and has finite, positive-definite covariance matrix. (A better way is to write $W * U(t)$ instead of the notation $W(t) * U(t)$, which, however, we follow here as it is more commonly adopted in the statistical literature.) We shall adopt the natural condition (i) that the coefficient sequence $W(t), t \in H$, is assumed to be square-summable, i.e., $\sum_{t \in H} \|W(t)\|^2 < \infty$, where $\|\cdot\|$ denotes the L^2-norm. The random vectors X's generally represent some observed data whereas the random vectors U's are referred to as stochastic errors or noises postulated to generate the observed data according to the specification (1). Model (1) includes many interesting stochastic processes. For example, it includes time series where the index set is Z, the set of integers equipped with the usual addition as the commutative group operation. Other significant examples include random fields where the index set is Z^2, the two-dimensional integer lattice equipped with vector addition as the commutative group operation.

A fundamental question is whether such a representation is essentially unique. The uniqueness question is an important question in the statistical inference of stochastic process; see, e.g., Donoho (1981), Lii and Rosenblatt (1982) and Rosenblatt (1985). In the one-dimensional case, it is well-known that for Gaussian X, the moving-average representation is generally non-unique (see, e.g., Lii and Rosenblatt, 1982 and Findley, 1986.) But, interestingly, for non-Gaussian 1-dimensional X, Findley (1986, 1990) and Cheng (1992) have established uniqueness under various conditions. Cheng (1992) proved uniqueness under the minimal condition that

$$\widehat{w}(\gamma) \neq 0, \qquad d\gamma \text{ (a.e.)}$$

where, in the case $H = Z$, $\widehat{w}(\lambda) = \sum_m W_m \exp(im\lambda)$ is the Fourier transform of $w(t)$, also known as the transfer function in time series; see Definition 1 below for the more general definition.

We consider the case that U and X are multidimensional, non-Gaussian processes. Under various conditions on the inter-component dependence structure of $U(t)$, we will give an affirmative answer to the uniqueness question under the additional mild regularity conditions that (ii) the transfer function is of full-rank and (iii) the dimension of X is not smaller than that of U. Indeed, regularity conditions (i)-(iii) guarantee that equation (1) can be inverted to obtain $U(t) = G(t) * X(t)$, for some square-summable G, which we will refer to as the (generalized) invertibility of the model. The (generalized) invertibility condition differs from the ordinary invertibility condition for time series (see Definition 3.1.4 in Brockwell and Davis, 1991) in that the latter condition requires that G be one-sided, i.e. $G(k) = 0$ for negative k so that the noise terms $U(t)$ can be recovered from current and past data. Henceforth, by invertibility we mean generalized invertibility. It is clear that condition (iii) is generally required for the invertibility of model (1) to hold. We shall furthermore show that conditions (i)-(iii) imply that the dimension of U occurring in a moving-average representation of X is unique.

This paper is organized as follows. We give a short discussion of Fourier analysis on groups in Section 2. In Section 3, we prove the invertibility of the model under conditions (i)-(iii). The uniqueness of the moving-average representation is discussed in Section 4, under the additional condition that $U(t)$ consists of iid components. The iid-component assumption is relaxed to independent but non-identically distributed components in Section 5, at the expense of requiring that there exists an integer $r \geq 3$ such that the rth cumulant of each component of $U(t)$ is non-zero. See Jammalamadaka, Rao and Terdik (2004) for the definition and properties of cumulants of multivariate processes. Then in section 6, we consider the case of dependent components of $U(t)$. We impose the dependence condition that a certain matrix of rth order cumulants is non-singular, where $r \geq 3$, and that any two non-zero linear combinations of $U(t)$ are stochastically dependent. These inter-component dependence conditions are rather mild as, e.g., they are valid for multivariate t-distributions. Under these inter-component dependence conditions and conditions (i)-(iii), we show that the moving-average representation is unique up to some shift in the index and an invertible multiplicative factor. It is pertinent to note that if the components of $U(t)$ are independent, the moving-average representation is, however, unique up to component-specific shifts in the index and component-specific non-zero multiplicative factors.

2 Fourier Analysis on Groups

Here we give a brief introduction to Fourier analysis on groups. We refer the reader to Rudin (1962) for details.

Let G be a locally compact Abelian group, on which there exists a non-negative regular measure m, the Haar measure on G. This measure is non-degenerate, i.e. not identically 0, translation invariant, and it is unique up to a multiplicative positive constant. If f and g are measurable functions on G, then the convolution $f * g$ is

$$(f * g)(x) = \int_G f(x-y)\, g(y)\, dy,$$

defined when the above integral is absolutely integrable. Here dy is the Haar measure on G.

A character on G is a continuous homomorphism of G into the multiplicative group of complex numbers of modulus 1. We denote $\widehat{G}$ the set of all characters on G. It is an Abelian group under pointwise multiplication. i.e. $(\gamma_1\gamma_2)(g) = \gamma_1(g)\,\gamma_2(g)$ for $\gamma_1, \gamma_2 \in \widehat{G}, g \in G$. It is customary to write $(g, \gamma) = \gamma(g)$. The topology on $\widehat{G}$ is induced by the topology of uniform convergence on compact subsets of G. Hence a neighborhood basis of 0 is given by sets of the form $\{\gamma : |(g,\gamma) - 1| < \varepsilon \text{ for all } g \in K\}$ where K is a compact set of G and $\varepsilon > 0$. $\widehat{G}$ is called the dual group of G. It is also a locally compact Abelian group, and hence endowed with a Haar measure. The Pontryagin duality theorem states that the dual group of $\widehat{G}$ is G. If G is a discrete group, then it can be shown that $\widehat{G}$ is compact. Also, if G is compact, then $\widehat{G}$ is discrete. The Fourier transform of $f \in L^1(G)$ is defined by

$$\widehat{f}(\gamma) = \int_G f(x)\,(-x, \gamma)\, dx.$$

Note that if G is a discrete group, then the Haar measure of any single point is assigned to be 1. In this case the Fourier transform is

$$\widehat{f}(\gamma) = \sum_{x \in G} f(x)\,(-x, \gamma).$$

The Fourier transform of $f * g$ is $\widehat{f}\widehat{g}$. i.e.

$$\widehat{f * g}(\gamma) = \widehat{f}(\gamma)\,\widehat{g}(\gamma).$$

The Haar measure on $\widehat{G}$ can be normalized so that the following inversion formula is valid for certain class of functions f. (It includes functions f which are continuous, with $\widehat{f} \in L^1(G)$.)

$$f(x) = \int_{\widehat{G}} \widehat{f}(\gamma)(x, \gamma)\, d\gamma, \quad x \in G.$$

Examples:

1. The additive group $\mathbf{R}$ of the real numbers, with the natural topology of the real line. The dual group is $\mathbf{R}$ itself. If we equip $G = \mathbf{R}$ with the Haar measure $\frac{1}{2\pi}dx$ where dx is the Lebesgue measure (the factor $\frac{1}{2\pi}$ is the normalized factor to make the inversion formula true), and the Lebesgue measure dy on $\widehat{G} = \mathbf{R}$, then the Fourier transform is

$$\widehat{f}(y) = \frac{1}{2\pi} \int_{\mathbf{R}} f(x)\, e^{-iyx} dx,$$

and the Fourier inversion formula is

$$f(x) = \int_{\mathbf{R}} \widehat{f}(y)\, e^{ixy} dy.$$

2. The additive group $\mathbf{Z}$ of integers. The dual group is T, the additive group of real numbers modulo 2π, endowed with the natural topology. The Haar measure on $\mathbf{Z}$ is the counting measure, and the Fourier transform is

$$\widehat{f}(\lambda) = \sum_{n=-\infty}^{\infty} f(n)\, e^{-in\lambda}$$

for $\lambda \in [0, 2\pi]$. The Fourier inversion is

$$f(n) = \frac{1}{2\pi} \int_0^{2\pi} \widehat{f}(\lambda)\, e^{in\lambda} d\lambda.$$

3. The case of $\mathbf{Z}^2$ is subsumed by the general case of $G_1 \oplus G_2$, the direct sum of two locally compact Abelian groups G_1 and G_2. The topology on $G_1 \oplus G_2$ is the product topology and the Haar measure is the product measure of the Haar measures on G_1 and G_2. The dual group is identified with $\widehat{G}_1 \oplus \widehat{G}_2$ through

$$((x_1, x_2), (\gamma_1, \gamma_2)) = (x_1, \gamma_1)(x_2, \gamma_2)$$

for $x_1 \in G_1, x_2 \in G_2, \gamma_1 \in \widehat{G}_1, \gamma_2 \in \widehat{G}_2$. The Fourier transform is

$$\widehat{f}\left(\gamma_1, \gamma_2\right) = \int_{G_1 \oplus G_2} f\left(x_1, x_2\right)\left(\left(-x_1, -x_2\right), \left(\gamma_1, \gamma_2\right)\right) dx_1 dx_2.$$

The Fourier inversion is

$$f\left(x_1, x_2\right) = \int_{\widehat{G}_1 \oplus \widehat{G}_2} \widehat{f}\left(\gamma_1, \gamma_2\right)\left(\left(x_1, x_2\right), \left(\gamma_1, \gamma_2\right)\right) d\gamma_1 d\gamma_2.$$

Definition 1. *Let H be a countable Abelian group, and $\widehat{H}$ the dual group of H. Let $W(t)$ be $p \times q$ constant matrices defined on $t \in H$. The $p \times q$ matrix transfer function is defined by*

$$\widehat{w}(\lambda) = \sum_{t \in H} W(t)(-t, \lambda)$$

for $\lambda \in \widehat{H}$.

We note that $W(t), t \in H$ can be recovered from its transfer function $\widehat{w}$ by the Fourier inversion formula, under suitable regularity conditions.

3 Invertibility

The following lemma states the invertibility of model (1) and the uniqueness of the dimension of error process U in the moving-average representation under some mild regularity conditions.

Lemma 2. *Let $X(t) = W(t) * U(t)$, with the dimension of X being p and that of U being q. Assume that (i) the coefficient sequence W is square-summable, (ii) the $p \times q$ matrix transfer function $\widehat{w}(\lambda)$ is of full-rank for almost all $\lambda \in \widehat{H}$ (w.r.t. the Haar measure), and (iii) $q \leq p$. Then there exists a square-summable sequence $G(t), t \in H$ such that $U(t) = G(t) * X(t)$. Furthermore, the dimension q is uniquely determined.*

Proof. Let $\mathcal{U}$ denote the Hilbert space consisting of series of the form $\sum_{j \in H} D(j)U(j)$ where D's are $q \times q$ real-valued matrices, and $\sum_j \|D(j)\|^2 < \infty$. For any $V, W \in \mathcal{U}$, the inner product $< U, V >= E(V^T W)$. Let $\mathcal{X}$ be the space spanned by $K(t)X(t), t \in H$ where K's are arbitrary real-valued $q \times p$ matrices. Clearly, $\mathcal{X} \subseteq \mathcal{U}$. We prove the existence of G such that $U = G * X$ by showing that $\mathcal{U} = \mathcal{X}$. It suffices to show that any $U(t)$ belongs to $\mathcal{X}$. Consider an $U(t)$ and subtract from it its projection on $\mathcal{X}$. Write the difference as $\sum_{j \in H} D(j)U(j)$ which is orthogonal to all $K(t)X(t), t \in Z$ with $K(t)$ being arbitrary. Hence we have, for any $s \in H$, $\sum D(j)E(U(j)X(s)^T) = 0$, implying that $\sum D(j)\Lambda W^T(s-j) = 0$ where Λ is the covariance matrix of $U(t)$. Note that Λ is positive definite. It is then readily seen that, for any character λ, $\widehat{D * W^T}(\lambda) = \sum_{s \in H}(-s, \lambda) \sum D(j)\Lambda W^T(s-j) = 0$, where we abuse the notation D to stand for the sequence $(D(j)\Lambda, j \in H)$. That is, $\widehat{D}(\lambda)\widehat{w^T}(\lambda) = 0$, by the convolution theorem. Because $\widehat{w^T}(\lambda)$ is of full rank almost everywhere, $\widehat{D}(\lambda) = 0$ almost everywhere, and hence $D = 0$ so that $U(t)$ belongs to $\mathcal{X}$. This completes the proof that $\mathcal{U} = \mathcal{X}$. We now prove the uniqueness of q, the dimension of the noise terms. Let $F_X()$ and $F_U()$ be the spectral distribution functions of X and U respectively, and $F_{UX}()$ be the co-spectrum of U and X. Because $X(t) = W(t) * U(t)$, and $U(t) = G(t) * X(t)$, it follows that (c.f. Theorem 9, p.58 of Hannan, 1970, which can be extended to the current setting)

$$F_X(\lambda) = \widehat{w}(\lambda)F_U(\lambda)\widehat{w}^*(\lambda) \tag{2}$$

$$F_{UX}(\lambda) = F_U(\lambda)\widehat{w}^*(\lambda), \tag{3}$$

$$F_U(\lambda) = \widehat{g}(\lambda)F_X(\lambda)\widehat{g}^*(\lambda). \tag{4}$$

$$F_{XU}(\lambda) = F_X(\lambda)\widehat{g}^*(\lambda). \tag{5}$$

where the superscript $*$ denotes transposition combined with complex conjugation, and $\widehat{w}(\lambda)$ is the Fourier transform of W, etc. Suppose that $X(t) = W(t) * U(t)$ and $X(t) = W'(t) * U'(t)$, where the dimension of $U(t)$ is q and that of U' is q'. Moreover, the transfer functions are assumed to be of full rank. But then an analogue to (4) for the second moving-average representation of X yields that $F_{U'}(\lambda) = \widehat{g'}(\lambda)F_X(\lambda)(\widehat{g'})^*(\lambda) = \widehat{g'}(\lambda)\widehat{w}(\lambda)F_U(\lambda)\widehat{w}^*(\lambda)(\widehat{g'})^*(\lambda)$, implying that the rank of $F_{U'}(\lambda)$ is not greater than that of $F_U(\lambda)$. By symmetry, these two ranks must be equal so that $q = \text{rank}\{F_U(\lambda)\} = \text{rank}\{F_{U'}(\lambda)\} = q'$. Note that the assumption that $U(t)$ has a finite, positive-definite covariance matrix implies the equalities $q = \text{rank}\{F_U(\lambda)$ and $\text{rank}\{F_{U'}(\lambda)\} = q'$. Recall the iid nature of U and U' implies that their spectral distribution functions are constant. □

4 Uniqueness

Recall that $X(t)$ and $U(t)$ are $p-$ and $q-$dimensional respectively, and the matrices $W(t)$ are $p \times q$. For a matrix $W(t), W_{ij}(t)$ represents the (i,j) entry, $W_i(t)$ the ith row, and $W^j(t)$ the jth column of $W(t)$. Similarly, $U_i(t)$ is the ith component of $U(t)$. The moving-average representation (1) is not unique as alternative representations can be obtained by (a) permuting the components of the noise terms, (b) shifting the components of the U's systematically, (c) multiplying the components by a scaler and dividing the corresponding coefficients by the same constant, and (d) combinations of these procedures. We shall show that under conditions (i)-(iii) and the non-Gaussianity of X the moving-average representation defined by (1) is unique up to variations induced by (a)-(d).

THEOREM 3. *Let*

$$X(t) = W(t) * U(t) = W'(t) * U'(t) \qquad t \in H$$

where all components of $U(t), U'(t)$ are i.i.d., H is a countable Abelian group, and conditions (i)-(iii) in Lemma 2 hold. If some component of $X(t)$ is non-Gaussian, then there exist a permutation π of the set of integers $\{1, 2, \cdots, q\}$, non-zero scalers β_i and integers $m(i)$ for $i = 1, \cdots, q$ such that for all t

$$U_i'(t) = U_{\pi(i)}(t + m(i))/\beta_i, \tag{6}$$

$$W'^{,i}(t) = \beta_i W^{\pi(i)}(t - m(i)). \tag{7}$$

Proof. Let

$$X(t) = \sum_k W(k) U(t-k) = \sum_k W'(k) U'(t-k)$$

It follows from Lemma 1 that model (1) is invertible, so we can express

$$U(t) = \sum_k C(k) U'(t-k) \tag{8}$$

and

$$U'(t) = \sum_k D(k) U(t-k) \tag{9}$$

where $C(k)$ and $D(k)$ are $q \times q$ square-summable matrices. We can construct independent $q-$dimensional random vectors $Z(s,t), s,t \in H$ on a certain probability space with all components $Z_m(s,t), 1 \le m \le q$ i.i.d. and have the same probability distribution as $U_i(s-t)$. Now define $q-$dimensional random vectors $Y(t), t \in H$ so that the ith component is

$$Y_i(t) = \sum_{1 \le m \le q, k \in H} D_{im}(k) Z_m(t,k).$$

Then for any $1 \le i \le q, t \in H, Y_i(t)$ have the same distribution as the components of $U_i'(t)$. Define $V(t)$ so that the ith component is

$$V_i(t) = \sum_{1 \le l \le q, n} C_{il}(n) Y_l(t-n) = \sum_{1 \le l,m \le q,n,k} C_{il}(n) D_{lm}(k) Z_m(t-n,k), \tag{10}$$

where the indices n and k range over the group H. Then the components $V_i(t)$ have the same distribution as the components of Z. Hence, $V(t)$ and $U(t)$ have identical distribution. $X(t)$ having a non-Gaussian component implies that $U_i(t)$ is non-Gaussian, and consequently $Z_m(s,t), V_i(t)$ are all non-Gaussian. Consider $E\left(V_i(t)^2\right)$ on the two sides of (10), we obtain

$$\sum_{n,k} \left\|\left(C(n) D(k)\right)_i\right\|^2 = 1$$

where the subscript means the ith row of the matrix. Consequently the dot products of the ith row of $C(n)$ and the jth column of $D(k)$ satisfy

$$\sum_{1 \le j \le q,n,k} \left(C(n) D(k)\right)_{ij}^2 = \sum_{1 \le j \le q,n,k} \left(C_i(n) \cdot D^j(k)\right)^2 = 1. \tag{11}$$

Since $V_i(t)$ and $Z_m(s,t)$ are non-Gaussian and identically distributed, in view of (11) and (10) we can apply Theorem 5.6.1 of Kagan *et. al.* (1973) and conclude that for each $1 \le i \le q$ there exist integers $n(i), k(i), 1 \le j(i) \le q$ such that

$$C_i(n(i)) \cdot D^{j(i)}(k(i)) = 1 \tag{12}$$

and

$$C_i(n) \cdot D^j(k) = 0 \tag{13}$$

for $(n,k,j) \neq (n(i), k(i), j(i))$. Hence, equation (10) becomes

$$V_i(t) = Z_{j(i)}(t - n(i), k(i)).$$

Because for fixed t, $V_i(t), 1 \leq i \leq q$, are independent, the triples $(j(i), n(i), k(i))$ are distinct for different i. Let $\mathcal{C}$ be the $q \times q$ matrix whose ith row equals $C_i(n(i))$ and $\mathcal{D}$ the $q \times q$ matrix whose ith column equals $D^{j(i)}(k(i))$. It can be readily checked by (12-13) that $\mathcal{CD}$ equals the $q \times q$ identity matrix. In particular, $\mathcal{D}$ is invertible and (13) implies that $C_i(t) = 0$ unless $t = n(i)$ for $1 \leq i \leq q$. Equation (8) entails that $U_i(t + n(i)) = C_i(n(i))U'(t), 1 \leq i \leq q$. It follows from the independence of $U_i(t + n(i)), 1 \leq i \leq q$, Theorem 3.1.1 of Kagan *et al.* (1973) and the non-Gaussianity of U's that $C_i(n(i))$ are orthogonal vectors each of which has only one non-zero element. Consequently, there exist a permutation γ of $\{1, 2, \cdots, q\}$, and constants α_i such that $U_i(t) = \alpha_i U'_{\gamma(i)}(t - n(i))$. Clearly, α_i are non-zero as the U's are non-degenerate. Let π be the inverse permutation of γ. We have

$$\begin{aligned}
X(t) &= \sum_k W(k)U(t-k) \\
&= \sum_{k, 1 \leq i \leq q} W^i(k)U_i(t-k) \\
&= \sum_{k,i} W^i(k)\alpha_i U'_{\gamma(i)}(t - n(i) - k) \\
&= \sum_{k,i} W^i(k - n(i))\alpha_i U'_{\gamma(i)}(t-k) \\
&= \sum_{k,i} W^{\pi(i)}(k - n(\pi(i)))\alpha_{\pi(i)} U'_i(t-k),
\end{aligned}$$

from which (7) follows, where $\beta_i = \alpha_{\pi(i)}$ and $m(i) = n(\pi(i))$. Similarly obtained is (6). $\square$

We note that $m(i)$ in the preceding theorem need not be identical, as demonstrated by the following example.

Example

Let

$$X(t) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} U(t) + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} U(t+1) + \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} U(t+2).$$

Then clearly for $U'_1(t) = U_1(t-1)$ and $U'_2(t) = U_2(t+1)$, (i.e. $m(1) = -1$

and $m(2)=1$) we have

$$X(t)=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} U'(t)+\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} U'(t+1)+\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} U'(t+2).$$

It is easily seen that

$$U(t)=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} X(t-1)+\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} X(t)+\begin{pmatrix} -\frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} X(t+1)+.....$$

5 Extension to Non-identically-distributed Components

The assumption that the components of $U(t)$ are iid is quite strong. Here, we extend the unique-representation result to the case of independent but possibly non-identically-distributed components, by adapting the cumulant techniques in Findley (1986, 1990). Suppose that X is non-Gaussian, has finite rth moments where r is an integer greater than two, and that the rth cumulants of $X(t)$ are not identically zero. Then, so is the rth cumulant of some component of $U(t)$ by the independent-component assumption, although it is not *a priori* true that each component of $U(t)$ must have non-zero rth cumulant. If some component of $U(t)$ is Gaussian, the moving-average representation will not be unique. Thus, we need to impose in the following theorem the stronger condition that all components of $U(t)$ have a non-zero rth cumulant.

THEOREM 4. *Let*

$$X(t)=W(t)*U(t)=W'(t)*U'(t) \qquad t\in H$$

where $U(t)$ has independent components and so has $U'(t)$, and H is a countable Abelian group. Assume conditions (i)-(iii) in Lemma 2 hold. Assume further that there is an even integer m such that $U(t)$ and $U'(t)$ have finite moments up to order m, and that there is an integer $3\le r\le m$ such that the rth cumulants of each component of $U(t)$ are non-zero. Then, the conclusions of Theorem 3 are valid, i.e. equations (6) and (7) hold.

Proof. For ease of exposition, we give the proof for the case when $H=\mathbf{Z}$, as the proof for the general case is similar. It follows from the first part of the proof of Theorem 3 that

$$U'(t) = \sum_j D(j)\, U(t-j),$$

where $D(j)$ is an $q \times q$ matrix for each j, and $U(t), U'(t)$ are $q-$vectors. Note that it can be seen from the proof of Lemma 2 that $\widehat{D}(\lambda) = \sum_j D(j)\, e^{-i\lambda j}$ is non-singular for each $0 \le \lambda \le 2\pi$, which holds because $\widehat{D}(\lambda) = \widehat{g'}(\lambda)\widehat{w}(\lambda)$, a product of two full-rank matrices. That $\widehat{g'}(\lambda)$ is of full-rank can be argued as follows. Equation (2) implies that $F_X(\lambda)$ is of rank q for all λ, so that (4) implies that $\widehat{g}(\lambda)$ is of full-rank, hence so is $\widehat{g'}(\lambda)$. Let $\ell \neq 0$, then by independence of $U'(t)$ for any m (can take the value 0 or ℓ), the cumulant

$$\operatorname{cum}\left(U'(t), U'(t+\ell), ..., U'(t+\ell), U'(t+m)\right) = 0$$

where $U'(t+\ell)$ is repeated $r-2$ times. By Lemma 5 below we can expand in terms of U and we get

$$\sum_{j_1,..,j_r} \operatorname{cum}\left(D(j_1)\, U(t-j_1), D(j_2)\, U(t+\ell-j_2), ..., D(j_{r-1})\, U(t+\ell-j_{r-1}), D(j_r)\, U(t+m-j_r)\right) = 0$$

Hence

$$\sum_{j_1,\cdots,j_r} \operatorname{cum}\left(\begin{array}{c} (D_1(j_1)\cdot U(t-j_1), ..., D_q(j_1)\cdot U(t-j_1)), (D_1(j_2)\cdot U(t+\ell-j_2), ..., D_q(j_2)\cdot U(t+\ell-j_2), \\ ..., (D_1(j_{r-1})\cdot U(t+\ell-j_{r-1}), ..., D_q(j_{r-1})\cdot U(t+\ell-j_{r-1}), (D_1(j_r)\cdot U(t+m-j_r), ..., D_q(j_r)\, U(t+m-j_r)) \end{array}\right) = 0$$

According to Jammalamadaka, Rao and Terdik (p.7, 2004) there are n^r components in the above cumulant, each one is of the form

$$\sum_{j_1,..,j_r} \operatorname{cum}\left(D_{i_1}(j_1)\cdot U(t-j_1), D_{i_2}(j_2)\cdot U(t+\ell-j_2), ..., D_{i_{r-1}}(j_{r-1})\cdot U(t+\ell-j_{r-1}), D_{i_r}(j_r)\cdot U(t+m-j_r)\right) = 0$$

for some $1 \le i_1, i_2, ...i_r \le q$, where $\cdot$ denotes the inner product. Since $U(i)'s$ are independent for different i, any non-zero term in the above sum must have $t-j_1, t+\ell-j_2, ..., t+\ell-j_{r-1}, t+m-j_r$ all agree, hence it is reduced to

$$\sum_j \operatorname{cum}\left(D_{i_1}(j)\cdot U(t-j), D_{i_2}(j+\ell)\cdot U(t-j), ..., D_{i_{r-1}}(j+\ell)\cdot U(t-j), D_{i_r}(j+m)\cdot U(t-j)\right) = 0.$$

The above cumulant consists of linear combination of terms $\operatorname{cum}\left(U_{k_1}(t), U_{k_2}(t), ..., U_{k_r}(t)\right)$ for $1 \le k_1, k_2, \cdots, k_r \le q$. Because of the independent-component assumption, the only non-zero terms are $c_i = \operatorname{cum}\left(U_i(t), U_i(t), ..., U_i(t)\right)$ where $i = 1, 2, ..., q$. Hence for each $1 \le i_1, i_2, ..., i_r \le q$ we get

$$\sum_{j, 1\le k\le q} D_{i_1 k}(j)\, D_{i_2 k}(j+\ell) ... D_{i_{r-1} k}(j+\ell)\, D_{i_r k}(j+m)\, c_k = 0. \tag{14}$$

Let's denote

$$f_{i_1,i_2,...,i_{r-1}}(j) = \left(D_{i_11}(j)\, D_{i_21}(j+\ell) ... D_{i_{r-1}1}(j+\ell), ..., D_{i_1q}(j)\, D_{i_2q}(j+\ell) ... D_{i_{r-1}q}(j+\ell)\right)^T$$

and

$$h_{i_r}(j) = \left(D_{i_r1}(-j)\, c_1, D_{i_r2}(-j)\, c_2, ..., D_{i_rn}(-j)\, c_n\right)^T.$$

Define

$$(f * h)(t) = \sum_j f(j) \cdot h(t-j).$$

Equation (14) is equivalent to

$$\left(f_{i_1,i_2,...,i_{r-1}} * h_{i_r}\right)(-m) = 0.$$

Let

$$\widehat{f}(\lambda) = \sum_j f(j)\, e^{-i\lambda j}$$

which is again a q-vector, and it is readily seen that

$$\widehat{(f * g)}(\lambda) = \widehat{f}(\lambda) \cdot \widehat{g}(\lambda).$$

Hence $\left(f_{i_1,i_2,...,i_{r-1}} * h_{i_r}\right)(-m) = 0$ for all m implies that the product $\widehat{f}_{i_1,i_2,...,i_{r-1}}(\lambda) \cdot \widehat{h}_{i_r}(\lambda) = 0$ for all λ. This is true for all $1 \leq i_r \leq q$. Let $\mathrm{diag}(c_1, c_2, \cdots, c_q)$ denote a diagonal matrix with c's being its diagonal elements. Now for each λ the matrix $\mathrm{diag}(1/c_1, 1/c_2, \cdots, 1/c_q)\left(\widehat{h}_1(\lambda), \widehat{h}_2(\lambda), ..., \widehat{h}_q(\lambda)\right)$ is non-singular as it equals the transpose of $\widehat{D}(-\lambda)$, implying that

$$\widehat{f}_{i_1,i_2,...,i_{r-1}}(\lambda) = 0 \text{ for all } \lambda.$$

Hence $f_{i_1,i_2,...,i_{r-1}}(j) = 0$ for all j, i.e. for each $1 \leq k \leq q$ and for all $1 \leq i_1, ..., i_{r-1} \leq q, l \neq 0$

$$D_{i_1k}(j)\, D_{i_2k}(j+\ell) ... D_{i_{r-1}k}(j+\ell) = 0. \tag{15}$$

The non-singularity of $\widehat{D}(\lambda)$ entails that for each i there exists j_i such that the column $D^i(j_i)$ is not a zero vector. By (15) it follows that $D^i(k) = 0$ for any $k \neq j_i$. Now the representation of U' is

$$U_i'(t) = \sum_k D_{ik}(j_k)\, U_k(t - j_k).$$

Thanks to Theorem 3.1.1 of Kagan *et al.* (1973), the independence of U_i', and the non-Gaussianity and independence of U_k's imply that there is only one non-zero D_{ik} for each k. The conclusion of the theorem can then be obtained by adapting the arguments at the end of Theorem 3. □

The following lemma is used to justify the convergence of an infinite sum of cumulant terms.

Lemma 5. *Let $U(t)$ and $U'(t)$ satisfy the same properties as in the previous theorem. Then*

$$\begin{aligned} & cum\left(\sum_k D(k)\,U(t-k)\,, U'(s_1)\,, ..., U'(s_{r-1})\right) \\ = & \sum_k D(k)\,cum\left(U(t-k)\,, U'(s_1)\,, ..., U'(s_{r-1})\right) \end{aligned}$$

Proof. Let m be an even integer such that $U(t)$ and $U'(t)$ have finite moments up to order m, and $U'(t) = \sum_{k=0}^{\infty} D(k)\,U(t-k)$ where $D(k)$ is square summable. We assume that the sum is one-sided for ease of notations. We will first show that given $\varepsilon > 0$, there exists N such that

$$E\left|\left(\sum_{k\geq N} D(k)\,U(t-k)\right)\right|^p < \varepsilon \tag{16}$$

for all $p = 2, 3, \cdots, m$.

First assume p is even. Using the fact that $U(t)$ are independent for different t and that $E(U_i(t)) = 0$, we obtain $E\left(\sum_{j,N\leq k\leq M} D_{ij}(k)\,U_j(t-k)\right)^p$ equals to a finite sum of terms whose number of terms is independent of N and M and where each term is, up to a multiplicative constant, a product of terms of the form

$$\sum_{1\leq j_i,\cdots,j_q\leq q, N\leq k\leq M} D_{ij_1}(k)\cdots D_{ij_q}(k)\,E\left(U_{j_1}(t-k)\cdots U_{j_{r_1}}(t-k)\right)\cdots E\left(U_{j_{r_s+1}}(t-k)\cdots U_{j_q}(t-k)\right),$$

with at least two terms inside each expectation. To see this, first, let's consider the case that $p = 4$. Then

$$E\left|\left(\sum_{N\leq k\leq M} D_{ij}(k)\, U_j(t-k)\right)\right|^4$$
$$= \sum_{1\leq j_1,\cdots,j_4\leq q, N\leq k_i\leq M} D_{ij_1}(k_1)\, D_{ij_2}(k_2)\, D_{ij_3}(k_3)\, D_{ij_4}(k_4)\, E\left(U_{j_1}(t-k_1)\, U_{j_2}(t-k_2)\, U_{j_3}(t-k_3)\, U_{j_4}(t-k_4)\right)$$
$$= \sum_{1\leq j_1,\cdots,j_4\leq q, N\leq k\leq M} D_{ij_1}(k)\, D_{ij_2}(k)\, D_{ij_3}(k)\, D_{ij_4}(k)\, E\left(U_{j_1}(t-k)\, U_{j_2}(t-k)\, U_{j_3}(t-k)\, U_{j_4}(t-k)\right)$$
$$+ \left(\sum_{1\leq j_1,j_2\leq q, N\leq k_1\leq M} D_{ij_1}(k_1)\, D_{ij_2}(k_1)\, E\left(U_{j_1}(t-k_1)\, U_{j_2}(t-k_1)\right)\right)$$
$$\times \left(\sum_{1\leq j_3,j_4\leq q, N\leq k_2\leq M} D_{ij_3}(k_2)\, D_{ij_4}(k_2)\, E\left(U_{j_3}(t-k_2)\, U_{j_4}(t-k_2)\right)\right)$$
$$- \sum_{1\leq j_1,\cdots,j_4\leq q, N\leq k\leq M} D_{ij_1}(k)\, D_{ij_2}(k)\, D_{ij_3}(k)\, D_{ij_4}(k)\, E\left(U_{j_1}(t-k)\, U_{j_2}(t-k)\right) E\left(U_{j_3}(t-k)\, U_{j_4}(t-k)\right)$$

We need to include the last term because in the product term of the second last line we include terms with $k_1 = k_2$, which are absent in the original expression. The computations for larger p is similar. Since $D(k)$ is square summable and all the expectations are finite and have the same value for all k, we can make each such sum small by choosing N large enough. Hence we obtain (16).

If p is odd, then we can apply the above result to $p-1, p+1$ separately. Write

$$p = \left(\frac{p-1}{2}\right) + \left(\frac{p+1}{2}\right),$$

and apply Holder's inequality we get (16).

To prove the convergence of the infinite sum of cumulants we need to show that given $\varepsilon > 0$, there exists N such that

$$\left|\operatorname{cum}\left(\sum_{k\geq N} D(k)\, U(t-k), U'(s_1), ..., U'(s_{r-1})\right)\right| < \varepsilon. \tag{17}$$

Note that the cumulant for any random variables $Y_1, Y_2, ..., Y_r$ can be expressed in terms of moments:

$$\operatorname{cum}(Y_1, Y_2, ..., Y_r) = \sum (-1)^{\ell-1}(\ell-1)!\left(E\left(\prod_{j\in\nu_1} Y_j\right) ....E\left(\prod_{j\in\nu_\ell} Y_j\right)\right)$$

where the summation extends over all partitions $(\nu_1, ..., \nu_\ell)$, $\ell = 1, 2, ..., r$, of $(1, 2, ..., r)$. Thus to prove (17) it suffices to prove that for any $i = 1, 2, ..., q$,

$$E\left(\left|\left(\sum_{k\geq N} D_{ij}(k)\, U_j(t-k)\right) f_1 f_2 \cdots f_p\right|\right) < \varepsilon$$

where $E\left(|f_i|^\ell\right) < \infty$ for all $i = 1, \cdots, p, \ell = 2, \cdots, m$. Here $p \geqslant 1$ since $E(U_i) = E(U_i') = 0$. Also, $p \leq r-1$. By the generalized Holder's inequality,

$$E\left(\left|\left(\sum_{k\geq N} D(k)\,U(t-k)\right) f_1 f_2 \cdots f_p\right|\right)$$

$$\leq \left(E\left|\left(\sum_{k\geq N} D(k)\,U(t-k)\right)^{p+1}\right|\right)^{1/(1+p)} \left(E\left(|f_1|^{1+p}\right)\right)^{1/(1+p)} \cdots \left(E\left(|f_i|^{1+p}\right)\right)^{1/(1+p)}.$$

By (16) the first term on the right hand side can be made arbitrarily small by choosing N large enough, hence we obtain the desired conclusion. □

6 Dependent Components

The results obtained so far establish that, under mild regularity conditions, the moving-average representation of a non-Gaussian linear process is unique up to some translation in the index and some multiplicative constant, both of which may be component specific. That the index can be shifted differently for different components without affecting the probabilistic structure of the underlying process is tied to the fact that the components of each noise term are independent. Intuitively, if the components of each noise term are jointly dependent, then any component-specific shifts in the index will alter the probabilistic structure of the underlying process. In this section, we consider the case when the (vector) noise terms contain dependent and possibly non-identically distributed components. We impose two conditions on the between-component dependence structure of the noise terms plus other mild regularity conditions to show that a non-Gaussian linear process admits a unique moving-average representation up to a shift in the index (being identical for all components) and a multiplicative invertible matrix. The dependence conditions to be imposed are rather mild as we shall show below that they are satisfied by, e.g., multivariate t-distributions.

Before we introduce the dependence conditions, first consider the following definition for the noise term in some moving-average representation.

Definition 6. *Let $U(t)$ be $q-$dimensional random vectors that are independent and identically distributed. Let $r \geq 3$ be an integer and $K =$*

$(k_3, k_4, ..., k_r)$ *be a multi-index, and* $1 \le k_i \le q$ *for all* $3 \le i \le r$. *Then the matrix* B_K *is the* $q \times q$ *matrix where the* (i,j)*th entry of* B_K *is* $\alpha_{ijK} = cum(U_i, U_j, U_{k_3}, U_{k_4}, ..., U_{k_r})$.

Note that B_K is a zero matrix for any Gaussian distribution. We shall impose the following two conditions on the noise process in some moving-average representation of the observed process:
(D1) The noise process $\{U(t)\}$ admits an invertible B_K for some K with an $r \ge 3$.
(D2) Any two linear combinations of $U(t)$ with non-zero coefficients must be stochastically dependent.

THEOREM 7. *Let*

$$X(t) = W(t) * U(t) = W'(t) * U'(t) \qquad t \in H$$

where $U(t)$ *satisfies condition (D1), for some* $r \ge 3$, *and (D2), and* H *is a countable Abelian group. Assume conditions (i)-(iii) in Lemma 2 hold, and there is an even integer* $m \ge r$ *such that* $U(t)$ *and* $U'(t)$ *have finite moments up to order* m *Then, there exists a* $c \in H$ *and an invertible matrix* Q *such that for all* $t \in H$,

$$U'(t) = Q^{-1}U(t+c), \tag{18}$$

$$W'(t) = W(t-c)Q. \tag{19}$$

Proof. The proof of Theorem 3 implies the validity of the relationship between the two noise processes stated in equation (9). The rest of the proof follows from Theorem 8 and Theorem 12 below. (Note that the conditions of Theorem 8 hold, thanks to (D1) and the arguments in the beginning of the proof of Theorem 4.) □

Example

We now exhibit an example that satisfies conditions (D1) and (D2). Let U be a random variable whose distribution is identical to that of the noise terms $U(t)$. Suppose that $U = \sqrt{\lambda}Z$ where Z has a multivariate normal distribution $N(0, \Sigma)$, where Σ is a positive-definite matrix, and λ is a positive random variable with finite, positive variance. This example includes multivariate t-distributions if the reciprocal of λ has a normalized Gamma distribution. The cumulants can be calculated by using a conditioning argument (Brillinger, 1969) that for any (scalar) random variables $X_1, X_2, \cdots, X_m$ and Y, we have

$$\text{cum}(X_1, X_2, \cdots, X_m) = \sum_{\pi} \text{cum}(\text{cum}(X_{\pi_1}|Y), \cdots, \text{cum}(X_{\pi_b}|Y)),$$

where the sum is over all partitions π of the set of integers $\{1, 2, \cdots, m\}$, X_{π_j} denotes the set of X_k where $k \in \pi_j \subseteq \{1, 2, \cdots, m\}$ with π_j being an element of the partition π. Simplification can be further achieved by using the result that for multivariate normal distributions, cumulants of order 3 or higher vanish.

We now show that any two non-trivial linear combinations of U must be dependent. Consider $\gamma^T U$ and $\beta^T U$ where both γ and β are non-zero. We shall prove that they are stochastically dependent by contradiction. Suppose these two linear combinations of U are independent. Then by the conditional cumulant formula,

$$0 = \mathrm{cum}(\gamma^T U, \beta^T U) = \mathrm{cum}(\mathrm{cum}(\gamma^T U|\lambda), \mathrm{cum}(\beta^T U|\lambda)) + \mathrm{cum}(\mathrm{cum}(\gamma^T U, \beta^T U|\lambda)) = \gamma^T \Sigma \beta E(\lambda).$$

So, $\gamma^T \Sigma \beta = 0$ as $E(\lambda) > 0$. (Here, we have made use of the fact that first-order cumulants are first moments.) But we also have, via conditioning,

$$\begin{aligned} 0 &= \mathrm{cum}(\gamma^T U, \gamma^T U, \beta^T U, \beta^T U) \\ &= \mathrm{cum}(\lambda, \lambda)\{\gamma^T \Sigma \gamma \beta^T \Sigma \beta + 2(\gamma^T \Sigma \beta)^2\} \\ &= \mathrm{var}(\lambda) \gamma^T \Sigma \gamma \beta^T \Sigma \beta, \end{aligned}$$

implying that either γ or β must be zero, deriving a contradiction. Hence, any two non-trivial linear combinations of U are dependent.

Next, we show that $B_{1,1}$ is invertible where we write $B_{1,1}$ for $B_{(1,1)}$. It is clear that $B_{1,1}$ is a symmetric matrix, so it suffices to show that the quadratic form $\gamma^T B_{1,1} \gamma$ is positive for any non-zero vector γ. But $\gamma^T B_{1,1} \gamma = \mathrm{cum}(\gamma^T U, \gamma^T U, e_1^T U, e_1^T U)$ where $e_1 = (1, 0, \cdots, 0)^T$, i.e., its first component is 1 and other components are zero. Using the conditional cumulant formula, we have $\gamma^T B_{1,1} \gamma = \mathrm{var} \lambda \{\gamma^T \Sigma \gamma e_1^T \Sigma e_1 + 2(\gamma^T \Sigma e_1)^2\}$ which is strictly positive for any non-zero γ. Hence, $B_{1,1}$ is invertible. □

We now state two theorems needed in the proof of Theorem 7, which are also of some independent interest. Both theorems concerns the moving-average structure of an iid process. The first result states that under some regularity conditions (mainly D1), a possibly infinite moving-average representation of an iid process is, in fact, a finite moving-average representation. We will further prove in the second result that with the additional condition (D2), any moving-average representation of an iid process is trivial, i.e. the process is identical to the noise process, up to some shift in the index and a multiplicative invertible matrix.

THEOREM 8. *Let*

$$U'(t) = D(t) * U(t) \qquad t \in H$$

where $U(t), U'(t)$ are $q-$dimensional random vectors that are iid respectively and there is an even integer m such that $U(t)$ and $U'(t)$ have finite moments up to order m. Here $D(t)$ is a sequence of square summable matrices, and H is a countable Abelian group. Assume that

1. *$\widehat{D}(\lambda) = \sum_{t \in H} D(t)(-t, \lambda)$ is non-singular for almost all $\lambda \in \widehat{H}$.*
2. *Condition (D1) holds for the process $\{U(t)\}$, for some multi-index K with $m \geq r \geq 3$.*

   Then there exists at most q non-zero matrices $D(t)$ in the above representation.

We need two lemmas for the proof of the theorem.

Lemma 9. *Let $a_{ij}, u_i \in \mathbf{R}^n$, where $1 \leq i \leq n, j \in \mathbf{N}$ such that $\sum_j a_{ij} \to u_i$ in $\boldsymbol{R}^n$. If $u_1, u_2, \cdots, u_n$ are linearly independent, then there exist $j_1, j_2, \cdots, j_n \in \mathbf{N}$ such that $a_{1j_1}, a_{2j_2}, \cdots, a_{nj_n}$ are linearly independent.*

Proof. Since $u_1, u_2, \cdots, u_n$ are linearly independent in $\mathbf{R}^n$, we have

$$\det(u_1, u_2, \cdots, u_n) \neq 0.$$

Hence there exists $M > 0$ such that

$$\det\left(\sum_{j_1=1}^{M} a_{1j_1}, \sum_{j_2=1}^{M} a_{2j_2}, \cdots, \sum_{j_n=1}^{M} a_{nj_n}\right) \neq 0.$$

By the multi-linearity of the determinant we obtain

$$\sum_{1 \leq j_1, j_2, \cdots, j_n \leq M} \det(a_{1j_1}, a_{2j_2}, \cdots, a_{nj_n}) \neq 0.$$

It follows that one of the determinants $\det(a_{1j_1}, a_{2j_2}, \cdots, a_{nj_n}) \neq 0$. □

Lemma 10. *Let $M_1, ..., M_n$ be $m \times m$ matrices, then*

$$\begin{vmatrix} a_{11}M_1 & a_{12}M_2 & .... & a_{1n}M_n \\ a_{21}M_1 & a_{22}M_2 & .... & a_{2n}M_n \\ .... & .... & .... & .... \\ a_{n1}M_1 & a_{n2}M_2 & .... & a_{nn}M_n \end{vmatrix} = (\det A)^m \det M_1 ... \det M_n$$

where $A = \begin{pmatrix} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ .... & .... & .... & .... \\ a_{n1} & a_{n2} & .... & a_{nn} \end{pmatrix}$.

Proof. We will prove by induction on n. It is clearly true for $n = 1$. Assume it is true for n. We may assume $a_{11} \neq 0$, then

$$\begin{vmatrix} a_{11}M_1 & a_{12}M_2 & .... & a_{1n}M_n & a_{1,n+1}M_{n+1} \\ a_{21}M_1 & a_{22}M_2 & .... & a_{2n}M_n & a_{2,n+1}M_{n+1} \\ .... & .... & .... & .... & .... \\ a_{n1}M_1 & a_{n2}M_2 & .... & a_{nn}M_n & a_{n,n+1}M_{n+1} \\ a_{n+1,1}M_1 & a_{n+1,2}M_2 & .... & a_{n+1,n}M_n & a_{n+1,n+1}M_{n+1} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11}M_1 & a_{12}M_2 & .... & a_{1n}M_n & a_{1,n+1}M_{n+1} \\ 0 & \left(-\frac{a_{21}a_{12}}{a_{11}} + a_{22}\right)M_2 & .... & \left(-\frac{a_{21}a_{1n}}{a_{11}} + a_{2n}\right)M_n & \left(-\frac{a_{21}a_{1,n+1}}{a_{11}} + a_{2,n+1}\right)M_{n+1} \\ .... & .... & .... & .... & .... \\ 0 & \left(-\frac{a_{n1}a_{12}}{a_{11}} + a_{n2}\right)M_2 & .... & \left(-\frac{a_{n1}a_{1n}}{a_{11}} + a_{nn}\right)M_n & \left(-\frac{a_{n1}a_{1,n+1}}{a_{11}} + a_{n,n+1}\right)M_{n+1} \\ 0 & \left(-\frac{a_{n+1,1}a_{12}}{a_{11}} + a_{n+1,2}\right)M_2 & .... & \left(-\frac{a_{n+1,1}a_{1n}}{a_{11}} + a_{n+1,n}\right)M_n & \left(-\frac{a_{n+1,1}a_{1,n+1}}{a_{11}} + a_{n+1,n+1}\right)M_{n+1} \end{vmatrix}$$

$$= a_{11}^m \det M_1 \begin{vmatrix} \left(-\frac{a_{21}a_{12}}{a_{11}} + a_{22}\right)M_2 & .... & \left(-\frac{a_{21}a_{1n}}{a_{11}} + a_{2n}\right)M_n & \left(-\frac{a_{21}a_{1,n+1}}{a_{11}} + a_{2,n+1}\right)M_{n+1} \\ .... & .... & .... & .... \\ \left(-\frac{a_{n1}a_{12}}{a_{11}} + a_{n2}\right)M_2 & .... & \left(-\frac{a_{n1}a_{1n}}{a_{11}} + a_{nn}\right)M_n & \left(-\frac{a_{n1}a_{1,n+1}}{a_{11}} + a_{n,n+1}\right)M_{n+1} \\ \left(-\frac{a_{n+1,1}a_{12}}{a_{11}} + a_{n+1,2}\right)M_2 & .... & \left(-\frac{a_{n+1,1}a_{1n}}{a_{11}} + a_{n+1,n}\right)M_n & \left(-\frac{a_{n+1,1}a_{1,n+1}}{a_{11}} + a_{n+1,n+1}\right)M_{n+1} \end{vmatrix}$$

$$= a_{11}^m \det M_1 \det M_2 ... \det M_{n+1} \begin{vmatrix} \left(-\frac{a_{21}a_{12}}{a_{11}} + a_{22}\right) & .... & \left(-\frac{a_{21}a_{1n}}{a_{11}} + a_{2n}\right) & \left(-\frac{a_{21}a_{1,n+1}}{a_{11}} + a_{2,n+1}\right) \\ .... & .... & .... & .... \\ \left(-\frac{a_{n1}a_{12}}{a_{11}} + a_{n2}\right) & .... & \left(-\frac{a_{n1}a_{1n}}{a_{11}} + a_{nn}\right) & \left(-\frac{a_{n1}a_{1,n+1}}{a_{11}} + a_{n,n+1}\right) \\ \left(-\frac{a_{n+1,1}a_{12}}{a_{11}} + a_{n+1,2}\right) & .... & \left(-\frac{a_{n+1,1}a_{1n}}{a_{11}} + a_{n+1,n}\right) & \left(-\frac{a_{n+1,1}a_{1,n+1}}{a_{11}} + a_{n+1,n+1}\right) \end{vmatrix}^m$$

$$= \det M_1 \det M_2 ... \det M_{n+1} \begin{vmatrix} a_{11} & a_{12} & .... & a_{1n} & a_{1,n+1} \\ a_{21} & a_{22} & .... & a_{2n} & a_{2,n+1} \\ .... & .... & .... & .... & ...... \\ a_{n1} & a_{n2} & .... & a_{nn} & a_{n,n+1} \\ a_{n+1,1} & a_{n+1,2} & .... & a_{n+1,n} & a_{n+1,n+1} \end{vmatrix}^m$$

$$= (\det A)^m \det M_1 ... \det M_{n+1}.$$

□

Proof of theorem 8. We will only give the proof when $H = \mathbf{Z}$ here. The proof is similar for the general case of a countable Abelian group H. Since $\widehat{D}(\lambda)$ is non-singular, by lemma 9 there exists n linearly independent column vectors $D^i(t_i), i = 1, 2, ..., n$ where the integers t_i may be repeated. Let $t \neq t_i, 1 \leq i \leq q$, and w arbitrary, by independence of $U(t)$ we have

$$\text{cum}\left(U'(t), U'(s_2), \cdots, U'(s_{r-1}), U'(w)\right) = 0$$

where $s_j \in \{t_1, \cdots, t_q\}, j = 2, \cdots, r-1$ and they may be repeated. Since $U'(t) = D(t) * U(t)$, by lemma 5 we obtain

$$\sum_{j_1, j_2 \ldots,, j_r \in \mathbf{Z}} \text{cum}\left(D(j_1) U(t - j_1), D(j_2) U(s_2 - j_2), ..., D(j_{r-1}) U(s_{r-1} - j_{r-1}), D(j_r) U(w - j_r)\right) = 0.$$

According to Jammalamadaka, Rao and Terdik (2004, p.7) there are q^r components in the each of the above cumulants that are indexed by $(i_1, i_2, ..., i_r)$, where $1 \leq i_k \leq q$ for each $k = 1, ..., r$. Each component equals 0 from above, hence

$$\sum_{j_1, j_2, \ldots, j_r \in \mathbf{Z}} \text{cum}\left(D_{i_1}(j_1) \cdot U(t - j_1), D_{i_2}(j_2) \cdot U(s_2 - j_2), ..., D_{i_r}(j_r) \cdot U(w - j_r)\right) = 0.$$

Since $U(t)' s$ are independent for different t, any non-zero term in the above sum must have $t - j_1, s_2 - j_2, ..., s_{r-1} - j_{r-1}, w - j_r$ all agree, which is then reduced to

$$\sum_{j \in \mathbf{Z}} \text{cum}\left(D_{i_1}(t - j) \cdot U(j), D_{i_2}(s_2 - j) \cdot U(j), ..., D_{i_r}(w - j) \cdot U(j)\right) = 0.$$

Expanding the above summation we get

$$\sum_{j, 1 \leq k_1, k_2, \ldots, k_r \leq q} D_{i_1 k_1}(t - j) D_{i_2 k_2}(s_2 - j) \ldots D_{i_{r-1} k_{r-1}}(s_{r-1} - j) D_{i_r k_r}(w - j) \alpha_{k_1 k_2 \ldots k_r} = 0. \tag{20}$$

where $\alpha_{k_1 k_2 \ldots k_r} = \text{cum}\left(U_{k_1}, U_{k_2}, ..., U_{k_r}\right)$. For each $(i_1, i_2, ..., i_{r-1})$ where $1 \leq i_k \leq q$ for $k = 1, ..., r-1, j \in \mathbf{Z}$, we define a $q-$vector

$$f_{i_1, i_2, \ldots, i_{r-1}}(j) = \begin{pmatrix} \sum_{1 \leq k_1, .., k_{r-1} \leq q} D_{i_1 k_1}(t - j) \ldots D_{i_{r-1} k_{r-1}}(s_{r-1} - j) \alpha_{k_1 k_2 .. k_{r-1} 1}, \ldots\ldots\ldots, \\ \sum_{1 \leq k_1, .., k_{r-1} \leq q} D_{i_1 k_1}(t - j) \ldots D_{i_{r-1} k_{r-1}}(s_{r-1} - j) \alpha_{k_1 k_2 .. k_{r-1} q} \end{pmatrix}^T.$$

For each $1 \leq i_r \leq q, j \in \mathbf{Z}$, define

$$g_{i_r}(j) = \left(D_{i_r 1}(j), ..., D_{i_r q}(j)\right)^T.$$

Also define

$$(f * g)(w) = \sum_j f(j) \cdot g(w - j).$$

It is not hard to see that (20) is equivalent to

$$\left(f_{i_1, i_2, \ldots, i_{r-1}} * g_{i_r}\right)(w) = 0.$$

Let

$$\widehat{f}(\lambda) = \sum_j f(j) e^{-i\lambda j}$$

which is again a q-vector, it is readily seen that

$$\widehat{(f * g)}(\lambda) = \widehat{f}(\lambda) \cdot \widehat{g}(\lambda).$$

Hence $\left(f_{i_1, i_2, \ldots, i_{r-1}} * g_{i_r}\right)(w) = 0$ for all w implies that $\widehat{f}_{i_1, i_2, \ldots, i_{r-1}}(\lambda) \cdot \widehat{g}_{i_r}(\lambda) = 0$ for almost all λ. This is true for all $1 \leq i_r \leq q$. Now for each λ by assumption the matrix $(\widehat{g}_1(\lambda), \widehat{g}_2(\lambda), \ldots, \widehat{g}_q(\lambda)) = \widehat{D}(\lambda)^T$ is non-singular. Therefore

$$\widehat{f}_{i_1, i_2, \ldots, i_{r-1}}(\lambda) = 0 \text{ for almost all } \lambda.$$

Hence $f_{i_1, i_2, \ldots, i_{r-1}}(j) = 0$ for all j, which implies that for each $t \notin \{t_1, \cdots, t_q\}$, $1 \leq i_1, i_2, \cdots, i_{r-1} \leq q$, and $1 \leq k \leq q$,

$$\sum_{1 \leq k_1, k_2, \cdots, k_{r-1} \leq q} D_{i_1 k_1}(t) D_{i_2 k_2}(t_{k_2}) \cdots D_{i_{r-1} k_{r-1}}\left(t_{k_{r-1}}\right) \alpha_{k k_1 k_2 \cdots k_{r-1}} = 0.$$

For each $1 \leq i_1 \leq q$, define a $q^{r-1}-$vector

$$v_{i_1} = (x_{i_1}^T, x_{i_1}^T, \cdots, x_{i_1}^T)^T,$$

where $x_{i1} = (D_{i_1 1}(t), D_{i_1 2}(t), \cdots, D_{i_1 q}(t))^T$ is repeated q^{r-2} times. Also define (writing D_{kl} for $D_{kl}(t_l)$ henceforth)

$$u_{i_2 \cdots i_{r-1} k} = \left(D_{i_2 k_2} \cdots D_{i_{r-1} k_{r-1}} \alpha_{k k_1 k_2 \cdots k_{r-1}}\right)$$

where the row vectors are indexed by $1 \leq i_2, \cdots, i_{r-1}, k \leq q$. There are q^{r-1} components in each vector and indexed by $(k_1, k_2, \cdots, k_{r-1})$. The above summation implies that the inner product

$$v_{i_1} \cdot u_{i_2 \cdots i_{r-1} k} = 0$$

for all $1 \leq i_1, i_2, \cdots, i_{r-1}, k \leq q$. For each $1 \leq i_1 \leq q$, there are q^{r-1} equations, indexed by $i_2, \cdots, i_{r-1}, k$. Each equation has q^{r-1} terms, indexed by $(k_1, k_2, \cdots, k_{r-1})$. Let

$$U = \left(u_{i_2 \cdots i_{r-1} k}\right)$$

which is a $q^{r-1} \times q^{r-1}$ matrix where the rows are indexed by $i_2, \cdots, i_{r-1}, k$, and the columns are indexed by $k_1, k_2, \cdots, k_{r-1}$. Consider a block of q columns, where $k_2, \cdots, k_{r-1}$ remain fixed and $1 \leq k_1 \leq q$. For these columns we consider a block of q rows that $i_2, \cdots, i_{r-1}$ remains fixed and $1 \leq k \leq q$. It is not hard to see that such a $q \times q$ block is of the form

$$D_{i_2 k_2} \cdots D_{i_{r-1} k_{r-1}} \begin{pmatrix} \alpha_{11 k_2 \cdots k_{r-1}} & \alpha_{12 k_2 \cdots k_{r-1}} & \cdots & \alpha_{1n k_2 \cdots k_{r-1}} \\ \alpha_{21 k_2 \cdots k_{r-1}} & \alpha_{22 k_2 \cdots k_{r-1}} & \cdots & \alpha_{2n k_2 \cdots k_{r-1}} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1 k_2 \cdots k_{r-1}} & \alpha_{n2 k_2 \cdots k_{r-1}} & \cdots & \alpha_{nn k_2 \cdots k_{r-1}} \end{pmatrix} = D_{IK} B_K$$

where $I = (i_2, i_3, \cdots, i_{r-1})$, $K = (k_2, k_3, \cdots, k_{r-1})$ are multi-indices, and B_K is a $q \times q$ matrix. Hence

$$U = \begin{pmatrix} D_{I_1 K_1} B_{K_1} & D_{I_1 K_2} B_{K_2} & \cdots & D_{I_1 K_l} B_{K_l} \\ D_{I_2 K_1} B_{K_1} & D_{I_2 K_2} B_{K_2} & \cdots & D_{I_2 K_l} B_{K_l} \\ \cdots & \cdots & \cdots & \cdots \\ D_{I_l K_1} B_{K_1} & D_{I_l K_2} B_{K_2} & \cdots & D_{I_l K_l} B_{K_l} \end{pmatrix}$$

where $l = q^{r-2}, I_j, K_m$ runs through all multi-indices. Then the equation $Uv_{i_1} = 0$ is equivalent to the equations

$$\sum_{i=1}^{l} D_{I_j K_i} w_i = 0, j = 1, 2, \cdots, l,$$

where $w_i = B_{K_i} x_{i_1}$. Let

$$\mathcal{D} = \begin{pmatrix} D_{I_1 K_1} & D_{I_1 K_2} & \cdots & D_{I_1 K_l} \\ D_{I_2 K_1} & D_{I_2 K_2} & \cdots & D_{I_2 K_l} \\ \cdots & \cdots & \cdots & \cdots \\ D_{I_l K_1} & D_{I_l K_2} & \cdots & D_{I_l K_l} \end{pmatrix}.$$

We claim that $\mathcal{D}$ is invertible. Hence, $w_i = 0$ for all i. Consequently, the invertibility of some B_K implies that $x_{i_1} = (D_{i_1 1}(t), D_{i_1 2}(t), \cdots, D_{i_1 q}(t))^T = 0$. It follows that $D(t) = 0$ for all $t \neq t_j, j = 1, \cdots, q$. We conclude that

$$U'(t) = \sum_{j=1}^{q} D(t_j) U(t - t_j).$$

To complete the proof, it remains to prove that $\mathcal{D}$ is invertible. We require the following lemma.

Lemma 11. *Let* $I = (i_2, i_3, \cdots, i_{r-1})\,, K = (k_2, k_3, \cdots, k_{r-1})\,, D_{IK} = D_{i_2k_2} \cdots D_{i_{r-1}k_{r-1}}$ *where* I_j, K_m *runs through all* $1 \le i_p, k_\ell \le q$*,*

$$\begin{vmatrix} D_{I_1K_1} & D_{I_1K_2} & \cdots & D_{I_1K_l} \\ D_{I_2K_1} & D_{I_2K_2} & \cdots & D_{I_2K_l} \\ \cdots & \cdots & \cdots & \cdots \\ D_{I_lK_1} & D_{I_lK_2} & \cdots & D_{I_lK_l} \end{vmatrix} = (\det D)^{n^{r-3}(r-2)}\,,$$

where D *is the* $q \times q$ *matrix whose* (i,j)*th element is* D_{ij}*.*

Proof. We prove by induction on r. Note that the determinant is of dimension $q^{r-2} \times q^{r-2}$. If $r = 3$, then the above determinant is $\det D$. Assume the formula is true for r. When r is increased by 1, there will be one more index. The determinant is

$$\begin{vmatrix} D_{I_1K_1}D & D_{I_1K_2}D & \cdots & D_{I_1K_l}D \\ D_{I_2K_1}D & D_{I_2K_2}D & \cdots & D_{I_2K_l}D \\ \cdots & \cdots & \cdots & \cdots \\ D_{I_lK_1}D & D_{I_lK_2}D & \cdots & D_{I_lK_l}D \end{vmatrix},$$

which upon applying Lemma 10 we obtain

$$\begin{aligned} &\begin{vmatrix} D_{I_1K_1} & D_{I_1K_2} & \cdots & D_{I_1K_l} \\ D_{I_2K_1} & D_{I_2K_2} & \cdots & D_{I_2K_l} \\ \cdots & \cdots & \cdots & \cdots \\ D_{I_lK_1} & D_{I_lK_2} & \cdots & D_{I_lK_l} \end{vmatrix}^q (\det D)^{q^{r-2}} \\ &= (\det D)^{q\left(q^{r-3}(r-2)\right) + q^{r-2}} \\ &= (\det D)^{n^{r-2}(r-1)} \end{aligned}$$

which is the required conclusion. □

Now, the invertibility of $\mathcal{D}$ follows from the preceding lemma and the fact that D is invertible, according to lemma 9. This completes the proof of Theorem 8

THEOREM 12. *With the same notations and assumptions as in Theorem 8, we assume further that there exist no two non-zero vectors* α *and* β *such that* $\alpha^T U(t)$ *is stochastically independent of* $\beta^T U(t)$*. Then there is only one non-zero matrix* $D\,(t)$ *in* $U'\,(t) = D\,(t) * U\,(t)\,.$

Proof. We shall prove by contradiction. Suppose there is more than one non-zero matrix. Then there are integers $q > p$ such that $D(p), D(p+1), \cdots, D(q)$ are the only possibly non-zero matrices, and that $D(p), D(q)$ are non-zero. Since $U'(t+q)$ is independent of $U'(t+p), D(q)U(t)$ is independent of $D(p)U(t)$. As $D(q)$ and $D(p)$ each contains a non-zero row vector, we derive a contradiction to the assumption that no two non-trivial linear combinations of $U(t)$ are stochastically independent of each other. □

References

[1] Brillinger, D. R. (1969). The calculation of cumulants via conditioning. *Annals of the Institute of Statistical Mathematics*, **21**, 215-218.

[2] Brockwell, P. J. & Davis, R. A. (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.

[3] Cheng, Q. (1990) Maximum standardized cumulant deconvolution of non-Gaussian linear processes. *Ann. Statist.*, **18**, 1774-1783.

[4] Cheng, Q. (1992) On the unique representation of non-Gaussian linear processes. *Ann. Statist.*, **20**, 1143-1145.

[5] Donoho, D. L. (1981). On minimum entropy deconvolution. In Applied Time Series Analysis II (D. F. Findley, ed.) 565-608. Academic, New York.

[6] Findley, D. (1986). The uniqueness of moving average representation with independent and identically distributed random variables for non-Gaussian stationary time series. *Biometrika*, **73**, 520-521.

[7] Findley, D. (1990). Amendments and corrections: The uniqueness of moving average representation with independent and identically distributed random variables for non-Gaussian stationary time series. *Biometrika*, **77**, 235.

[8] Hannan, E. J. (1970). *Multiple Time Series*. Wiley, New York.

[9] Jammalamadaka, S. R., Subba Rao, T. and Terdik, G. (2004). Higher order cumulants of random vectors, differential operators, and applications to statistical inference and time series. Can be downloaded from `http://www.pstat.ucsb.edu/faculty/jammalam/html/research%20publication_files/DiffOpApplAMSV12.pdf`

[10] Kagan, A. M., Linnik, Y. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics*. Wiley, New York.

[11] Lii, K. S. and Rosenblatt, M. (1982). Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes. *Ann. Statist.*, **10**, 1195-1208.

[12] Loeve, M. (1963). *Probability Theory*. Van Nostrand, New York.

[13] Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Birkhauser, Boston.

[14] Rudin, W. (1962). *Fourier Analysis on Groups*. Interscience, New York.