Probab. Th. Rel. Fields 79, 1–27 (1988)

Probability Theory and Related Fields

Innovations and Wold Decompositions of Stable Sequences*

Stamatis Cambanis, Clyde D. Hardin Jr.[1], and Aleksander Weron[2]

Center for Stochastic Processes, Department of Statistics, University of North Carolina, Chapel Hill, NC 27599-3260, USA

Summary. For symmetric stable sequences, notions of innovation and Wold decomposition are introduced, characterized, and their ramifications in prediction theory are discussed. As the usual covariance orthogonality is inapplicable, the non-symmetric James orthogonality is used. This leads to right and left innovations and Wold decompositions, which are related to regression prediction and least p^{th} moment prediction, respectively. Independent innovations and Wold decompositions are also characterized; and several examples illustrating the various decompositions are presented.

0. Introduction

The problem of prediction for processes with infinite variance is of compelling practical and theoretical interest, although very little work on this subject exists. The early work of Urbanik [19] shows that the classical theory is limited to the important class of nonanticipating moving averages. But, while all regular stationary Gaussian processes are indeed nonanticipating moving averages, among the non-Gaussian stable stationary processes those that are nonanticipating moving averages form a thin class [1]. Another class of stationary stable processes, which is disjoint from the nonanticipating moving averages, consists of the hamonizable stable processes whose prediction was considered by Hosoya [6] and in [1]. The prediction of autoregressive moving averages was considered by Cline and Brockwell [3]. Except for the work of Urbanik all other papers treat the discrete time case.

This paper describes the extent to which a reasonable theory of prediction holds for discrete time stable sequences, and also uncovers some intriguing behavior, unsuspected from the classical Gaussian theory. In view of how little is currently known, and especially how differently innovations are built up for non-Gaussian random models, it is hoped that this paper may serve as a first step in approaching the more difficult problem of prediction for continuous

* Research supported by the Air Force Office of Scientific Research Contract F49620 82 C 0009

[1] Now at The Analytic Sciences Corporation, 55 Walkers Brook Drive, Reading, MA 01867, USA

[2] Now at the Institute of Mathematics, Technical University, 50-370 Wroclaw, Poland

time stable processes, as well as for other non-Gaussian (e.g., general infinitely divisible) processes.

The Wold (or orthogonal) decomposition of Gaussian and other second-order stochastic processes is a fundamental tool in their study, and in particular in their predictions. For stable and other p^{th}-order processes (with $p<2$) the lack of second moments renders the usual L^2 notion of orthogonality inapplicable, and thus orthogonal decomposition of these processes does not even make sense a priori. There are, however, notions of orthogonality in Banach spaces; and one of these, due to G. Birkhoff and popularized by R.C. James [7], seems appropriate in this context. Still, the situation is much more complex than in the second-order case, as we shall see shortly.

The purpose of this paper is to examine James' orthogonality as an appropriate prediction theoretic tool for symmetric α-stable (SαS) random variables and processes. Using this orthogonality we define appropriate notions of Wold decomposition for SαS sequences and characterize those sequences which can be so decomposed. The role of independence is also examined. (Orthogonality implies independence in Gaussian systems, but not in stable systems!)

The organization of the paper is as follows. Section 1 includes some preliminary facts, which clarify the role of orthogonality in stable systems. We give some characterizations of orthogonality (Corollary 1.3); for example, we find that for jointly SαS r.v.'s X and Y, X is orthogonal to Y if and only if $E(Y|X)=0$. We also characterize the linearity of a conditional expectation in a stable system in terms of an appropriate orthogonality.

In Sect. 2 we define two kinds of innovations, right orthogonal and left orthogonal, and Wold decompositions for SαS sequences, and give necessary and sufficient conditions for their existence. It turns out that a right Wold decomposition exists if and only if right innovations exist, if and only if the regressions on the past are linear (Theorem 2.3). Left innovations always exist (Proposition 2.8), while a left Wold decomposition exists if and only if the metric projections on the past are linear (Theorem 2.10); and these results hold for general p^{th} order processes. We also define "non-linear" innovations and Wold decompositions. Right nonlinear innovations and Wold decompositions always exist (Theorem 2.2). Left nonlinear innovations always exist (Proposition 2.8) and we note that a left nonlinear Wold decomposition exists whenever a left Wold decomposition exists. The right and left innovations and Wold decompositions have precisely the properties required to solve the problem of predicting m-steps ahead based on past observations, and they correspond to regression prediction and best prediction in the usual p^{th} order moment sense $(1<p<\alpha)$ respectively. Thus when a right or left Wold decomposition exists, the m-step linear regression prediction or best linear prediction has a fairly simple solution. However, when a Wold decomposition does not exist, then the prediction problem becomes difficult indeed as is illustrated by the case of harmonizable stable sequences (cf. [1]).

In Sect. 3, an independent decomposition is introduced and spectral necessary and sufficient conditions are given for its existence. Section 4 consists entirely of examples, intended to illustrate the various decompositions and some of the complexities involved.

After this work was submitted we received the paper by Miamee and Pourahmadi "Wold decomposition, prediction and parametrization of stationary processes with infinite variance", Probab. Theor. Rel. Fields 79, 145–164 (1988), where a *finite* term left Wold decomposition is also given for p^{th} order processes.

1. Orthogonality and Stable Systems

A collection of random variables $\{X_t\colon t\in T\}$ defined on (Ω, Σ, P) will be called *jointly symmetric* α*-stable* or a *symmetric* α*-stable process* if each finite real-linear combination $\sum \lambda_j X_{t_j}$ has a symmetric stable distribution of index α. We abbreviate "symmetric α-stable" by SαS. If X is SαS, then for $0<p<\alpha$, we have $E|X|^p<\infty$, so that a SαS process $\{X_t\}$ is a p^{th} order process, i.e., $\{X_t\}\subseteq L^p(\Omega, \Sigma, P)$. A useful tool in the analysis of SαS processes is the so-called spectral representation theorem. The version we will need here says that if $\{X_n\colon n\in A\}$ (where A is finite or denumerably infinite) is a SαS process, then there exist functions $\{f_n\colon n\in A\}\subseteq L^\alpha[0,1]$ such that

$$-\log E\exp\left(i\sum_{j=1}^{n}\lambda_j X_{n_j}\right)=\left\|\sum_{j=1}^{n}\lambda_j f_{n_j}\right\|_\alpha^\alpha.$$

Further, if $\{Z(s)\colon s\in[0,1]\}$ is "α-stable motion", i.e., an independent increments SαS process with $-\log E\exp itZ(s)=s|t|^\alpha$, then the process $\{Y_n\}$ defined by

$$Y_n=\int_0^1 f_n(s)\,dZ(s)$$

is stochastically equivalent to $\{X_n\}$, and we say that $\{X_n\}$ *is represented by* $\{f_n\}$. The spectral representation was first expressed in this form by Kuelbs [9]; for more information consult [5].

Now let $\mathscr{L}$ be a normed linear space, with norm $\|\cdot\|$. For $x, y\in\mathscr{L}$, we say that x is *(James) orthognal to* y, written $x\perp y$, if

$$\|x+\lambda y\|\geqq\|x\|$$

for all scalars λ. For subspaces M and N of $\mathscr{L}$, we say $M\perp N$ if $m\perp n$ for all $m\in M$ and $n\in N$. If $\mathscr{L}$ is in fact a Hilbert space, this defines the usual "inner product" orthogonality. For general Banach spaces, however, this is a nonsymmetric notion, i.e., x may be orthogonal to y, but not vice versa.

This definition makes sense for random variables with p^{th} moments in that we may take $(\mathscr{L}, \|\cdot\|)$ to be $L^p(\Omega)$ with the usual norm. For X and Y in $L^p(\Omega)$, if X is orthogonal to Y, we will write $X\perp_p Y$. The relation $\perp_p$ is well-defined for jointly SαS random variables as long as $1\leqq p<\alpha$.

The following known characterization of orthogonality will be useful for us. For a proof, consult [15; Thm. 1.11, p. 56 and Lemma 1.14, p. 92]

Lemma 1.1. *Let X and Y be random variables with p^{th} moments, $p>1$. Then $X\perp_p Y$ if and only if $EX^{\langle p-1\rangle}Y=0$.*

Here, we use the convention that for complex z and real q, $z^{\langle q\rangle}$ denotes $|z|^{q-1}\bar{z}$. (We take $0^{\langle q\rangle}=0$.)

A point evident from this lemma and crucial for us is that the orthogonality relation is "linear" in the second argument, but not in the first, i.e., $X\perp_p Y$ and $X\perp_p Z$ implies $X\perp_p(aY+bZ)$ for all a, b – but we may have $X\perp_p Z$ and $Y\perp_p Z$ without $X+Y\perp_p Z$.

The next lemma is somewhat curious.

Lemma 1.2. *Let $\alpha>1$ and $\{X, Y\}$ be jointly SαS represented by $\{f, g\}$. Then for all $p\in(1,\alpha)$,*

$$\frac{EX^{\langle p-1\rangle}Y}{E|X|^p}=\frac{\int f^{\langle\alpha-1\rangle}g\,dm}{\int|f|^\alpha dm}$$

(m is Lebesgue measure on $[0,1]$).

Remark. Note that the right-hand side does not depend on p. It follows from this and Lemma 1.1 that for such X and Y, $X\perp_p Y$ for *some* $p\in(1,\alpha)$ if and only if $X\perp_p Y$ for *all* such p, if and only if $f\perp_p g$. We shall henceforth say in this case simply that X is orthogonal to Y, omitting mention of p, and write $X\perp Y$.

Proof of Lemma 1.2. Let X_0 be SαS with $Ee^{itX_0}=e^{-|t|^\alpha}$. Now, $E\exp[it(X+\lambda Y)]=\exp[-\|f+\lambda g\|_\alpha^\alpha|t|^\alpha]$, which shows that $X+\lambda Y$ is distributed as $\|f+\lambda g\|_\alpha X_0$. Therefore,

$$E|X+\lambda Y|^p=\|f+\lambda g\|_\alpha^p E|X_0|^p.$$

Differentiating this expression with respect to λ and putting $\lambda=0$, we obtain when $1<p<\alpha$ that

$$EX^{\langle p-1\rangle}Y=E|X_0|^p\|f\|_\alpha^{p-\alpha}\int f^{\langle\alpha-1\rangle}g\,dm=E|X|^p\|f\|_\alpha^{-\alpha}\int f^{\langle\alpha-1\rangle}g\,dm,$$

proving the lemma. □

It follows from Lemma 1.2 that

$$E(Y|X)=\frac{\int f^{\langle\alpha-1\rangle}g\,dm}{\int|f|^\alpha dm}X=\frac{EX^{\langle p-1\rangle}Y}{E|X|^p}X$$

where the first equality is established by Kanter [8]. This combined with Lemma 1.1 shows

Corollary 1.3. *For $1<p<\alpha$ and SαS $\{X, Y\}$ represented by $\{f, g\}$ we have*

$$X\perp Y-\frac{EX^{\langle p-1\rangle}Y}{E|X|^p}X=Y-E(Y|X);$$

and the following are equivalent:

(i) $X \perp Y$,

(ii) $E(Y|X)=0$,

(iii) $EX^{\langle p-1\rangle} Y=0$,

(iv) $\int f^{\langle\alpha-1\rangle} g\, dm=0$.

We note that if X and Y are independent SαS variables, then necessarily $X \perp Y$ and $Y \perp X$. The converse is not true, however, since Schilder [14] has shown that X and Y are independent if and only if their representatives f and g have a.e. disjoint support (i.e., $f \cdot g=0$ a.e.). Clearly there exist f and g with $\int f^{\langle\alpha-1\rangle} g\, dm=0$ yet $f \cdot g \neq 0$ a.e. In fact, orthogonality implies independence *only* in Gaussian systems, in the following sense.

Proposition 1.4. *Let $1<\alpha \leqq 2$, and let L be a closed linear space of SαS random variables with $\dim(L)>1$. Suppose that whenever $X, Y \in L$ and $X \perp Y$, then X is independent of Y. Then $\alpha=2$, i.e. L consists of mean-zero Gaussian random variables.*

Proof. Choose an arbitrary non-zero $X \in L$ and let $1<p<\alpha$. By the hypothesis $\dim(L)>1$ we may find $Z \in L$ such that $Z \neq \lambda X$ for any $\lambda \in \mathbb{R}$. Let $\beta = EX^{\langle p-1\rangle} Z/E|X|^p$. This gives that $EX^{\langle p-1\rangle}(Z-\beta X)=0$. Since $Z-\beta X \neq 0$, we may find a constant b so that $Y \triangleq b(Z-\beta X)$ is distributed as X. Since $EX^{\langle p-1\rangle} Y=0$, Corollary 1.3 shows $X \perp Y$, and so X and Y are independent by hypothesis. This implies that (X, Y) is distributed as (Y, X), and hence that

$$E(X+Y)^{\langle p-1\rangle}(X-Y)=E(X+Y)^{\langle p-1\rangle} X-E(X+Y)^{\langle p-1\rangle} Y=0.$$

Hence $X+Y \perp X-Y$, and again this means that $X+Y$ is independent of $X-Y$.

Now let c be such that $\phi(t) \triangleq E\exp(itX)=\exp(-c|t|^\alpha)=E\exp(itY)$. Then by independence we have that for all t,

$$\begin{aligned} E\exp\{i[t(X+Y)+t(X-Y)]\} &= E\exp\{it(X+Y)\} \cdot E\exp\{it(X-Y)\} \\ &= \phi^4(t)=\exp(-4c|t|^\alpha), \end{aligned}$$

and

$$\begin{aligned} E\exp\{i[t(X+Y)+t(X-Y)]\} &= E\exp\{i2tX\} \\ &= \phi(2t)=\exp(-2^\alpha c|t|^\alpha). \end{aligned}$$

Therefore $2^\alpha=4$ and $\alpha=2$. □

The equivalences of Corollary 1.3 can be seen in a broader context. Let $1<p<\alpha$, and let $\{X_t\colon t \in T\}$ be any SαS process represented by $\{f_t\colon t \in T\}$ (T here is arbitrary). Fix an arbitrary subset S of T, let $t \in T \setminus S$, and define $L(S) = \overline{sp}\{X_s\colon s \in S\}_{L^p(\Omega)}$ and $L'(S)=\overline{sp}\{f_s\colon s \in S\}_{L^\alpha}$. The following result gives necessary and sufficient conditions for the conditional expectation to be linear, i.e. to belong to $L(S)$.

Proposition 1.5. *The following are equivalent:*

(i) $E(X_t|X_s\colon s\in S)\in L(S)$.

(ii) *There exists* $\tilde{X}\in L(S)$ *such that* $L(S)\perp X_t-\tilde{X}$ *(in which case* $E(X_t|X_s\colon s\in S)=\tilde{X}$).

(iii) *There exists* $\tilde{g}\in L'(S)$ *such that* $L'(S)\perp_\alpha f_t-\tilde{g}$ *(in which case* $E(X_t|X_s\colon s\in S)$ *is represented by* $\tilde{g}$*).*

Proof. Let J_0 be the spectral representation map for $\{X_t\}$, i.e., $J_0(X_t)=f_t$. Use the same argument as in the proof of Lemma 1.2 to see that for any $t_j\in T$ and scalars λ_j,

$$\left\|\sum_{j=1}^{n}\lambda_j X_{t_j}\right\|_{L^p(\Omega)}=\|\sum\lambda_j f_{t_j}\|_{L^\alpha}\cdot\|X_0\|_{L^p(\Omega)}$$

where X_0 is as in Lemma 1.2. Putting $c=\|X_0\|_{L^p(\Omega)}$, this shows cJ_0 extends by linearity and continuity to an isometry cJ of $L(S)$ onto $L'(S)$. Hence (ii) and (iii) are equivalent.

We show (i) and (iii) are equivalent. Let Y by any arbitrary element of $L(S)$, and define $h=J(Y)$. (Or equivalently, let h be arbitrary in $L'(S)$ and define $Y=J^{-1}(h)$.) For $\phi(u)\triangleq E\exp[i(uX_t+Y)]$, we have $\phi(u)=\exp[-\|uf_t+h\|_\alpha^\alpha]$, and thus, putting $\hat{X}=E(X_t|X_s\colon s\in S)$, that

$$E e^{iY}\hat{X}=EX_t e^{iY}=-i\phi'(0)=i\alpha\exp[-\|h\|_\alpha^\alpha]\int h^{\langle p-1\rangle}f_t\,dm.$$

Now for arbitrary $\tilde{X}\in L(S)$, let $\tilde{g}=J(\tilde{X})$. (Again, we may let $\tilde{g}\in L'(S)$ and define $\tilde{X}=J^{-1}(\tilde{g})$.) Define $\psi(u)=E\exp[i(u\tilde{X}+Y)]$, and note that $\psi(u)=\exp[-\|u\tilde{g}+h\|_\alpha^\alpha]$, and

$$E e^{iY}\tilde{X}=-i\psi'(0)=i\alpha\exp[-\|h\|_\alpha^\alpha]\int h^{\langle\alpha-1\rangle}\tilde{g}\,dm.$$

This gives

$$E e^{iY}(\hat{X}-\tilde{X})=i\alpha\exp[-\|h\|_\alpha^\alpha]\int h^{\langle\alpha-1\rangle}(f_t-\tilde{g})\,dm.$$

Since both $\hat{X}$ and $\tilde{X}$ are measurable with respect to $\sigma\{X(s)\colon s\in S\}$, we have that $\hat{X}=\tilde{X}$ if and only if $E e^{iY}(\hat{X}-\tilde{X})=0$ for all $Y\in L(S)$ (see, e.g., [8] or [10]). This fact and Lemma 1.1 applied to the last equation give us the equivalence of (i) and (iii), proving the proposition. □

In particular, this shows how the linearity of regression is related to orthogonality.

Corollary 1.6. *The following are equivalent.*

(i) $E(X_t|X_s\colon s\in S)=0$.

(ii) $\overline{sp}\{X_s\colon s\in S\}_{L^p(\Omega)}\perp X_t$.

(iii) $\overline{sp}\{f_s\colon s\in S\}_{L^\alpha}\perp_\alpha f_t$.

2. Orthogonal Decomposition of Stable Sequences

Throughout this section we assume $1<\alpha<2$ and take p such that $1<p<\alpha$. Also we let $\{X_n: -\infty<n<\infty\}$ be a SαS sequence on (Ω, Σ, P). We define the linear spaces of the sequence:

$$L_n=\overline{sp}\{X_k: k\leqq n\}_{L^p(\Omega)},$$
$$L_{-\infty}=\bigcap_n L_n,$$

and the corresponding nonlinear spaces:

$$\mathscr{L}_n=L^p(\Omega, \Sigma_n, P),$$
$$\mathscr{L}_{-\infty}=\bigcap_n \mathscr{L}_n,$$

where $\Sigma_n=\sigma\{X_k, k\leqq n\}$. Note that L_n consists of SαS random variables, while $\mathscr{L}_n$ contains much more. Note also that since for every $X\in\overline{sp}\{X_n: -\infty<n<\infty\}$ with representative $f\in L^\alpha$ we have, as in the proof of Lemma 1.2, $\|X\|_{L^p(\Omega)}=C_p\|f\|_{L^\alpha}$ for some constant $C_p=\|X_0\|_{L^p(\Omega)}$ depending only on p and not on X, the choice of p in $(1, \alpha)$ throughout the following is immaterial.

We will be concerned with the orthogonal decomposition of these spaces. Our notation, which is somewhat non-standard, is as follows. For a Banach space $\mathscr{M}$ and closed subspaces $M_1, M_2, \ldots$, the symbol $M_1+\ldots+M_n$ $\left(\text{or } \sum_{j=1}^{n} M_j\right)$ denotes the subspace $\{x_1+\ldots+x_n: x_j\in M_j, 1\leqq j\leqq n\}$. Also, $M_1+M_2+\ldots$ $\left(\text{or } \sum_{j=1}^{\infty} M_j\right)$ is defined to be the subspace $\overline{\bigcup_n \sum_{j=1}^{n} M_j}$. Writing $\mathscr{M}=M_1 \underset{\rightarrow}{\oplus} \ldots \underset{\rightarrow}{\oplus} M_n$ $\left(\text{or } \mathscr{M}=\sum_{j=1}^{n} \underset{\rightarrow}{\oplus} M_j\right)$ means that $\mathscr{M}=M_1+\ldots+M_n$ and also that

$$(M_1+\ldots+M_k)\perp(M_{k+1}+\ldots+M_n) \qquad \text{for all } 1\leqq k<n. \tag{2.1}$$

Writing $\mathscr{M}=M_1 \underset{\leftarrow}{\oplus} \ldots \underset{\leftarrow}{\oplus} M_n$ $\left(\text{or } \mathscr{M}=\sum_{j=1}^{n} \underset{\leftarrow}{\oplus} M_j\right)$ means that $\mathscr{M}=M_1+\ldots+M_n$ and that

$$(M_n+\ldots+M_{k+1})\perp(M_k+\ldots+M_1) \qquad \text{for all } 1\leqq k<n, \tag{2.2}$$

i.e., that $\mathscr{M}=M_n \underset{\rightarrow}{\oplus} \ldots \underset{\rightarrow}{\oplus} M_1$. Thus the statements $\mathscr{M}=M_1 \underset{\rightarrow}{\oplus} M_2$ and $\mathscr{M}=M_1 \underset{\leftarrow}{\oplus} M_2$ are, in general, distinct. Writing $\mathscr{M}=\sum_{j=1}^{\infty} \underset{\rightarrow}{\oplus} M_j$ $\left(\text{respectively, } \mathscr{M}=\sum_{j=1}^{\infty} \underset{\leftarrow}{\oplus} M_j\right)$ will denote that $\mathscr{M}=\sum_{j=1}^{\infty} M_j$ and further that (2.1) (respectively (2.2)) holds for all n.

If $\mathscr{M}=\sum_{j=1}^{\infty}\underset{\rightarrow}{\oplus} M_j$ and we pick $0\neq x_j\in M_j$, it follows that $\{x_j\}$ *forms a basis for its closed linear span*, i.e., each $x\in\overline{sp}\{x_j: j=1,2,\ldots\}$ has a unique norm-convergent expansion $x=\sum_{j=1}^{\infty}\lambda_j x_j$ for some scalars λ_j. This is so because a necessary and sufficient condition for $\{x_j\}$ to be a basis for its closed linear span is the existence of $K<\infty$ such that for all n, $m\leqq n$, and scalars β_j, $\left\|\sum_{j=1}^{m}\beta_j x_j\right\| \leqq K\left\|\sum_{j=1}^{n}\beta_j x_j\right\|$ (see, e.g., [16]); and, because of orthogonality,

$$\left\|\sum_{j=1}^{n}\beta_j x_j\right\|=\left\|\sum_{j=1}^{m}\beta_j x_j+\sum_{j=m+1}^{n}\beta_j x_j\right\|\geqq\left\|\sum_{j=1}^{m}\beta_j x_j\right\|.$$

The same argument cannot be made in the case $\mathscr{M}=\sum_{j=1}^{\infty}\underset{\leftarrow}{\oplus} M_j$.

Right Innovations and Wold Decomposition

We will say that $\{X_n\}$ has *right innovations* if for each n there is a subspace N_n so that $L_n=L_{n-1}\underset{\rightarrow}{\oplus} N_n$. N_n is necessarily of dimension one or zero (by an elementary argument). Similarly, we say that $\{X_n\}$ has *right non-linear innovations* if for each n there is a subspace $\mathscr{N}_n$ so that $\mathscr{L}_n=\mathscr{L}_{n-1}\underset{\rightarrow}{\oplus}\mathscr{N}_n$.

We say that $\{X_n\}$ has a *right Wold decomposition* if there are subspaces N_n, $-\infty<n<\infty$, so that for each n, $L_n=\left(\sum_{k=0}^{\infty}\underset{\leftarrow}{\oplus} N_{n-k}\right)\underset{\leftarrow}{\oplus} L_{-\infty}$, $L_n\perp N_m$ for all $m>n$, and further each $Z\in\sum_{k=0}^{\infty}\underset{\leftarrow}{\oplus} N_{n-k}$ has an L^p-convergent expansion $Z=\sum_{k=0}^{\infty}W_{n-k}$, $W_j\in N_j$, which is then necessarily unique. In this case it is easy to see that we can write $X_n=Y_n+Z_n$, where

(i) $\{Y_n\}$ and $\{Z_n\}$ are jointly SαS processes,
(ii) $\{Y_n\}\subseteq L_{-\infty}$ (the "remote past") and $\{Y_n\}\perp\{Z_n\}$,
(iii) there exist $\xi_j\in N_j$ and $a_{k,n}\in\mathbb{R}$ so that $Z_n=\sum_{k=0}^{\infty}a_{k,n}\xi_{n-k}$.

In the case that $\{X_n\}$ is stationary and not completely deterministic (i.e., $L_{-\infty}\neq L_0$), we may choose $\|\xi_j\|_{L^p(\Omega)}=1$ and claim that $a_{k,n}$ is independent of n, i.e., Z_n is a moving average of an "orthonormal sequence".

Similarly, we can define *right non-linear Wold decomposition* by requiring the existence of $\mathscr{N}_n$ so that $\mathscr{L}_n=\left(\sum_{k=0}^{\infty}\underset{\leftarrow}{\oplus}\mathscr{N}_{n-k}\right)\underset{\leftarrow}{\oplus}\mathscr{L}_{-\infty}$, $\mathscr{L}_n\perp\mathscr{N}_m$ for $m>n$, and with the property that each $\mathscr{X} \in \sum_{k=0}^{\infty} \overleftarrow{\oplus} \mathscr{N}_{n-k}$ has a norm convergent expansion $\mathscr{X} = \sum_{k=0}^{\infty} \mathscr{W}_{n-k}$, $\mathscr{W}_j \in \mathscr{N}_j$, which is then unique.

The first result, Proposition 2.1, is the key ingredient to the proof that right innovations, linear or non-linear, imply the corresponding Wold decomposition (see Theorems 2.2 and 2.3). This proposition is implicit in [2] and [4]; we include a proof here for completeness.

Proposition 2.1. *Suppose that $\mathscr{M}$ is a closed subspace of some L^p space, $p>1$, and that there exist closed subspaces $\mathscr{M}_n$ and $\mathscr{O}_n$ of $\mathscr{M}$ with $\mathscr{M} = \mathscr{M}_n \overrightarrow{\oplus} \mathscr{O}_n \overrightarrow{\oplus} \ldots \overrightarrow{\oplus} \mathscr{O}_1$ for each $n \geqq 1$. Then $\mathscr{M} = \left(\sum_{n=1}^{\infty} \overleftarrow{\oplus} \mathscr{O}_n\right) \overleftarrow{\oplus} (\bigcap_n \mathscr{M}_n)$ and each $k \in \sum_{n=1}^{\infty} \overleftarrow{\oplus} \mathscr{O}_n$ has a unique norm convergent expansion $k = \sum_{n=1}^{\infty} o_n$, $o_n \in \mathscr{O}_n$.*

Proof. Define $\mathscr{M}_\infty = \bigcap_n \mathscr{M}_n$, $\mathscr{K}_n = \mathscr{O}_n \overrightarrow{\oplus} \ldots \overrightarrow{\oplus} \mathscr{O}_1$, and $\mathscr{K}_\infty = \overline{\bigcup_n \mathscr{K}_n}$. We first show that $\mathscr{M} = \mathscr{M}_\infty \overrightarrow{\oplus} \mathscr{K}_\infty$. Clearly, $\mathscr{M}_\infty \perp \mathscr{K}_n$ for each n, and by continuity, $\mathscr{M}_\infty \perp \mathscr{K}_\infty$. Now for $x \in \mathscr{M}$, write $x = m_n + k_n$ with $m_n \in \mathscr{M}_n$, $k_n \in \mathscr{K}_n$. Since $m_n \perp k_n$ we have that

$$\|m_n\| \leqq \|m_n + k_n\| = \|x\| \quad \text{and} \quad \|k_n\| \leqq \|x - m_n\| \leqq 2\|x\|.$$

The sequences m_n and k_n, being norm bounded in a reflexive Banach space, have simultaneously weakly convergent subsequences, say $\{m_{n_i}\}$ and $\{k_{n_i}\}$ with weak limits m_∞ and k_∞, respectively. It is clear that $x = m_\infty + k_\infty$, and that $k_\infty \in \mathscr{K}_\infty$, proving $\mathscr{M} = \mathscr{M}_\infty \overrightarrow{\oplus} \mathscr{K}_\infty$.

It remains to show that each element $k \in \mathscr{K}_\infty$ has a unique norm convergent expansion $k = \sum_{n=1}^{\infty} o_n$, $o_n \in \mathscr{O}_n$. For each n we can write $k = m_n + k_n$ uniquely where $m_n \in \mathscr{M}_n$ and $k_n \in \mathscr{K}_n$. In turn we may write $k_n = o_1 + \ldots + o_n$ uniquely with $o_j \in \mathscr{O}_j$. Define the operator Q_n: $\mathscr{K}_\infty \to \mathscr{K}_n$ by $Q_n k = k_n$. It is easy to see that $Q_n Q_l = Q_{n \wedge l}$. Also by orthogonality we have that

$$\|Q_n k\| \leqq \|k - Q_n k\| + \|k\| = \|m_n\| + \|k\| \leqq \|m_n + k_n\| + \|k\| = 2\|k\|$$

so that $\{Q_n\}$ is a bounded sequence. Clearly, $s\text{-}\lim_{n \to \infty} Q_n k = k$ for any $k \in \bigcup_m \mathscr{K}_m$. Hence, by continuity, we have for any $k \in \mathscr{K}_\infty$ that

$$k = s\text{-}\lim_{n \to \infty} Q_n k = s\text{-}\lim_{n \to \infty} \sum_{i=1}^{n} o_i = \sum_{n=1}^{\infty} o_n. \quad \square$$

Theorem 2.2. *$\{X_n\}$ has right non-linear innovations and a right non-linear Wold decomposition.*

Proof. Note that for each n, $\mathscr{L}_{n-1}=\{E(X|\Sigma_{n-1}): X\in\mathscr{L}_n\}$. Define $\mathscr{N}_n=\{X-E(X|\Sigma_{n-1}): X\in\mathscr{L}_n\}$. Clearly, each element of $\mathscr{L}_n$ is the sum of an element of $\mathscr{L}_{n-1}$ and an element of $\mathscr{N}_n$. To see $\mathscr{L}_{n-1}\perp\mathscr{N}_n$, let $X\in\mathscr{L}_{n-1}$ and $Y\in\mathscr{N}_n$, and note that $E(Y|\Sigma_{n-1})=0$. Hence,

$$EX^{\langle p-1\rangle}Y=EE(X^{\langle p-1\rangle}Y|\Sigma_{n-1})=EX^{\langle p-1\rangle}E(Y|\Sigma_{n-1})=0,$$

and thus $\{X_n\}$ has right non-linear innovations.

To see that $\{X_n\}$ has a right non-linear Wold decomposition, fix k and note that by the argument above

$$\begin{aligned}\mathscr{L}_k&=\mathscr{L}_{k-1}\underset{\rightarrow}{\oplus}\mathscr{N}_k\\&=(\mathscr{L}_{k-2}\underset{\rightarrow}{\oplus}\mathscr{N}_{k-1})\underset{\rightarrow}{\oplus}\mathscr{N}_k\\&=\mathscr{L}_{k-2}\underset{\rightarrow}{\oplus}(\mathscr{N}_{k-1}\underset{\rightarrow}{\oplus}\mathscr{N}_k)\\&\;\vdots\\&=\mathscr{L}_{k-n}\underset{\rightarrow}{\oplus}(\mathscr{N}_{k-n+1}\underset{\rightarrow}{\oplus}\dots\underset{\rightarrow}{\oplus}\mathscr{N}_k).\end{aligned}$$

Now identify $\mathscr{L}_{k-n}$ with $\mathscr{M}_n$ and $\mathscr{N}_{k-n+1}$ with $\mathscr{O}_n$ of Proposition 2.1. □

The next result is somewhat more interesting.

Theorem 2.3. *The following are equivalent.*

(i) $\{X_n\}$ *has a right Wold decomposition.*

(ii) $\{X_n\}$ *has right innovations.*

(iii) $E(X_{n+1}|X_n, X_{n-1},\dots)\in L_n$ *for each* n, *i.e. regressions on the past are linear.*

Proof. The "linear version" of the proof of the second statement of Theorem 2.2 shows that if $\{X_n\}$ has right innovations it has a right Wold decomposition. The converse follows by definition, so (i) and (ii) are equivalent.

We show the equivalence of (ii) and (iii). Take, in the notation of Proposition 1.5, $X_t\triangleq X_{n+1}$ and $S=\{n, n-1,\dots\}$. Then $L(S)=L_n$, and by that Proposition we have that (iii) is equivalent to the existence of $\tilde{X}\in L_n$ such that $L_n\perp X_{n+1}-\tilde{X}$. The latter, clearly, is in turn equivalent to the existence of the required innovation space N_{n+1}. □

Remark. By Theorem 2.2, we may write $X_n=\mathscr{Y}_n+\mathscr{Z}_n$ where $\mathscr{Y}_n\in\mathscr{L}_{-\infty}$ and $\mathscr{Z}_n\in\sum_{k=0}^{+\infty}\underset{\leftarrow}{\oplus}\mathscr{N}_{n-k}$. If $\{X_n\}$ has a right Wold decomposition, we have $X_n=Y_n+Z_n$ as in the comment following the definition. In this case, we must have $\mathscr{Y}_n=Y_n$ and $\mathscr{Z}_n=Z_n$, since $L_{-\infty}\subseteq\mathscr{L}_{-\infty}$ and $N_k=\{X-E(X|\Sigma_{k-1}): X\in L_k\}\subseteq\{X-E(X|\Sigma_{k-1}): X\in\mathscr{L}_k\}=\mathscr{N}_k$, and the decomposition is unique.

Left Innovations and Wold decomposition

We now examine *left* innovations and Wold decompositions. Their definitions are obtained by reversing the arrows in the definitions of their right counterparts, and ignoring the requirement $L_n \perp N_m$ (or $\mathscr{L}_n \perp \mathscr{N}_m$) for the Wold decomposition $\left(\text{as } N_m \perp L_n \text{ follows from } L_n = \left(\sum_{k=0}^{\infty} \underset{\rightarrow}{\oplus} N_{n-k}\right) \underset{\rightarrow}{\oplus} L_{-\infty}\right)$. Also, the "basis property" of $\sum_{k=0}^{\infty} \underset{\rightarrow}{\oplus} N_{n-k}$ is automatically satisfied, as can be seen from the argument following the definitions of $\underset{\rightarrow}{\oplus}$ and $\underset{\leftarrow}{\oplus}$.

As conditional expectation is an appropriate notion for the study of right orthogonality, the metric projection in L^p is an appropriate notion for the study of left orthogonality. For completeness we include the needed definitions and preliminary results here in a compact, self-contained way (see [7, 11, 15]).

Let $(\mathscr{L}, \|\cdot\|)$ be a Banach space, and M a closed subspace of $\mathscr{L}$. For $x \in \mathscr{L}$, an element $m_x \in M$ is called a best approximation to x in M if $\|x - m_x\| \leqq \|x - m\|$ for all $x \in M$. If $\mathscr{L}$ is reflexive and strictly convex (as we henceforth assume throughout) m_x exists and is unique (see [15]). In this case we define $P_M x = m_x$ and call P_M the metric projection onto M. P_M is continuous, bounded, and idempotent, but not in general linear. In fact, if P_M is a linear operator for all closed subspaces M of $\mathscr{L}$, $\mathscr{L}$ must be isometrically isomorphic to a Hilbert space (see [15]).

The relation between orthogonality and metric projection is illustrated by the following two standard results.

Proposition 2.4. *Let Q: $\mathscr{L} \to M$ be an operator (not necessarily linear). Then $Q = P_M$ if and only if $(I-Q)\mathscr{L} \perp M$.*

Proof. $Q = P_M$ if and only if $\|x - Qx\| \leqq \|x - m\|$ for all $m \in M$ and $x \in \mathscr{L}$, if and only if $\|x - Qx\| \leqq \|x - Qx + m\|$ for all $m \in M$ and $x \in \mathscr{L}$, if and only if $(I-Q)\mathscr{L} \perp M$. □

Proposition 2.5. $x \perp M$ *if and only if* $P_M x = 0$.

Proof. $x \perp M$ if and only if $\|x - m\| \geqq \|x\| = \|x - 0\|$ for all $m \in M$, if and only if $P_M x = 0$. □

Although P_M is not a linear operator in general, the following known "quasilinearity" properties are true and will be needed for the proof of Theorem 2.10.

Proposition 2.6. *$P_M(ax) = a P_M x$ for scalars a and $x \in \mathscr{L}$. Also, $P_M(x+m) = P_M x + P_M m$ for all $x \in \mathscr{L}$ and $m \in M$.*

Proof. The homogeneity is obvious. Also, for fixed $x \in \mathscr{L}$, $m \in M$,

$$\|(x+m) - (P_M x + P_M m)\| = \|x - P_M x\| \leqq \|x + m - m'\|$$

for all $m' \in M$, showing $P_M(x+m) = P_M x + P_M m$. □

Proposition 2.7. *If M has codimension one in $\mathscr{L}$, then P_M is a linear operator.*

Proof. We show additivity; the homogeneity follows from Proposition 2.6. Let $z_0 \in \mathscr{L} \setminus M$ be non-zero. Then for $x_1, x_2 \in \mathscr{L}$ there are unique $m_j \in M$ and scalars a_j such that $x_j = m_j + a_j z_0$. Then by Proposition 2.6,

$$\begin{aligned} P_M(x_1+x_2) &= P_M((m_1+m_2)+(a_1+a_2)\,z_0) \\ &= m_1+m_2+(a_1+a_2)\,P_M\,z_0 \\ &= P_M(m_1+a_1 z_0)+P_M(m_2+a_2 z_0) \\ &= P_M x_1 + P_M x_2. \quad \square \end{aligned}$$

We now apply these facts to our situation. Call $L = L_{+\infty} = \overline{sp}\{X_n\colon -\infty < n < \infty\}_{L^p(\Omega)}$, $1<p<\alpha$. Since L^p is reflexive, so is L. Denote by P_n the metric projection of L onto L_n. It turns out that *every* SαS sequence has left innovations:

Proposition 2.8. *$\{X_n\}$ has left innovations, both linear and non-linear.*

Proof. Define $N_n = (I - P_{n-1})\,L_n$. By Proposition 2.4, $N_n \perp L_{n-1}$; and since $P_{n-1} L_n = L_{n-1}$ we have $L_n = L_{n-1} + N_n$. The proof for left non-linear innovations is identical. $\square$

Thus no conditions are needed to split off a "left orthogonal" innovation space. But, unlike the case of right innovations, this is not enough to produce a left Wold decomposition. The problem lies in the impossibility of developing an argument like that of Theorem 2.2, as the following example shows.

Example 2.9. There exist one-dimensional SαS subspaces M_1, M_2, M_3 such that $M_j \perp M_k$ for all $j \neq k$, yet $M_1 + M_2$ is *not* orthogonal to M_3; hence $(M_3 \underset{\leftarrow}{\oplus} M_2) \underset{\leftarrow}{\oplus} M_1 \neq M_3 \underset{\leftarrow}{\oplus} M_2 \underset{\leftarrow}{\oplus} M_1$.

Proof. Let $1<\alpha<2$ and define the functions

$$\begin{aligned} f_1 &= 1_A - 1_B + 1_C - 1_D, \\ f_2 &= 1_A + 2\cdot 1_B - 1_C - 2\cdot 1_D, \\ f_3 &= 1_{[0,1]}, \end{aligned}$$

where $A=[0,\frac{1}{4})$, $B=[\frac{1}{4},\frac{1}{2})$, $C=[\frac{1}{2},\frac{3}{4})$, and $D=[\frac{3}{4},1]$. It may be easily checked that

$$\begin{gathered} \int f_j^{\langle \alpha-1\rangle} f_k\,dm = 0 \qquad \text{for } j \neq k, \\ \int (f_1+f_2)^{\langle \alpha-1\rangle} f_3\,dm = \tfrac{1}{4}(2^{\alpha-1}+1-3^{\alpha-1}) > 0. \end{gathered}$$

By Corollary 1.3 this implies that the SαS subspaces $M_j \triangleq \left\{\lambda \int_0^1 f_j(s)\,dZ(s)\colon \lambda \in \mathbb{R}\right\}$ have the advertised properties. $\square$

There is still, however, a nice characterization, in terms of the metric projections P_n, of those processes having a left Wold decomposition.

Theorem 2.10. *The following are equivalent.*

(i) $\{X_n\}$ *has a left Wold decomposition.*

(ii) *The metric projection operators* $P_n\colon L_{+\infty}\to L_n$ *are linear.*

(iii) *The operators* P_n *commute.*

(iv) *Denoting by* $P_{n,m}$ *the restriction of* P_n *to* L_m, *we have that for all* $k\geqq 1$,

$$P_{n,n+1}\,P_{n+1,n+2}\,\dots\,P_{n+k-1,n+k}=P_{n,n+k}.$$

Proof. We show (iv)$\Rightarrow$(ii)$\Rightarrow$(i)$\Rightarrow$(iv) and (ii)$\Leftrightarrow$(iii).

Assume (iv) holds. By Proposition 2.7, each operator $P_{n+l,n+l+1}$ is linear, implying each $P_{n,n+k}$ is linear. P_n, being linear on each L_{n+k}, is by continuity linear on all of $L_{+\infty}$, giving (ii).

Assume (ii). Define $N_n=(I-P_{n-1})\,L_n$, and let $Z_n\in N_n$. Then $Z_n\perp L_{n-1}$ by Proposition 2.4 and thus $Z_n\perp L_{n-l}$ for $l\geqq 1$. By Proposition 2.5, $P_{n-l}Z_n=0$. The linearity of P_n shows $P_{n-k}(Z_n+Z_{n-1}+\dots+Z_{n-k+1})=0$, giving us by Proposition 2.5 that $N_n+N_{n-1}+\dots+N_{n-k+1}\perp L_{n-k}$, and hence that $L_n=\left(\sum_{l=0}^{k-1}\overrightarrow{\oplus}N_{n-l}\right)\overrightarrow{\oplus}L_{n-k}$. We now note that Propositon 2.1 and its proof are valid with all arrows and orthogonalities reversed, provided we change the estimates on m_n and k_n to read $\|k_n\|\leqq\|m_n+k_n\|=\|x\|$ and $\|m_n\|\leqq\|x-k_n\|\leqq 2\|x\|$. (Also, we may ignore the proof of the basis property of $\Sigma\overrightarrow{\oplus}\mathcal{O}_n$ by our remarks following the definition of left Wold decomposition.) Applying this version of Proposition 2.1, then, we have that (i) holds.

Assume (i). Then we may write for all n and $l\geqq 1$,

$$L_{n+l}=N_{n+l}\overrightarrow{\oplus}N_{n+l-1}\overrightarrow{\oplus}\dots\overrightarrow{\oplus}N_{n+1}\overrightarrow{\oplus}L_n.$$

This means that writing $Y\in L_{n+l}$ (uniquely) as $Y=Z_{n+l}+\dots+Z_{n+1}+Y_n$ with $Z_j\in N_j$ and $Y_n\in L_n$, we have $P_{n,n+l}(Y)=Y_n$. Then

$$\begin{aligned}
&P_{n,n+1}\dots P_{n+k-1,n+k}(Z_{n+k}+\dots+Z_{n+1}+Y_n)\\
&\quad=P_{n,n+1}\dots P_{n+k-2,n+k-1}(Z_{n+k-1}+\dots+Z_{n+1}+Y_n)\\
&\quad\ \vdots\\
&\quad=P_{n,n+1}(Z_{n+1}+Y_n)\\
&\quad=Y_n\\
&\quad=P_{n,n+k}(Z_{n+k}+\dots+Z_{n+1}+Y_n).
\end{aligned}$$

Thus, (iv) holds.

We now show (ii)$\Leftrightarrow$(iii). Assuming first that (ii) holds, we note that for arbitrary $W\in L_{+\infty}$ and $m\leqq n$,

$$P_mP_nW=P_m(W-(W-P_nW))=P_mW-P_m(W-P_nW)=P_mW=P_nP_mW$$

since $P_m(W - P_n W) = 0$ by Propositions 2.4 and 2.5. Hence (ii) implies (iii). Conversely, assume (iii) holds. We show by induction on k that P_n is linear on each L_{n+k}, whence it is linear on $L_{+\infty}$ by continuity. P_n is homogeneous by Proposition 2.6; we show additivity. Clearly, P_n is additive on L_n. Assume it is additive on L_{n+k-1}. Let W_1, W_2 be arbitrary in L_{n+k}, and define $Y_j = P_{n+k-1} W_j$ and $Z_j = W_j - Y_j$. Note $Z_j \perp L_{n+k-1}$. By Proposition 2.7, P_{n+k-1} is a linear operator on L_{n+k}, and this coupled with (iii) and our induction assumption gives

$$\begin{aligned} P_n(W_1 + W_2) &= P_n P_{n+k-1}(Y_1 + Y_2 + Z_1 + Z_2) \\ &= P_n(Y_1 + Y_2) \\ &= P_n Y_1 + P_n Y_2 \\ &= P_n P_{n+k-1}(Y_1 + Z_1) + P_n P_{n+k-1}(Y_2 + Z_2) \\ &= P_n W_1 + P_n W_2. \end{aligned}$$

Thus (iii) implies (ii). The proof is complete. □

Remark. The observant reader will have noticed that we make no use whatsoever of the SαS property of $\{X_n\}$ in Proposition 2.8 and Theorem 2.10. Thus *these results are true for any* p^{th} *order process* $\{X_n\}$ (i.e. $E|X_n|^p < \infty$ for all n) with $p > 1$. Of course, the definitions of innovation and Wold decomposition in this case are with respect to the L^p orthogonality $\perp_p$.

We do not have a characterization of $\{X_n\}$ for which a left non-linear Wold decomposition exists. The method of proof of Theorem 2.10 will not work to prove a non-linear version of that theorem, as it uses the property that L_n is codimension one in L_{n+1}. However, the non-linear analog of the proof that (ii) implies (i) *is* valid, so that *linearity of the metric projections* P_n: $\mathscr{L}_{+\infty} \to \mathscr{L}_n$ *implies that* $\{X_n\}$ *has a left non-linear Wold decomposition.*

Innovations: Right and Left, Nonlinear and Linear

It is of interest to compare the various types of innovations introduced earlier, which unfold the information of a SαS sequence $\{X_n\}$.

The right nonlinear innovations $\{\mathscr{I}_n^r\}$ are given by the residuals of the regression of X_n on the past $X_{n-1}, \ldots$:

$$\mathscr{I}_n^r = X_n - E(X_n | X_{n-1}, \ldots).$$

The right (linear) innovations $\{I_n^r\}$ exist precisely when these regressions are linear (Theorem 2.3) and then equal the right nonlinear innovations, $I_n^r = \mathscr{I}_n^r$.

The nonlinear left innovations $\{\mathscr{I}_n^l\}$ and the (linear) left innovations $\{I_n^l\}$ are given by the nonlinear and the linear prediction errors of X_n from the past $X_{n-1}, \ldots$:

$$\mathscr{I}_n^l = X_n - NL(X_n | X_{n-1}, \ldots),$$
$$I_n^l = X_n - L(X_n | X_{n-1}, \ldots),$$

where $NL(X_n|X_{n-1}, \ldots)$ is the "best" nonlinear and $L(X_n|X_{n-1}, \ldots)$ is the "best" linear predictor of X_n from the past $X_{n-1}, \ldots$, i.e., the element of the nonlinear span of the past $\mathscr{L}_n$ and of the linear span of the past L_n, which is nearest to X_n in L_p-norm $(1<p<\alpha)$.

The nonlinear and linear left innovations coincide, $I_n^l = \mathscr{I}_n^l$, if and only if the best nonlinear and linear predictors of X_n from the past $X_{n-1}, \ldots$ coincide, i.e., if and only if the metric projection of X_n onto L_{n-1} coincides with its metric projection onto $\mathscr{L}_{n-1}$, if and only if there is a $Y_n \in L_{n-1}$ such that

$$E(X_n - Y_n)^{\langle p-1 \rangle} Z = 0 \qquad \text{for all } Z \in \mathscr{L}_{n-1},$$

or equivalently

$$E\{(X_n - Y_n)^{\langle p-1 \rangle} | X_{n-1}, \ldots\} = 0,$$

(in which case of course $Y_n = I_n^l = \mathscr{I}_n^l$).

The nonlinear left and right innovations coincide, $\mathscr{I}_n^r = \mathscr{I}_n^l$, if and only if the regression predictors from the past coincide with the best nonlinear predictors from the past, $E(X_n|X_{n-1}, \ldots) = NL(X_n|X_{n-1}, \ldots)$, if and only if

$$E\{[X_n - E(X_n|X_{n-1}, \ldots)]^{\langle p-1 \rangle} | X_{n-1}, \ldots\} = 0.$$

This condition is a form of weak conditional symmetry, and is clearly satisfied if the conditional law of X_n given $X_{n-1}, \ldots$ is symmetric (since it will then be necessarily symmetric about its conditional mean $E(X_n|X_{n-1}, \ldots)$).

The right linear innovations exist and equal the left linear innovations, $I_n^r = I_n^l$, if and only if the regression predictors from the past coincide with the best linear predictors from the past, $E(X_n|X_{n-1}, \ldots) = L(X_n|X_{n-1}, \ldots)$, if and only if $E(X_n|X_{n-1}, \ldots)$ is linear and

$$E\{[X_n - E(X_n|X_{n-1}, \ldots)]^{\langle p-1 \rangle} X_k\} = 0 \qquad \text{for all } k<n.$$

This is weaker than the previous conditional symmetry condition and is likewise satisfied whenever the conditional law of X_n given $X_{n-1}, \ldots$ is symmetric. Thus symmetry of the conditional laws and linearity of the regressions implies that all types of innovations coincide.

So far we have limited the discussion to one-step ahead prediction. But the Wold decompositions and innovations introduced here are precisely tailored to handle the general m-step ahead prediction; and indeed any estimation problem based on observations of the past of X. To simplify the notation we will write the expression of the m-step predictors and their errors in terms of innovations only in the *stationary* case. Let

$$X_n = \mathscr{Y}_n^r + \sum_{k=0}^{\infty} a_k^r \Xi_{n-k}^r$$

be the right nonlinear Wold decomposition of X, which always exists. Then $\mathcal{I}_n^r = a_0^r \Xi_n$, and thus it can be written in the form

$$X_n = \mathcal{Y}_n^r + \sum_{k=0}^{\infty} \frac{a_k^r}{a_0^r} \mathcal{I}_{n-k}^r.$$

It then follows (as in the proof of Theorem 2.2) that the m-steps ahead ($m \geqq 1$) regression predictor is

$$E(X_n | X_{n-m}, \ldots) = \mathcal{Y}_n^r + \sum_{k=m}^{\infty} \frac{a_k^r}{a_0^r} \mathcal{I}_{n-k}^r$$

and the regression prediction error is

$$X_n - E(X_n | X_{n-m}, \ldots) = \sum_{k=0}^{m-1} \frac{a_k^r}{a_0^r} \mathcal{I}_{n-k}^r.$$

If a linear right Wold decomposition exists (cf. Theorem 2.3), then one simply replaces $\mathcal{Y}$ by Y and $\mathcal{I}$ by I.

Now assume a left linear Wold decomposition exists (cf. Theorem 2.10). Then we obtain likewise

$$X_n = Y_n^l + \sum_{k=0}^{\infty} \frac{a_k^l}{a_0^l} I_{n-k}^l$$

from which it follows that the m-step ahead best linear predictor is

$$L(X_n | X_{n-m}, \ldots) = Y_n^l + \sum_{k=m}^{\infty} \frac{a_k^l}{a_0^l} I_{n-k}^l$$

and the linear prediction error is

$$X_n - L(X_n | X_{n-m}, \ldots) = \sum_{k=0}^{m-1} \frac{a_k^l}{a_0^l} I_{n-k}^l.$$

3. Independent Decomposition of Stable Sequences

As we observed following Corollary 1.3, independence implies two-sided orthogonality for SαS random variables, but not conversely. Thus we should not, in general, expect as in the Gaussian case that the innovation subspaces in a Wold decomposition are independent. In this section, we study those processes for which this is the case.

Using the notation of Sect. 2, we say that a SαS sequence $\{X_n\}$ has *independent innovations* if for each n we can find a subspace N_n so that $L_n = L_{n-1} + N_n$, with L_{n-1} and N_n independent. To symbolize this we write $L_n = L_{n-1} \oplus N_n$. We say that $\{X_n\}$ has an *independent Wold decomposition* if there exist subspaces $\{N_k\}$ so that for each n, $L_n = \sum_{k=0}^{\infty} N_{n-k} + L_{-\infty}$ where $\{L_{-\infty}, N_k\colon k \in \mathbb{Z}\}$ are mutually independent $\left(\text{in symbols, } L_n = \left(\sum_{k=0}^{\infty} \oplus N_{n-k}\right) \oplus L_{-\infty}\right)$. If $\{X_n\}$ has an independent Wold decomposition then clearly it has both right and left Wold decompositions and all three coincide.

The independent Wold decomposition for stochastic processes with infinite variance was studied by Urbanik [18–20] for strictly stationary processes "admitting prediction", and by Thu [17] for random fields. Here we give spectral necessary and sufficient conditions for the existence of such a decomposition for SαS sequences.

Theorem 3.1. *Let* $1<\alpha<2$ *and let* $\{X_n\}$ *be a SαS sequence, represented by* $\{f_n\}$. *The following are equivalent.*

(i) $\{X_n\}$ *has independent innovations.*

(ii) $\{X_n\}$ *has an independent Wold decomposition.*

(iii) *For all* n, $f_n = g_n + h_n$, *where* $g_n \in \overline{sp}\{f_k\colon k \leqq n-1\}_{L^\alpha}$ *and* $f_k \cdot h_n = 0$ *a.e. for* $k \leqq n-1$.

(iv) *For all* n, $f_n = \sum_{k=0}^{\infty} a_{k,n} \phi_{n-k} + \psi_n$, *where* $\psi_n \in \bigcap_m \overline{sp}\{f_k\colon k \leqq m\}$, $\overline{sp}\{f_k\colon k \leqq n\} = \bigcap_m \overline{sp}\{f_k\colon k \leqq m\} + \overline{sp}\{\phi_k\colon k \leqq n\}$, $\psi_k \cdot \phi_l = 0$ *a.e. for all* k, l, *and* $\phi_k \cdot \phi_l = 0$ *a.e. for all* $k \neq l$.

Proof. We show first that (i) is equivalent to (ii). Assume (i) holds, and observe that for fixed n we may write

$$\begin{aligned} L_n &= L_{n-1} \bar{\oplus} N_n \\ &= L_{n-2} \bar{\oplus} N_{n-1} \bar{\oplus} N_n \\ &\ \ \vdots \\ &= L_{n-k-1} \bar{\oplus} N_{n-k} \bar{\oplus} \dots \bar{\oplus} N_n. \end{aligned}$$

Choosing $1<p<\alpha$, and applying Proposition 2.1 (remembering that independence implies orthogonality), we get that $L_n = \left(\sum_{k=0}^{\infty} \underset{\leftarrow}{\oplus} N_{n-k}\right) \underset{\leftarrow}{\oplus} L_{-\infty}$ and that each $Z \in \sum_{k=0}^{\infty} \underset{\leftarrow}{\oplus} N_{n-k}$ has the appropriate unique expansion. Since the spaces $L_{n-k-1}, N_{n-k}, \dots, N_n$ are mutually independent by the construction above, the mutual independence of $\{L_{-\infty}, N_k\colon k \in \mathbb{Z}\}$ follows. So (i) implies (ii). Also, (ii) implies (i) by definition.

We now deal with the spectral conditions (iii) and (iv). Recall Schilder's result that SαS variables are independent if and only if their spectral representations have almost disjoint support.

Assume (i). We may then write $X_n = Y_n + Z_n$ with $Y_n \in L_{n-1}$, and Z_n independent of L_{n-1}. Denoting by $\{g_n, h_n\}$ the representatives of $\{Y_n, Z_n\}$, we see (iii) holds. Conversely, if (iii) holds, we let Z_n be the random variable in L_n which is represented by h_n, and let $N_n = sp\{Z_n\}$. N_n is independent of L_{n-1} since $f_k \cdot h_n = 0$ for $k \leqq n-1$. It is clear that $L_n = L_{n-1} + N_n$, so (i) holds. This shows (i) is equivalent to (iii).

Assume (ii). If $\dim(N_j) \neq 0$, choose a non-zero $W_j \in N_j$. Otherwise, let $W_j = 0$. Let $\{\phi_j\}$ be the representatives of $\{W_j\}$. By hypothesis, X_n has an independent expansion $X_n = Y_n + \sum_{k=0}^{\infty} a_{k,n} W_{n-k}$, where $Y_n \in L_{-\infty}$. Letting $\{\psi_j\}$ represent $\{Y_j\}$, we have $f_n = \psi_n + \sum_{k=0}^{\infty} a_{k,n} \phi_{n-k}$, with $\psi_n \in \bigcap_m \overline{sp}\{f_k\colon\ k \leqq m\}$. The relation $L_n = L_{-\infty} \bar{\oplus} \sum_{k=0}^{\infty} \bar{\oplus} N_{n-k}$ translates in representation space to the remaining statements in (iv). Hence (ii) implies (iv). That (iv) implies (ii) is easily seen, since (iv) implies (iii) with $g_n \triangleq \psi_n + \sum_{k=1}^{\infty} a_{k,n} \phi_{n-k}$ and $h_n \triangleq a_{0,n} \phi_n$. Therefore (ii) is equivalent to (iv) and the theorem is proved. □

In the stationary case, this result takes on the following form, where we assume for simplicity that $\{X_n\}$ is completely non-deterministic, i.e., $L_{-\infty} = \{0\}$.

Theorem 3.2. *Let $1 < \alpha < 2$, and let $\{X_n\}$ be a $S\alpha S$ sequence represented by $\{f_n\}$. Then $\{X_n\}$ is stationary, completely non-deterministic, and has an independent Wold decomposition if and only if* (iv) *of Theorem 3.1 holds with $\psi_n = 0$, and with $a_{k,n}$ and $\|\phi_n\|_\alpha$ independent of n.*

Proof. Assume first that $\{X_n\}$ is stationary, completely non-deterministic, and has an independent Wold decomposition. Let S be the canonical shift of $\{X_n\}$, i.e., S is the isometric linear extension of the map $SX_n = X_{n-1}$ on $L^p(\Omega, \Sigma, P)$. Since S preserves joint distributions

$$
\begin{aligned}
L_{n-1} \bar{\oplus} N_n = L_n = SL_{n+1} &= S(L_n \bar{\oplus} N_{n+1}) = SL_n \bar{\oplus} SN_{n+1} \\
&= L_{n-1} \bar{\oplus} SN_{n+1}.
\end{aligned}
$$

This implies $SN_{n+1} = N_n$. Choosing a non-zero $W_0 \in N_0$ and defining $W_k = S^{-k} W_0$ we see that $\{W_k\}$ is an i.i.d. sequence with $W_k \in N_k$. By our assumption, then, we may find $a_{k,n}$ so that $X_n = \sum_{k=0}^{\infty} a_{k,n} W_{n-k}$ for each n. Note that

$$
\sum_{k=0}^{\infty} a_{k,n} W_{n-k} = X_n = SX_{n+1} = \sum_{k=0}^{\infty} a_{k,n+1} SW_{n+1-k} = \sum_{k=0}^{\infty} a_{k,n+1} W_{n-k},
$$

whence $a_k \triangleq a_{k,n}$ does not depend on n.

Letting $\{\phi_j\}$ be the representatives of $\{W_j\}$, we have that $f_n = \sum_{k=0}^{\infty} a_k \phi_{n-k}$. That $\|\phi_j\|_\alpha$ is constant in j follows from the fact that $\{W_j\}$ is identically distributed, and that $\phi_j \cdot \phi_k = 0$ a.e. for $k \neq j$ follows from the independence of $\{W_j\}$. $\overline{sp}\{f_k: k \leqq n\} = \overline{sp}\{\phi_k: k \leqq n\}$ since $L_n = \overline{sp}\{W_k: k \leqq n\}$, and the first implication is proved.

For the reverse implication, let $W_j \in L_j$ be the SαS random variable represented by ϕ_j, and let $N_j = sp\{W_j\}$. Clearly, $\{W_j\}$ is i.i.d. and $X_n = \sum_{k=0}^{\infty} a_k W_{n-k}$. This moving average is stationary and completely non-deterministic. $L_n = \overline{sp}\{W_k: k \leqq n\}$ since $\overline{sp}\{f_k: k \leqq n\} = \overline{sp}\{\phi_k: k \leqq n\}$, proving the theorem. □

4. Examples

We present here some examples of SαS processes having or not having various of the decompositions discussed in previous sections. They are intended to illustrate the theorems we have proved (although they do not exhaustively do so), and more importantly, to provide some feeling for what is and what is not possible regarding these decompositions. We should note at the outset that in the Gaussian case $\alpha = 2$, all aforementioned decompositions exist and coincide; and the situation for $\alpha < 2$ should be compared with this.

Example 4.1. Certain autoregressive and moving average SαS processes have Wold decompositions. Specifically, let $\{\xi_n\}$ be a sequence of i.i.d. SαS variables. If for all n, $\{X_n\}$ satisfies either

(i) $X_{n+1} = \sum_{k=0}^{K} \lambda_k X_{n-k} + \xi_{n+1}$ with ξ_n independent of L_{n-1}, or

(ii) $X_n = \sum_{k=0}^{K} \mu_k \xi_{n-k}$ with $\xi_n \in L_n$,

then $\{X_n\}$ has an independent Wold decomposition.

Proof. In the case (i), it is clear that $\{X_n\}$ has independent innovations, and so by Theorem 3.1 has an independent Wold decomposition. The existence of the decomposition in case (ii) follows by definition (with $N_n = sp\{\xi_n\}$).

Of course, left, right, or two-sided decompositions exist for such $\{X_n\}$ when the appropriate hypotheses of left, right, or two-sided orthogonality of $\{\xi_n\}$ are assumed. □

If $\xi_n \notin L_n$, however, a moving average as in 4.1 (ii) may not have a Wold decomposition, as the following example shows.

Example 4.2. There exists a stationary SαS moving average that has a left Wold decomposition, yet does not have a right or independent Wold decomposition. Specifically, let $\{\xi_n\}_{n=-\infty}^{\infty}$ be an i.i.d. sequence of SαS random variables, $1 < \alpha < 2$. Set $X_n = \xi_n - 2\xi_{n-1}$. Then $\{X_n\}$ does not have a right (linear) Wold decomposition, yet does have a left (linear) Wold decomposition.

Proof. To show $\{X_n\}$ does not have a right Wold decomposition, we proceed as follows. Assume that $\{X_n\}$ does have a right Wold decomposition, in which case we have $E(X_{n+1}|X_n, X_{n-1}, \dots) \in L_n$ by Theorem 2.3. We show that $\{X_k: k \leqq n\}$ forms a basis for its span, whereby we may write $E(X_{n+1}|X_n, X_{n-1}, \dots) = \sum_{k=0}^{\infty} \lambda_k X_{n-k}$ for some $\{\lambda_k\}$. We then determine, using orthogonality, the only possible choice for the sequence $\{\lambda_k\}$, and show that all necessary orthogonality relations do not hold with this choice, completing the first half of the proof.

To show that $\{X_k: k \leqq n\}$ forms a basis for its span, it suffices to show there exists K such that

$$\left\|\sum_{j=0}^{M} \beta_j X_{n-j}\right\|_p \leqq K \left\|\sum_{j=0}^{N} \beta_j X_{n-j}\right\|_p$$

for all β_j and all $M<N$. Recall that for any $1<p<\alpha$ there is a constant $c=c(p,\alpha)$ such that for all SαS variables X with representative f, $\|X\|_p = c\|f\|_\alpha$. Note also that we may represent $\{\xi_n\}$ by $\{1_{[n,n+1]}\}$ on $L^\alpha(\mathbb{R})$. Hence

$$\begin{aligned}
\left\|\sum_{j=0}^{L} \beta_j X_{n-j}\right\|_p^\alpha &= \left\|\sum_{j=0}^{L} \beta_j(\xi_{n-j} - 2\xi_{n-j-1})\right\|_p^\alpha \\
&= \left\|\beta_0 \xi_n + \sum_{j=1}^{L} (\beta_j - 2\beta_{j-1})\xi_{n-j} - 2\beta_L \xi_{n-L-1}\right\|_p^\alpha \\
&= c^\alpha \left[|\beta_0|^\alpha + \sum_{j=1}^{L} |\beta_j - 2\beta_{j-1}|^\alpha + |2\beta_L|^\alpha\right] \\
&\triangleq c^\alpha \cdot S_L.
\end{aligned}$$

It thus suffices to find K such that $\frac{S_N}{S_M} \geqq K^{-\alpha}$. We claim that $K=2$ will satisfy this requirement. To see this, call $\beta = |\beta_0|^\alpha + \sum_{j=1}^{M} |\beta_j - 2\beta_{j-1}|^\alpha$, and $\gamma_k = \frac{\beta_{M+k}}{\beta_M}$. Then

$$\begin{aligned}
\frac{S_N}{S_M} &= \frac{\beta + \sum_{j=M+1}^{N} |\beta_j - 2\beta_{j-1}|^\alpha + |2\beta_N|^\alpha}{\beta + |2\beta_M|^\alpha} \\
&\geqq \min\left(1, \frac{\sum_{j=M+1}^{N} |\beta_j - 2\beta_{j-1}|^\alpha + |2\beta_N|^\alpha}{|2\beta_M|^\alpha}\right) \\
&= \min\left(1, \sum_{j=1}^{N-M} |\gamma_{j-1} - \gamma_j/2|^\alpha + |\gamma_{N-M}|^\alpha\right).
\end{aligned}$$

Putting $n=N-M$, we have

$$\sum_{j=1}^{N-M} |\gamma_{j-1}-\gamma_j/2|^\alpha+|\gamma_{N-M}|^\alpha$$
$$=|1-\gamma_1/2|^\alpha+|\gamma_1-\gamma_2/2|^\alpha+\ldots+|\gamma_{n-1}-\gamma_n/2|^\alpha+|\gamma_n|^\alpha.$$

We may verify this is $\geqq 2^{-\alpha}$ as follows. If not, then all terms must be less than $2^{-\alpha}$, i.e., $|1-\gamma_1/2|<\frac{1}{2}$, $|\gamma_1-\gamma_2/2|<\frac{1}{2}$, etc. But $|1-\gamma_1/2|<\frac{1}{2}$ implies $\gamma_1>1$, and $|\gamma_1-\gamma_2/2|<\frac{1}{2}$ with $\gamma_1>1$ implies $\gamma_2>1$, and so on until we reach $\gamma_n>1$ in which case the last term is not less than $2^{-\alpha}$. We now have that $\{X_k\colon k\leqq n\}$ forms a basis for L_n.

Under the assumption that X_n does in fact have a right Wold decomposition, we may write $E(X_{n+1}|X_n, X_{n-1}, \ldots)=\sum_{k=0}^{\infty}\lambda_k X_{n-k}$ for some choice of $\{\lambda_k\}$. Also, the λ_k must satisfy $X_{n-j}\perp X_{n+1}-\sum_{k=0}^{\infty}\lambda_k X_{n-k}$ for all $j\geqq 0$ by Proposition 1.5. This requirement is equivalent, by Lemmata 1.1 and 1.2, to

$$\begin{aligned}0&=EX_{n-j}^{\langle p-1\rangle}\left(X_{n+1}-\sum_{k=0}^{\infty}\lambda_k X_{n-k}\right)\\
&=E(\xi_{n-j}-2\xi_{n-j-1})^{\langle p-1\rangle}\left(\xi_{n+1}-2\xi_n-\sum_{k=0}^{\infty}\lambda_k(\xi_{n-k}-2\xi_{n-k-1})\right)\\
&=c^p(1+2^\alpha)^{p/\alpha-1}\cdot\begin{cases}-2-(1+2^\alpha)\lambda_0+2^{\alpha-1}\lambda_1, & j=0,\\ 2\lambda_{j-1}-(1+2^\alpha)\lambda_j+2^{\alpha-1}\lambda_{j+1}, & j>0.\end{cases}\end{aligned}$$

Thus λ_k must satisfy

$$\begin{aligned}\lambda_1&=2^{2-\alpha}+2(1+2^{-\alpha})\lambda_0,\\ \lambda_{k+1}&=2(1+2^{-\alpha})\lambda_k-2^{2-\alpha}\lambda_{k-1}, \qquad k>0.\end{aligned}$$

A solution to these equations is determined by specifying λ_0. The solution for $k\geqq 0$ is

$$\lambda_k=2^k(1-2^{-\alpha})^{-1}[2^{1-\alpha}(1-2^{-\alpha k})+(1-2^{-\alpha(k+1)})\lambda_0].$$

It is easily seen that $\lim_{k\to\infty}|\lambda_k|=\infty$ unless $\lambda_0=-2^{1-\alpha}$. Hence we must have that

$$\lambda_k=-(2^{1-\alpha})^{k+1}$$

and furthermore that

$$X_{n+1}-\sum_{k=0}^{\infty}\lambda_k X_{n-k}=\xi_{n+1}-2(1-2^{-\alpha})\sum_{k=0}^{\infty}(2^{1-\alpha})^k\xi_{n-k}.$$

To obtain our contradiction, recall that *all* of L_n (not just each X_{n-j}) must be orthogonal to $X_{n+1}-\sum_{k=0}^{\infty}\lambda_k X_{n-k}$. Check that for $j>0$,

$$\begin{aligned}
&E(X_{n-j}+X_{n-j-1})^{\langle p-1\rangle}\left(X_{n+1}-\sum_{k=0}^{\infty}\lambda_k X_{n-k}\right)\\
&\quad=E(\xi_{n-j}-\xi_{n-j-1}-2\xi_{n-j-2})^{\langle p-1\rangle}\left(\xi_{n+1}-2(1-2^{-\alpha})\sum_{k=0}^{\infty}(2^{1-\alpha})^k\xi_{n-k}\right)\\
&\quad=(\text{const.}\neq 0)\,[1-2^{1-\alpha}-2^{\alpha-1}(2^{1-\alpha})^2]\\
&\quad=(\text{const.}\neq 0)\,[1-2^{2-\alpha}],
\end{aligned}$$

which is non-zero for $\alpha<2$, completing the proof that no right Wold decomposition exists.

To show $\{X_n\}$ has a left Wold decomposition, it suffices by Theorem 2.10 to show that the operators P_n are linear. Clearly these operators are linear if and only if they are linear on each L_M, $M<\infty$. By our arguments above, $\{X_n\}$ is a basic set. It is thus a simple matter to show, in view of Proposition 2.6, that P_n is linear on L_M if and only if

$$P_n\left(\sum_{k=n+1}^{M} a_k X_k\right)=\sum_{k=n+1}^{M} a_k P_n X_k. \qquad (\dagger)$$

Since X_k is by definition independent of (and thus orthogonal to) L_n for $k\geqq n+2$, the RHS of (†) is just $a_{n+1}P_n X_{n+1}$, or $P_n(a_{n+1}X_{n+1})$, by Propositions 2.5 and 2.6. Recall (Proposition 2.4) that $P_n(a_{n+1}X_{n+1})$ is the unique $Y\in L_n$ satisfying

$$E(a_{n+1}X_{n+1}-Y)^{\langle p-1\rangle}X_l=0 \qquad \text{for } l\leqq n.$$

The LHS of (†) is likewise the unique $Y'\in L_n$ satisfying

$$E\left(\sum_{k=n+1}^{M} a_k X_k-Y'\right)^{\langle p-1\rangle}X_l=0 \qquad \text{for } l\leqq n.$$

Now represent $\{Y', X_n\colon -\infty<n<\infty\}$ by $\{g', f_n\colon -\infty<n<\infty\}$ and recall that independence of X_k and X_l for $|k-l|\geqq 2$ is equivalent to f_k and f_l having almost disjoint support for like indices. Thus for $l\leqq n$,

$$\begin{aligned}
0&=E\left(\sum_{k=n+1}^{M} a_k X_k-Y'\right)^{\langle p-1\rangle}X_l\\
&=(\text{const.}\neq 0)\int\left(\sum_{k=n+1}^{M} a_k f_k-g'\right)^{\langle\alpha-1\rangle}f_l\,dm\\
&=(\text{const.}\neq 0)\int(a_{n+1}f_{n+1}-g')^{\langle\alpha-1\rangle}f_l\,dm\\
&=(\text{const.}\neq 0)\,E(a_{n+1}X_{n+1}-Y')^{\langle p-1\rangle}X_l.
\end{aligned}$$

Hence $Y=Y'$ (i.e. (†) holds) and $\{X_n\}$ has a left Wold decomposition.

For the sake of completeness, we also compute this left Wold decomposition. Since $\{X_n\}$ is basic, we must have that $P_n X_{n+1}=\sum_{k=0}^{\infty}\lambda_k X_{n-k}$ for some choice of $\{\lambda_k\}$. Analogous to what was done for the right Wold decomposition for this process, we may use the orthogonality relations $X_{n+1}-\sum_{k=0}^{\infty}\lambda_k X_{n-k}\perp X_l$ for $l\leqq n$ to derive equations which $\{\lambda_k\}$ must satisfy. We omit the details, and state only that the analogous arguments show that

$$\lambda_k=-\left(2^{\frac{1}{1-\alpha}}\right)^{k+1}$$

provides the unique solution to these equations for which $\sum\lambda_k X_{n-k}$ converges. Hence the left Wold decomposition for this process is given by: $L_{-\infty}=\{0\}$; $N_n=sp\{I_n^l\}$ where $I_n^l=\sum_{k=0}^{\infty}(2^{\frac{1}{1-\alpha}})^k X_{n-k}$; and $X_n=I_n^l-2^{\frac{1}{1-\alpha}}I_{n-1}^l$. □

Example 4.3. All sub-Gaussian sequences have identical right and left Wold decompositions, yet never have independent Wold decompositions.

Proof. Any sub-Gaussian process $\{X_n\}$ may be represented as $X_n=A^{\frac{1}{2}}G_n$, where $\{G_n\}$ is a mean-zero Gaussian process and A is a positive $\alpha/2$-stable variable, independent of $\{G_n\}$. Let $L_n'=\overline{sp}\{G_k\colon k\leqq n\}$, and let $L_n'=L_{-\infty}'\bar{\oplus}\sum_{k=0}^{\infty}\bar{\oplus}N_{n-k}'$ be the standard (independent) Wold decomposition of $\{G_n\}$. Then $L_n=A^{\frac{1}{2}}L_n'$, $L_{-\infty}=A^{\frac{1}{2}}L_{-\infty}'$; and letting $N_k=A^{\frac{1}{2}}N_k'$, we have the decomposition $L_n=L_{-\infty}+\sum_{k=0}^{\infty}N_{n-k}$. That this decomposition possesses the appropriate orthogonalities follows from the fact that if A, Z_1, and Z_2 are independent with Z_1, Z_2 mean-zero Gaussian, then $A^{\frac{1}{2}}Z_1$ and $A^{\frac{1}{2}}Z_2$ are two-sided orthogonal:

$$E(A^{\frac{1}{2}}Z_1)^{\langle p-1\rangle}(A^{\frac{1}{2}}Z_2)=EA^{p/2}Z_1^{\langle p-1\rangle}Z_2=EA^{p/2}\cdot EZ_1^{\langle p-1\rangle}EZ_2=0.$$

The decomposition cannot be independent, since L_n contains no non-trivial independent random variables (cf. Lemma 2.1 in [1]). □

Example 4.4. Let $\{\xi_n\}$ be i.i.d. SαS, $1<\alpha<2$. Let $0<|\lambda|<1$ and define $X_n=\sum_{k=0}^{\infty}\lambda^k\xi_{n+k}$. *Then $\{X_n\}$ has a right Wold decomposition, but has no independent or left Wold decomposition.*

Proof. Let $\mu=\lambda^{\langle\alpha-1\rangle}$ and define $Z_n=X_n-\mu X_{n-1}$. Since $X_n=\sum_{k=0}^{\infty}\mu^k Z_{n-k}$, we have $L_n=\overline{sp}\{X_k\colon k\leqq n\}=\overline{sp}\{Z_k\colon k\leqq n\}$. We claim that $\{Z_n\}$ is not an independent sequence, yet $\{X_n\}$ has a right Wold decomposition $L_n=\sum_{k=0}^{\infty}\overleftarrow{\oplus} N_{n-k}$, where $N_j=sp\{Z_j\}$.

Note $Z_n=-\lambda^{\langle\alpha-1\rangle}\xi_{n-1}+(1-|\lambda|^\alpha)\sum_{k=0}^{\infty}\lambda^k\xi_{n+k}$. Represent $\{\xi_j\}$ by $\{I_j\}$, $I_j\triangleq 1_{[j,j+1]}$, so that $\{Z_n\}$ is represented by $\{f_n\}$, where

$$f_n=-\lambda^{\langle\alpha-1\rangle}I_{n-1}+(1-|\lambda|^\alpha)\sum_{k=0}^{\infty}\lambda^k I_{n+k}.$$

Note also that since $\xi_j=X_j-\lambda X_{j+1}$, we have $L_n=\overline{sp}\{\dots,\xi_{n-2},\xi_{n-1},X_n\}$. The following calculation shows $a_0X_n+\sum_{k=1}^{N}a_k\xi_{n-k}\perp Z_{n+1}$ for any choice of a_j and hence that $L_n\perp Z_{n+1}$:

$$\begin{aligned}
&\int_{\mathbb{R}}\left[a_0\left(\sum_{j=0}^{\infty}\lambda^j I_{n+j}\right)+\sum_{k=1}^{N}a_k I_{n-k}\right]^{\langle\alpha-1\rangle} f_{n+1}\,dm\\
&=a_0^{\langle\alpha-1\rangle}\int_{\mathbb{R}}\left[\sum_{j=0}^{\infty}(\lambda^j)^{\langle\alpha-1\rangle}I_{n+j}\right]\left[-\lambda^{\langle\alpha-1\rangle}I_n+(1-|\lambda|^\alpha)\sum_{k=0}^{\infty}\lambda^k I_{n+k+1}\right]dm\\
&=a_0^{\langle\alpha-1\rangle}\left[-\lambda^{\langle\alpha-1\rangle}+(1-|\lambda|^\alpha)\sum_{k=0}^{\infty}\lambda^k(\lambda^{k+1})^{\langle\alpha-1\rangle}\right]\\
&=a_0^{\langle\alpha-1\rangle}\left[-\lambda^{\langle\alpha-1\rangle}+(1-|\lambda|^\alpha)\lambda^{-1}\sum_{k=0}^{\infty}|\lambda|^{\alpha(k+1)}\right]\\
&=0.
\end{aligned}$$

Now observe that

$$E(X_{n+1}|X_n,X_{n-1},\dots)=E(Z_{n+1}+\lambda^{\langle\alpha-1\rangle}X_n|X_n,X_{n-1},\dots)=\lambda^{\langle\alpha-1\rangle}X_n.$$

Hence $\{X_n\}$ has a right Wold decomposition by Theorem 2.3. However, the spaces N_k are *not* independent, since $f_k\cdot f_l\neq 0$ a.e. (It is also clear that $\{X_n\}$ is not a sub-Gaussian process, since L_n contains independent random variables.)

We now wish to show $\{X_n\}$ has no left Wold decomposition. We do this by showing that condition (iii) of Theorem 2.10 is violated. To this end, let P_n be the metric projection onto L_n. We show that there are constants b_1 and b_2 such that $P_nX_{n+1}=b_1X_n$ and $P_nX_{n+2}=b_2X_n$ yet

$$P_{n+1}P_nX_{n+2}=P_nX_{n+2}=b_2X_n\neq b_1^2X_n=P_nP_{n+1}X_{n+2},$$

showing P_n does not commute with P_{n+1}.

Let $Y_j=P_nX_{n+j}$, $j=1,2$. Then necessarily $Y_1=b_1X_n+\sum_{k=1}^{\infty}a_k\xi_{n-k}$, since $L_n=\overline{sp}\{X_n,\xi_{n-1},\xi_{n-2},\dots\}$ and $\{X_n,\xi_{n-1},\xi_{n-2},\dots\}$ is a basic set. By Proposition

2.4, Y_1 must satisfy $X_{n+1}-Y_1 \perp L_n$. The requirement $X_{n+1}-Y_1 \perp \xi_j$, $j \leqq n-1$, implies

$$0=E\left(X_{n+1}-b_1 X_n-\sum_{k=1}^{\infty} a_k \xi_{n-k}\right)^{\langle p-1\rangle} \xi_j$$

which in turn implies $a_j=0$ for all j, and $Y_1=b_1 X_n$. To find b_1, note that $X_n - b_1 X_n = -b_1 \xi_n+(\lambda^{-1}-b_1)\sum_{k=1}^{\infty} \lambda^k \xi_{n+k}$ and compute

$$\begin{aligned} 0&=E(X_{n+1}-b_1 X_n)^{\langle p-1\rangle} X_n \\ &=(\text{const.} \neq 0)\left[-b_1^{\langle \alpha-1\rangle}+(\lambda^{-1}-b_1)^{\langle \alpha-1\rangle} \sum_{k=1}^{\infty} (\lambda^k)^{\langle \alpha-1\rangle} \lambda^k\right] \\ &=(\text{const.} \neq 0)\left[-b_1^{\langle \alpha-1\rangle}+(\lambda^{-1}-b_1)^{\langle \alpha-1\rangle} |\lambda|^\alpha (1-|\lambda|^\alpha)^{-1}\right]. \end{aligned}$$

Solve for b_1 to get

$$b_1=\frac{1}{\lambda} \cdot \frac{(|\lambda|^\alpha)^q}{(1-|\lambda|^\alpha)^q+(|\lambda|^\alpha)^q}, \qquad q \triangleq \frac{1}{\alpha-1}.$$

Using the same methods, we find that $Y_2=b_2 X_n$ where

$$b_2=\frac{1}{\lambda^2} \frac{(|\lambda|^{2\alpha})^q}{(1-|\lambda|^{2\alpha})^q+(|\lambda|^{2\alpha})^q}.$$

Now, $b_2=b_1^2$ if and only if

$$(1-|\lambda|^{2\alpha})^q+(|\lambda|^{2\alpha})^q=[(1-|\lambda|^\alpha)^q+(|\lambda|^\alpha)^q]^2,$$

if and only if

$$(1+|\lambda|^\alpha)^q=(1-|\lambda|^\alpha)^q+2(|\lambda|^\alpha)^q.$$

Since $1<\alpha<2$ and $0<|\lambda|<1$, we have that $q>1$ and

$$\begin{aligned} (1-|\lambda|^\alpha)^q+2(|\lambda|^\alpha)^q &\leqq [1-|\lambda|^\alpha+2^{1/q}|\lambda|^\alpha]^q \\ &<[1+|\lambda|^\alpha]^q. \end{aligned}$$

This shows $b_2 \neq b_1^2$ and hence that $\{X_n\}$ cannot have a left Wold decomposition. □

Example 4.5. The stationary sequence $X_n=\int_{-\pi}^{\pi} e^{in\lambda} dZ(\lambda)$ *is orthogonal but has no right or left or independent Wold decomposition.*

Proof. Since for $m \neq n$, $\int_{-\pi}^{\pi} (e^{im\lambda})^{\langle \alpha-1\rangle} e^{in\lambda} d\lambda=\int_{-\pi}^{\pi} e^{-im\lambda} e^{in\lambda} d\lambda=0$, it follows that $X_m \perp X_n$.

We show that $sp\{X_{n-2}, X_{n-1}\}$ is not orthogonal to X_n, i.e., that $\int_{-\pi}^{\pi} (a + b\,e^{i\lambda})^{\langle\alpha-1\rangle} e^{i2\lambda}\, d\lambda$ does not equal zero for all a and b. Taking $a=b$, we have

$$I \triangleq \int_{-\pi}^{\pi} (1+e^{i\lambda})^{\langle\alpha-1\rangle} e^{i2\lambda}\, d\lambda = \int_{-\pi}^{\pi} \frac{(1+e^{-i\lambda})\, e^{i2\lambda}}{|1+e^{i\lambda}|^{2-\alpha}}\, d\lambda$$

$$= 2^{\alpha/2} \int_0^{\pi} \frac{\cos 2\lambda + \cos\lambda}{(1+\cos\lambda)^{1-\alpha/2}}\, d\lambda. \qquad (*)$$

The numerator of the integrand vanishes at θ with $\cos\theta = \frac{1}{2}$, is positive on $[0,\theta)$ and negative on (θ,π), and of course $\int_0^{\pi} (\cos 2\lambda + \cos\lambda)\, d\lambda = 0$. We thus have

$$I = 2^{\alpha/2} \left\{ \int_0^{\theta} + \int_{\theta}^{\pi} \right\} \frac{\cos 2\lambda + \cos\lambda}{(1+\cos\lambda)^{1-\alpha/2}}\, d\lambda$$

$$< \frac{2^{\alpha/2}}{(1+\cos\theta)^{1-\alpha/2}} \left\{ \int_0^{\theta} + \int_{\theta}^{\pi} \right\} (\cos 2\lambda + \cos\lambda)\, d\lambda$$

$$= 0. \qquad (**)$$

Assume that $\{X_n\}$ has right innovations, i.e., $L_n = L_{n-1} \underset{\rightarrow}{\oplus} N_n$. Then $X_n = Y_n + Z_n$ where $Y_n \in L_{n-1}$ and $L_{n-1} \perp Z_n = X_n - Y_n$. A straightforward adaptation of Theorem 7.1 of [11] shows that $\{X_k,\ k \leqq n\}$ forms a basis in L_n (see [1, p. 606]) so that $Y_n = \sum_{k \leqq n-1} a_k X_k$, the series converging in every L_p, $p < \alpha$. Then $X_k \perp X_n - Y_n$, $k \leqq n-1$, implies $a_k = 0$, i.e. $Y_n = 0$, so that $L_{n-1} \perp X_n$ and $sp\{X_{n-2}, X_{n-1}\} \perp X_n$ contradicting our earlier result. Thus $\{X_n\}$ has no right innovations and no right Wold decomposition.

The orthogonality of the X_n's implies $X_n \perp L_{n-1}$ and thus the best approximation to X_n in L_k, $k \leqq n-1$, is the zero element. It follows that the left innovations space $L_n = L_{n-1} \underset{\leftarrow}{\oplus} N_n$ is $N_n = sp\{X_n\}$. However $N_{n-1} \oplus N_n$ is not orthogonal to L_{n-2}, so no left Wold decomposition exists. This is so because $X_n + X_{n-1}$ is not orthogonal to X_{n-2} as

$$\int_{-\pi}^{\pi} (e^{in\lambda} + e^{i(n-1)\lambda})^{\langle\alpha-1\rangle} e^{i(n-2)\lambda}\, d\lambda = \int_{-\pi}^{\pi} (1+e^{i\lambda})^{\langle\alpha-1\rangle} e^{-i2\lambda}\, d\lambda$$

$$= I < 0$$

from $(*)$ and $(**)$.

That $\{X_n\}$ has no independent Wold decomposition follows from the above, but also follows immediately from part (iii) of Theorem 3.1 and the fact that each $f_n(\lambda) = e^{in\lambda}$ has as support the entire interval $[-\pi, \pi]$. □

References

1. Cambanis, S., Soltani, A.R.: Prediction of stable processes: Spectral and moving average representations. Z. Wahrscheinlichkeitstheor. Verw. Geb. **66**, 593–612 (1984)
2. Campbell, S., Faulkner, G., Sine R.: Isometries, projections, and Wold decompositions. In: I. Endelyi (ed.) Operator theory and functional analysis. San Francisco: Pitman 1979
3. Cline, D.B., Brockwell, P.J.: Linear prediction of ARMA processes with infinite variance. Stochastic Processes Appl. **19**, 281–296 (1985)
4. Faulkner, G.D., Huneycutt, J.E., Jr.: Orthogonal decomposition of isometries in a Banach space. Proc. Am. Math. Soc. **69**, 125–128 (1978)
5. Hardin, C.D.: On the spectral representation of symmetric stable processes. J. Multivariate Anal. **12**, 385–401 (1982)
6. Hosoya, Y.: Harmonizable stable processes. Z. Wahrscheinlichkeitstheor. Verw. Geb. **60**, 517–533 (1982)
7. James, R.C.: Orthogonality and linear functionals in a normed linear space. Trans. Am. Math. Soc. **61**, 265–292 (1947)
8. Kanter, M.: Linear sample spaces and stable processes. J. Funct. Anal. **9**, 441–456 (1972)
9. Kuelbs, J.: A representation theorem for symmetric stable processes and stable measures on *H*. Z. Wahrscheinlichkeitstheor. Verw. Geb. **26**, 259–271 (1973)
10. Lehmann, E.L.: Testing statistical hypotheses. New York: Wiley 1959
11. Lindenstrauss, J., Tzafriri, L.: Classical Banach spaces. Berlin-Heidelberg-New York: Springer 1977
12. Lukacs, E., Laha, R.G.: Applications of characteristic functions. New York: Hafner 1964
13. Rozanov, Yu.A.: Stationary random processes. San Francisco: Holden-Day 1967
14. Schilder, M.: Some structure theorems for the symmetric stable laws. Ann. Math. Statist. **41**, 412–421 (1970)
15. Singer, I.: Best approximation in normal linear space by elements of linear subspaces. New York: Springer 1970
16. Singer, I.: Bases in Banach spaces I. Berlin-Heidelberg-New York: Springer 1970
17. Thu, N.V.: Prediction problems. Dissertationes Math. **163**, 1–69 (1980)
18. Urbanik, K.: Prediction of strictly stationary sequences. Colloq. Math. **12**, 115–129 (1964)
19. Urbanik, K.: Some prediction problems for strictly stationary processes. In: Proc. Fifth Berkeley Symp. Math. Stat. Probab., Vol. 2, Part I, pp. 235–258. University of California Press 1967
20. Urbanik, K.: Harmonizable sequences of random variables. Colloques Internationaux du C.N.R.S. **186**, 345–361 (1970)

Received August 9, 1985; received in revised form April 7, 1988