The Annals of Probability
1991, Vol. 19, No. 2, 843–867

ON THE LINEAR PREDICTION OF MULTIVARIATE (2, p)-BOUNDED PROCESSES[1]

BY CHRISTIAN HOUDRÉ

University of Maryland

We develop the linear least squares prediction theory for some classes of nonstationary processes having a Fourier spectral representation. We study time domain as well as spectral domain properties for these processes, such as a Wold decomposition and a decomposition for matrix bimeasures. We also obtain an autoregressive representation for the optimum predictor.

0. Introduction. The object of this work is to develop a comprehensive study of the linear (least squares) prediction problem for some classes of multivariate discrete time nonstationary processes. The processes under study are the $(2, p)$-bounded processes, $1 \leq p \leq +\infty$, which have been introduced by Bochner (1956) and the author [see Houdré (1990a)]. The $(2, \infty)$-bounded class, also known as V-bounded, has recently seen a revival of popularity essentially due to the work of Niemi (1975a, b), (1977). Such is not the case for $1 \leq p < +\infty$, and our work will thus also include a study of some structural properties of these processes.

$(2, \infty)$-bounded processes are important since they contain wide sense stationary (WSS) processes and also retain some of the fundamental WSS characteristics, such as a Fourier integral representation. However, the $(2, \infty)$-bounded class is too broad. Typically, for WSS processes, one is often interested not in the full generality of the WSS class but rather in some particular subclass, say the WSS processes with absolutely continuous spectral measure. For $p < +\infty$, $(2, p)$-bounded processes play such a role within the $(2, \infty)$-bounded class. Furthermore, they provide a single framework to develop a prediction theory which recovers both the Wiener–Kolmogorov (stationary) theory as well as Kalman's (nonstationary) state space theory. The analogies between the stationary and $(2, p)$-bounded classes do not, however, carry over analytically and new methods have to be developed to study the prediction problem for these processes.

In its modern setting, and for discrete time univariate WSS processes, linear least squares prediction was initiated by Kolmogorov (1941), following the

Received February 1989; revised January 1990.

[1]Written while the author was at The Center for Stochastic Processes, Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27599-3260 and supported by AFOSR Contract F49620-85-C-0144 and ONR Contract N00014-86-C-0227. It was revised while the author was at the Center for Computational Statistics and Probability, George Mason University, Fairfax, Virginia 22030.

AMS 1980 *subject classifications*. Primary 60G12, 60G10; secondary 62M10.

Key words and phrases. Linear prediction, nonstationary processes, Wiener and Kalman filtering.

geometric approach of Wold. Independently, Wiener (1942) motivated by engineering applications, studied the prediction problem and obtained, for continuous time processes, an explicit formula for the optimal predictor. For multivariate WSS processes, Wiener and Masani (1957), (1958) also adopted the geometric framework and essentially exhausted the study of the prediction problem for full rank processes. Their fundamental work, which was partly duplicated by Helson and Lowdenslager (1959), (1961), generated subsequent studies too numerous to list here so we simply refer to Masani (1966) for a good introduction to the subject. For nonstationary processes, prediction has seen a shift in methods and emphasis, originating in the work of Kalman (1960), where the processes to be estimated have a state space representation. Time domain methods have replaced frequency domain methods and analytic tools have been replaced by recursive ones. Kalman's work also generated considerable interest. A good overview of this work as well as extensive references is in Kailath (1974). The reviews of Wiener's work in the third volume of Wiener's (1981) collected work will also provide the reader with recent insights and references.

We now give a brief description of the various parts of this work. In the first section, a few basic definitions are introduced and some theorems, mainly on matricial integration, recalled. The second section is essentially devoted to a characterization of $(2, p)$-boundedness in terms of dominating measures. This provides a unifying framework for the (matricial) Grothendieck's inequality. In the third section a matrix bimeasure integral is studied, a decomposition of matrix bimeasures obtained and other characterizations of $(2, p)$-boundedness provided. The results of the fourth section are built on the previous ones as well as the Wold decomposition. It can be called prediction à la Kolmogorov. Section 5 is prediction à la Wiener–Masani. A new factorization notion is introduced and a further decomposition of matrix bimeasures obtained. In the next to last section, we give an analytical expression for the optimal predictor and study the problem of obtaining this predictor in autoregressive form. In the last section, we present some open problems.

1. Preliminaries. $\mathbb{C}$ is the complex field, $\mathbb{R}$ the real one, $\mathbb{Z}$ the integers, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ and K denotes a generic absolute constant whose value might change from one expression to another. $\mathbb{C}^{k \times k}$ is the set of $k \times k$ matrices with complex entries, and det, tr and * refer to the determinant, the trace and the adjoint of matrices. $\mathbf{A}, \mathbf{B}$, etc., denote $k \times k$ matrices with entries $a_{i,j}$, $b_{i,j}$, etc., while $\mathbf{\Phi}, \mathbf{\Psi}, \boldsymbol{\beta}, \dots,$ are reserved for $\mathbb{C}^{k \times k}$-valued entities with components $\varphi_{i,j}, \psi_{i,j}, \beta_{i,j}, \dots$. The letters $\mathbf{x}, \mathbf{y}, \boldsymbol{\varphi}, \boldsymbol{\psi}$, etc., denote k-dimensional column vector (-valued entities) with components $x_i, y_i, \varphi_i, \psi_i$, etc. The exceptions to these rules generally involve complex valued quantities, for example, $e^{in\cdot}$ denotes both the complex exponentials and $e^{in\cdot}\mathbf{I}$, where $\mathbf{I}$ is the $k \times k$ identity matrix. For $k \times k$ matrices $\mathbf{A}$ and $\mathbf{B}$, we set $(\mathbf{A}, \mathbf{B}) = \operatorname{tr} \mathbf{A}\mathbf{B}^*$ and $\sqrt{\operatorname{tr} \mathbf{A}\mathbf{A}^*}$ is the associated Euclidean norm $\|\mathbf{A}\|_E$.

Next, if H is any complex Hilbert space with inner product $\langle \cdot , \cdot \rangle$ and norm $\| \cdot \|_{\mathbf{H}}$, $\mathbf{H}$ denotes the inflated space $\otimes_{i=1}^{k} \mathbf{H}$. For $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbf{H}$, the square matrices $[\boldsymbol{\varphi}, \boldsymbol{\psi}] = (\langle \varphi_i, \psi_j \rangle)_{i,j \in \{1,\dots,k\}}$ induce a Gram structure on **H**, and **H** is a *pre-Hilbert module* over $\mathbb{C}^{k\times k}$ [see Masani (1966)]. The topology on **H** is the one induced by the norm $\|\cdot\|_{\mathbf{H}}$, associated to the inner product $(\boldsymbol{\varphi}, \boldsymbol{\psi}) = \mathrm{tr}[\boldsymbol{\varphi}, \boldsymbol{\psi}] = \sum_{i=1}^{k} \langle \varphi_i, \psi_i \rangle$. For this topology, **H** is complete, that is, **H** is a *Hilbert module*. In **H**, two elements, $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$, are *orthogonal* ($\boldsymbol{\varphi} \perp \boldsymbol{\psi}$) if $[\boldsymbol{\varphi}, \boldsymbol{\psi}] = 0$, that is, $\langle \varphi_i, \psi_j \rangle = 0$, $i, j \in \{1, \dots, k\}$.

We now introduce a few basic spaces on $\Pi = \mathbb{R}/2\pi\mathbb{Z} =]-\pi, \pi]$. For μ a finite positive measure on Π, $L^p(\mu)$, $1 \le p \le +\infty$, are the usual Lebesgue spaces of complex valued functions. When μ is the normalized Lebesgue measure on Π, these spaces are simply denoted by $L^p(\Pi)$, while $C(\Pi)$, $B(\Pi)$, and $P(\Pi)$ denote, respectively, the sets of (complex valued) continuous functions on Π, Borel bounded ones and the trigonometric polynomials.

Let **M** be a $k \times k$ nonnegative Hermitian matrix valued measure over $(\Pi, \mathscr{B}(\Pi))$, where $\mathscr{B}(\Pi)$ is the Borel σ-algebra of Π. The space $\mathbf{L}^2(\mathbf{M})$ is defined à la Rosenberg (1964) and Rozanov (1967), that is, $\mathbf{L}^2(\mathbf{M})$ is the set of matrix functions $\boldsymbol{\Phi}$ such that

$$\int_{-\pi}^{\pi} \boldsymbol{\Phi}(\theta)\, d\mathbf{M}(\theta) \boldsymbol{\Phi}^*(\theta) = \int_{-\pi}^{\pi} \boldsymbol{\Phi}(\theta) \mathbf{M}'_\tau(\theta) \boldsymbol{\Phi}^*(\theta)\, d\tau(\theta)$$

exists, where $\mathbf{M}'_\tau$ is the Radon–Nikodym derivative of **M** with respect to the trace measure. For $\boldsymbol{\Phi}, \boldsymbol{\Psi} \in \mathbf{L}^2(\mathbf{M})$, $[\boldsymbol{\Phi}, \boldsymbol{\Psi}] = \int_\Pi \boldsymbol{\Phi}\, d\mathbf{M}\, \boldsymbol{\Psi}^* = \int_\Pi \boldsymbol{\Phi} \mathbf{M}'_\tau \boldsymbol{\Psi}^*\, d\tau$ exists and $\mathbf{L}^2(\mathbf{M})$ is a pre-Hilbert module over $\mathbb{C}^{k\times k}$. Under the associated norm $\|\boldsymbol{\Phi}\| = (\int_\Pi \|\boldsymbol{\Phi}\sqrt{\mathbf{M}'_\tau}\|_E^2\, d\tau)^{1/2}$, $\mathbf{L}^2(\mathbf{M})$ is complete ($\sqrt{\mathbf{M}'_\tau}$ is any nonnegative Hermitian square root of $\mathbf{M}'_\tau$). When $d\mathbf{M} = d\mu$, that is, $\mathbf{M} = \mu\mathbf{I}$, the space $\mathbf{L}^2(\mathbf{M})$ as defined above is just the space of matrix functions $\boldsymbol{\Phi} = (\varphi_{i,j})$ such that $\varphi_{i,j} \in L^2(\mu)$, $i, j \in \{1, \dots, k\}$, it is denoted by $\mathbf{L}^2(\mu)$ and the integral is just a componentwise integral. Similar definitions hold for $\mathbf{L}^p(\mu)$, $1 \le p \le +\infty$. When $d\mu = d\theta$, the normalized Lebesgue measure on Π, $\mathbf{L}^p(\mu)$ becomes $\mathbf{L}^p(\Pi)$ and the norms become $\|\boldsymbol{\Phi}\|_p = \{\int_{-\pi}^{\pi} \|\boldsymbol{\Phi}(\theta)\|_E^p\, d\theta\}^{1/p}$, $1 \le p < +\infty$, $\|\boldsymbol{\Phi}\|_\infty = \mathrm{ess\,sup}_{\theta \in]-\pi,\pi]} \|\boldsymbol{\Phi}(\theta)\|_E$. Similarly, $\mathbf{B}(\Pi)$ and $\mathbf{C}(\Pi)$ are the matrix-valued functions whose entries are in $B(\Pi)$ and $C(\Pi)$, while $\mathbf{P}(\Pi)$ is the set of $\mathbb{C}^{k\times k}$ matrix trigonometric polynomials. On $\mathbf{B}(\Pi)$ or $\mathbf{C}(\Pi)$, the norm $\sup_{\theta \in]-\pi,\pi]} \|\boldsymbol{\Phi}(\theta)\|_E$ is also denoted by $\|\boldsymbol{\Phi}\|_\infty$. For $d\mathbf{M} = \mathbf{M}'\, d\theta$, $\boldsymbol{\Phi} \in \mathbf{L}^2(\mathbf{M})$ if and only if $\boldsymbol{\Phi}\sqrt{\mathbf{M}'} \in \mathbf{L}^2(\Pi)$, equivalently, $\boldsymbol{\Phi} \in \mathbf{L}^2(\mathbf{M})$ if and only if the componentwise integral $\int_\Pi \boldsymbol{\Phi}\mathbf{M}'\boldsymbol{\Psi}^*\, d\theta$ exists (this last integral is in fact equal to $\int_\Pi \boldsymbol{\Phi}\mathbf{M}'_\tau \boldsymbol{\Psi}^*\, d\tau$). The spaces $\mathbf{L}^p(\mu)$ enjoy properties similar to the ones satisfied by the unidimensional $L^p(\mu)$. When needed, these properties of $\mathbf{L}^p(\mu)$ will be stated in the text.

When **M** is a σ-additive matrix valued function (not necessarily nonnegative definite), the above definition of integrability can be extended. This follows from the work of Robertson and Rosenberg (1968), and integrability of matrix functions with respect to matrix valued measures, that is, $\int_\Pi \boldsymbol{\Phi}\, d\mathbf{M}\, \boldsymbol{\Psi}$ is always taken in that sense. When $\boldsymbol{\Psi} = \mathbf{I}$, we get $\int_\Pi \boldsymbol{\Phi}\, d\mathbf{M}$ and say that $\boldsymbol{\Phi}$ is *left integrable* (with respect to **M**); $\int_\Pi d\mathbf{M}\, \boldsymbol{\Psi}$ is defined similarly and $\boldsymbol{\Psi}$ is said to be *right integrable* (with respect to **M**).

Let us make a last convention which holds throughout this work. A process is always understood to be of discrete time and **H**-bounded, that is, $\mathbf{x}: \mathbb{Z} \to \mathbf{H}$, with $\|\mathbf{x}_n\|_{\mathbf{H}} \le K$, $n \in \mathbb{Z}$. We now present our first definition and refer the reader to Houdré (1990a) for more details.

Definition 1.1. A process x is $(2, p)$*-bounded*, $1 \le p \le +\infty$, if there exists $K > 0$ such that

$$\left\| \sum_{j=1}^{N} \mathbf{P}_j \mathbf{x}_{n_j} \right\|_{\mathbf{H}} \le K \|\mathbf{P}\|_p \tag{1.1}$$

for all matrix trigonometric polynominals **P** of the form $\mathbf{P}(\theta) = \sum_{j=1}^{N} \mathbf{P}_j e^{-in_j\theta}$.

It is immediate to verify that $\mathscr{M}^{\mathbf{p}}$, $1 \le p \le +\infty$, denotes the $\mathbb{C}^{k\times k}$-linear space of $(2, p)$-bounded processes, we have $\mathscr{M}^1 \subset \mathscr{M}^{\mathbf{r}} \subset \mathscr{M}^{\mathbf{s}} \subset \mathscr{M}^{\infty}$, $1 \le r \le s \le +\infty$ and that WSS processes are $(2, \infty)$-bounded; however, for $p < +\infty$, this will not necessarily be the case. A WSS process is said to have L^p*-spectrum*, $1 \le p \le +\infty$, if its spectral matrix measure is absolutely continuous with a Radon–Nikodym derivative in $\mathbf{L}^{\mathbf{p}}(\Pi)$. With this notion, we have:

Theorem 1.2. *For $2 \le p < \infty$, a WSS process is $(2, p)$-bounded if and only if it has $L^{p/(p-2)}$-spectrum (L^∞-spectrum for $p = 2$). For $1 \le p < 2$, a WSS process is $(2, p)$-bounded if and only if it is the zero process.*

Proof. Let $2 < p < +\infty$ and let **x** be a WSS process with $L^{p/(p-2)}$-spectrum. Then, $[\mathbf{x}_n, \mathbf{x}_m] = \int_{-\pi}^{\pi} e^{i(n-m)\theta} \mathbf{\Phi}'(\theta)\, d\theta$, $n, m \in \mathbb{Z}$, with $\mathbf{\Phi}' \in \mathbf{L}^{\mathbf{p}/(\mathbf{p}-\mathbf{2})}(\Pi)$. Hence

$$\begin{aligned}
\left\| \sum_{j=1}^{N} \mathbf{P}_j \mathbf{x}_{n_j} \right\|_{\mathbf{H}}^2 &= \int_{-\pi}^{\pi} \operatorname{tr}\left(\sum_{j=1}^{N} \mathbf{P}_j e^{in_j\theta} \right) \mathbf{\Phi}'(\theta) \left(\sum_{j=1}^{N} \mathbf{P}_j e^{in_j\theta} \right)^{*} d\theta \\
&\le \int_{-\pi}^{\pi} \|\mathbf{P}(\theta)\|_E^2 \operatorname{tr} \mathbf{\Phi}'(\theta)\, d\theta \\
&\le \sqrt{k} \int_{-\pi}^{\pi} \|\mathbf{P}(\theta)\|_E^2 \|\mathbf{\Phi}'(\theta)\|_E\, d\theta \\
&\le \sqrt{k} \left(\int_{-\pi}^{\pi} \|\mathbf{P}(\theta)\|_E^p\, d\theta \right)^{2/p} \left(\int_{-\pi}^{\pi} \|\mathbf{\Phi}'(\theta)\|_E^{p/(p-2)}\, d\theta \right)^{(p-2)/p},
\end{aligned}$$

by Hölder's inequality. Thus, **x** is $(2, p)$-bounded. Similarly, for $p = 2$, $\|\sum_{j=1}^{N} \mathbf{P}_j \mathbf{x}_{n_j}\|_{\mathbf{H}}^2 \le \sqrt{k}\, \|\mathbf{\Phi}'\|_\infty^2 \|\mathbf{P}\|_2^2$. Let $2 < p < +\infty$ and let **x** be a $(2, p)$-bounded WSS process with matrix spectral measure $\mathbf{\Phi}$. Then by (1.1),

$$\operatorname{tr} \int_{\Pi} \mathbf{P} \mathbf{\Phi}'_{\tau} \mathbf{P}^*\, d\tau \le K \left(\int_{\Pi} \|\mathbf{P}\|_E^p\, d\theta \right)^{2/p}. \tag{1.2}$$

Since matrix trigonometric polynomials are dense in $\mathbf{L}^{\mathbf{p}}(\Pi)$, (1.2) can be extended to

$$\text{(1.3)} \qquad \operatorname{tr}\int_{\Pi} \mathbf{\Psi}\mathbf{\Phi}'_\tau\mathbf{\Psi}^*\,d\tau \le K\left(\int_{\Pi} \|\mathbf{\Psi}\|_E^p\,d\theta\right)^{2/p}, \qquad \mathbf{\Psi} \in \mathbf{L}^{\mathbf{p}}(\Pi).$$

Let $\mathbf{\Psi} = \chi_E \mathbf{I}$, with $E \in \mathscr{B}(\Pi)$ and $|E| = 0$, ($|E|$ denotes the Lebesgue measure of E), then $\int_{\Pi}\mathbf{\Phi}'_\tau\,d\tau = \operatorname{tr}\mathbf{\Phi}(E) = 0$, hence $\mathbf{\Phi}(E) = 0$ and $d\mathbf{\Phi} = \mathbf{\Phi}'\,d\theta$, $\mathbf{\Phi}' \in \mathbf{L}^1(\Pi)$. Now, for the matrix function $\mathbf{\Psi} = \|\mathbf{\Phi}'\|_E^{1/(p-2)}\mathbf{I} \in \mathbf{L}^{\mathbf{p}}(\Pi)$ and since $\operatorname{tr}\mathbf{\Phi}' \ge \|\mathbf{\Phi}'\|_E$, (1.3) becomes

$$\text{(1.4)} \qquad \int_{\Pi}\|\mathbf{\Phi}'\|_E^{p/(p-2)}\,d\theta \le K\left\{\int_{\Pi}\|\mathbf{\Phi}'\|_E^{p/(p-2)}\,d\theta\right\}^{2/p}.$$

In other words, if $\mathbf{x}$ is not the zero process, $\{\int_{\Pi}\|\mathbf{\Phi}'\|_E^{p/(p-2)}\,d\theta\}^{(p-2)/p} \le K$, which is the result. For $p = 2$, similar arguments show that, $(1/|E|)\int_E\|\mathbf{\Phi}'\|_E\,d\theta \le K$, for all $E \in \mathscr{B}(\Pi)$ with $|E| > 0$, that is, $\|\mathbf{\Phi}'\|_E \in L^\infty(\Pi)$. For $1 \le p < 2$, the disjointness of the two classes can be obtained (using $\|\cdot\|_E$) essentially as in Houdré (1990a). □

We now briefly sketch an integration theory of matrix valued functions with respect to stochastic measures. This can be seen as either the restriction to our framework of Bartle's (1956) general vector integral or as the matricial version of the Bartle, Dunford and Schwartz (1955) integral; details are also in Dunford and Schwartz [(1957), 4.10].

DEFINITION 1.3. A *stochastic measure* $\boldsymbol{\zeta}$ is an $\mathbf{H}$-valued function on $\mathscr{B}(\Pi)$ such that

$$\text{(1.5)} \qquad \boldsymbol{\zeta}(E) = \sum_{i=1}^{\infty}\boldsymbol{\zeta}(E_i), \qquad E \in \mathscr{B}(\Pi),$$

for every Borel partition $\{E_i\}_{i\in\mathbb{N}^*}$ of E, where the convergence in (1.5) is in the norm of $\mathbf{H}$.

Following Masani (1968), we also say that a stochastic measure $\boldsymbol{\zeta}$ is *orthogonally scattered* if $[\boldsymbol{\zeta}(E), \boldsymbol{\zeta}(F)] = 0$ whenever $E \cap F = \phi$, $E, F \in \mathscr{B}(\Pi)$. The set of stochastic measures is denoted by $\mathbf{M}^{\mathbf{s}}(\Pi)$; it is a $\mathbb{C}^{k\times k}$-linear space.

The Vitali variation (which for stochastic measures is not necessarily finite) is replaced by the *Fréchet* (or *semi-*) *variation*. For each E in $\mathscr{B}(\Pi)$,

$$\|\zeta\|(E) = \sup\left\{\left\|\sum_{i=1}^{N}\mathbf{A}_i\boldsymbol{\zeta}(E_i)\right\|_{\mathbf{H}}, \forall\, N \in \mathbb{N}^*, \forall\,\{E_i\}_{i=1}^N \text{ Borel partition of } E, \right.$$
$$\left. \forall\,\{\mathbf{A}_i\}_{i=1}^N, \mathbf{A}_i \in \mathbb{C}^{k\times k}, \|\mathbf{A}_i\|_E \le 1\right\}$$

is called the Fréchet variation of E. As defined, the Fréchet variation $\|\boldsymbol{\zeta}\|(\cdot)$ is monotone and σ-subadditive on $\mathscr{B}(\Pi)$. In addition, for each $E \in \mathscr{B}(\Pi)$ and $\mathbf{A} \in \mathbb{C}^{k\times k}$, we have $\|\boldsymbol{\zeta}(E)\|_{\mathbf{H}} \le \|\boldsymbol{\zeta}\|(E)$ and $\|\mathbf{A}\boldsymbol{\zeta}\|(E) \le \|\mathbf{A}\|_E\|\boldsymbol{\zeta}\|(E)$. Furthermore, $\|\boldsymbol{\zeta}\|(\Pi) < +\infty$.

With the help of the matricial semivariation, classical one-dimensional definitions such as *$\boldsymbol{\zeta}$-null set*, *$\boldsymbol{\zeta}$-almost everywhere* ($\boldsymbol{\zeta}$-a.e.), *$\boldsymbol{\zeta}$-measurable* and *$\boldsymbol{\zeta}$-essentially bounded* can be extended. These definitions as well as the results stated below can be obtained following, mutatis mutandis, the steps in the works previously mentioned, as such, they are only briefly stated. However, we sketch the appropriate integral. For $\boldsymbol{\Phi}$, a (Borel) matrix simple function, i.e., $\boldsymbol{\Phi} = \sum_{i=1}^{N} \mathbf{A}_i \chi_{E_i}$, $\mathbf{A}_i \in \mathbb{C}^{k \times k}$, $E_i \in \mathscr{B}(\Pi)$ disjoints, the $\boldsymbol{\zeta}$-integral of $\boldsymbol{\Phi}$ over $E \in \mathscr{B}(\Pi)$ is defined by $\int_E \boldsymbol{\Phi}\, d\boldsymbol{\zeta} = \sum_{i=1}^{N} \mathbf{A}_i \boldsymbol{\zeta}(E_i \cap E)$. Then, a function $\boldsymbol{\Phi}\colon \Pi \to \mathbb{C}^{k \times k}$ is $\boldsymbol{\zeta}$-integrable if there exists a sequence $\{\boldsymbol{\Phi}_n\}_n$ of matrix simple functions converging to $\boldsymbol{\Phi}$, $\boldsymbol{\zeta}$-a.e. and such that $\{\int_E \boldsymbol{\Phi}_n\, d\boldsymbol{\zeta}\}_n$ converges (is Cauchy) in $\mathbf{H}$ for each E in $\mathscr{B}(\Pi)$. This limit is defined to be the $\boldsymbol{\zeta}$-integrable of $\boldsymbol{\Phi}$ over E and is denoted by $\int_E \boldsymbol{\Phi}(\theta)\, d\boldsymbol{\zeta}(\theta)$. The $\boldsymbol{\zeta}$-integral shares most of the properties of the Lebesgue integral, in particular, a dominated convergence theorem holds. The $\mathbb{C}^{k \times k}$-linear space of $\boldsymbol{\zeta}$-integrable matrix functions is denoted by $\mathbf{L}^1(\boldsymbol{\zeta})$. The following result is fully stated since used in a few instances.

THEOREM 1.4. (i) *Let $\boldsymbol{\Phi}$ be $\boldsymbol{\zeta}$-measurable and $\boldsymbol{\zeta}$-essentially bounded on Π, then $\boldsymbol{\Phi}$ is $\boldsymbol{\zeta}$-integrable and* $\| \int_E \boldsymbol{\Phi}\, d\boldsymbol{\zeta} \| \le \boldsymbol{\zeta}\text{-ess}_E \sup \|\boldsymbol{\Phi}\|_E \|\boldsymbol{\zeta}\|(E)$.

(ii) *Let $\boldsymbol{\zeta} \in \mathbf{M}^{\mathbf{s}}(\Pi)$ and let $\mathbf{T}$ be a bounded $\mathbb{C}^{k \times k}$-linear operator on $\mathbf{H}$. Then* $\mathbf{T} \circ \boldsymbol{\zeta} \in \mathbf{M}^{\mathbf{s}}(\Pi)$ *with for each* $E \in \mathscr{B}(\Pi)$, $\|\mathbf{T} \circ \boldsymbol{\zeta}\|(E) \le \|\mathbf{T}\|_0 \|\boldsymbol{\zeta}\|(E)$ ($\| \cdot \|_0$ *is the norm of the corresponding operator*). *Furthermore, if $\boldsymbol{\Phi} \in \mathbf{L}^1(\boldsymbol{\zeta})$, then $\boldsymbol{\Phi} \in \mathbf{L}^1(\mathbf{T} \circ \boldsymbol{\zeta})$ and* $\mathbf{T} \int_E \boldsymbol{\Phi}\, d\boldsymbol{\zeta} = \int_E \boldsymbol{\Phi}\, d\mathbf{T} \circ \boldsymbol{\zeta}$.

As a particular case of (i): If $\boldsymbol{\Phi} \in \mathbf{B}(\Pi)$, then $\boldsymbol{\Phi} \in \mathbf{L}^1(\boldsymbol{\zeta})$. The $\mathbb{C}^{k \times k}$-linearity of $\mathbf{T}$ is essential to ensure $\mathbf{T} \circ \boldsymbol{\zeta}$-integrability in (ii). If $\mathbf{T}$ is only linear, this is no longer true.

To close the preliminary section, we state another folklore result: a characterization of $(2, \infty)$-boundedness. This is a direct consequence of the work of Grothendieck (1953) or of Bartle, Dunford and Schwartz (1955) on the representation of bounded linear operators from $C(\Pi)$ to $\mathbf{H}$. As is well known, for $\mathbf{x}$ is WSS, $\boldsymbol{\zeta}$ below is orthogonally scattered.

THEOREM 1.5. *A process $\mathbf{x}$ is $(2, \infty)$-bounded if and only if there exists a (unique) stochastic measure $\boldsymbol{\zeta}$ such that* $\mathbf{x}_n = \hat{\boldsymbol{\zeta}}(n) = \int_{-\pi}^{\pi} e^{in\theta}\, d\boldsymbol{\zeta}(\theta)$, $n \in \mathbb{Z}$.

Since $\mathscr{M}^{\mathbf{p}} \subset \mathscr{M}^{\infty}$, $p < \infty$, $(2, p)$-bounded processes can be represented as in Theorem 1.5, moreover the corresponding stochastic measure $\boldsymbol{\zeta}$ has some additional properties. For $E \in \mathscr{B}(\Pi)$, we define the *p-Fréchet* (or *-semi*) *variation* as $\|\boldsymbol{\zeta}\|_p(E) = \sup\{\|\sum_{i=1}^{N} \mathbf{A}_i \boldsymbol{\zeta}(E_i)\|_{\mathbf{H}}\}$, where the supremum is taken over all matrix simple functions $\boldsymbol{\Phi} = \sum_{i=1}^{N} \mathbf{A}_i \chi_{E_i}$ such that $\|\boldsymbol{\Phi}\|_p \le 1$. Then, it is not difficult to see that $\mathbf{x}$ is $(2, p)$-bounded if and only if $\mathbf{x}_n = \hat{\boldsymbol{\zeta}}(n)$ for some $\boldsymbol{\zeta}$ of finite p-Fréchet variation. To verify that $\|\boldsymbol{\zeta}\|_p(\Pi) < +\infty$ is quite impractical, and not much better than (1.1). To study the prediction problem we need to find more practical characterizations of the classes $\mathscr{M}^{\mathbf{p}}$. This is the object of our next section.

2. Domination results. Results on vector measures often rely on Grothendieck's inequality. Such is the case of the theory of $(2,p)$-bounded processes. Among the various equivalent forms of this inequality, the one of particular interest is the form due to Pietsch (1969). Before presenting these and connected results we need:

DEFINITION 2.1. A stochastic measure $\boldsymbol{\zeta}$ is *majorizable* if there exists a finite positive Borel measure μ on Π (called a *Grothendieck's measure* or a *dominating measure*) such that

$$\left\| \int_{\Pi} \mathbf{\Phi}\, d\boldsymbol{\zeta} \right\|_{\mathbf{H}} \le \left\{ \int_{\Pi} \|\mathbf{\Phi}\|_E^2\, d\mu \right\}^{1/2} \tag{2.1}$$

for all continuous matrix functions $\mathbf{\Phi}$ on Π.

Pietsch's univariate result is to the effect that a stochastic measure $\boldsymbol{\zeta}$ is *always* majorizable. Multivariate versions (even infinite dimensional) of this domination result have also been obtained by Truong-Van (1981) and Rosenberg (1982). In its full generality (2.1) gives very little precise information about μ. One of the great features of the $(2,p)$-bounded classes, $p < +\infty$, is the possibility to *characterize* the dominating measures. First we need a lemma. Its proof can be obtained by modifications of arguments given, for $p = +\infty$, by Truong-Van (1981) (using Proposition 2 there, and replacing $\|\cdot\|_\infty$ by $\|\cdot\|_{p/2}$).

LEMMA 2.2. *Let* $\mathbf{x}$ *be* $(2,p)$*-bounded*, $2 \le p < +\infty$, *with spectral stochastic measure* $\boldsymbol{\zeta}$. *Then there exists* $K > 0$ *such that* $\sum_{i=1}^N \|\int_\Pi \mathbf{\Phi}_i\, d\boldsymbol{\zeta}\|_{\mathbf{H}}^2 \le K \|\sum_{i=1}^N \|\mathbf{\Phi}_i\|_E^2\|_{p/2}$ *for all* $N \in \mathbb{N}^*$, $\mathbf{\Phi}_1, \mathbf{\Phi}_2, \dots, \mathbf{\Phi}_N \in \mathbf{C}(\Pi)$.

THEOREM 2.3. *Let* $\mathbf{x}$ *be* $(2,p)$*-bounded*, $2 \le p < +\infty$ [*resp*., $(2,\infty)$*-bounded*], *with spectral stochastic measure* $\boldsymbol{\zeta}$. *Then there exists a nonnegative function* μ *in* $L^{p/(p-2)}(\Pi)$ (*nonnegative finite Borel measure* μ) *such that*

$$\left\|\int_{\Pi} \mathbf{\Phi}\, d\boldsymbol{\zeta}\right\|_{\mathbf{H}}^2 \le \int_{\Pi} \|\mathbf{\Phi}\|_E^2 \mu\, d\theta \qquad \left(resp.,\ \int_{\Pi} \|\mathbf{\Phi}\|_E^2\, d\mu \right) \tag{2.2}$$

for all continuous matrix functions $\mathbf{\Phi}$ *on* Π.

PROOF. The case $p = +\infty$ is Truong-Van's. Let $p < +\infty$ and let φ be a continuous real valued function on Π and let K be any constant appearing in Lemma 2.2. Let finally

$$Q(\varphi) = \inf \left\{ \left(\int_{\Pi} \left| \varphi + K \sum_{i=1}^N \|\mathbf{\Phi}_j\|_E^2 \right|^{p/2} d\theta \right)^{2/p} - \sum_{j=1}^N \left\| \int_{\Pi} \mathbf{\Phi}_j\, d\boldsymbol{\zeta} \right\|_{\mathbf{H}}^2 \right\},$$

where the infimum is taken over all the finite sets $\{\mathbf{\Phi}_1, \mathbf{\Phi}_2, \dots, \mathbf{\Phi}_N\} \subset \mathbf{C}(\Pi)$. As in Houdré (1990a), Q is an inhomogeneous, subadditive functional on $C^r(\Pi)$ (the space of real-valued continuous functions on Π) such that $-\|\varphi\|_{p/2} \le Q(\varphi) \le \|\varphi\|_{p/2}$. By two applications of the Hahn–Banach theorem, there exists a complex linear functional L on $C(\Pi)$ such that $|L(\varphi_1 + i\varphi_2)| \le \sqrt{2}\,\|\varphi_1 + i\varphi_2\|_{p/2}$ L can be extended to $L^{p/2}(\Pi)$. By the Riesz representation theorem, there exists $\mu_0 \in L^{p/(p-2)}(\Pi)$ such that $L(\varphi) = \int_\Pi \varphi(\theta)\mu_0(\theta)\,d\theta$, for all $\varphi \in L^{p/2}(\Pi)$. Then, $L(\varphi) \ge 0$ for $\varphi \ge 0$ in $C(\Pi)$ and thus for $\varphi \ge 0$ in $L^{p/2}(\Pi)$. This in turn gives $\mu_0 \ge 0$. Finally, for $\mathbf{\Phi} \in \mathbf{C}(\Pi)$, $Q(-K\|\mathbf{\Phi}\|_E^2) \le -\|\int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}\|_{\mathbf{H}}^2$, hence $-KL(\|\mathbf{\Phi}\|_E^2) = L(-K\|\mathbf{\Phi}\|_E^2) \le Q(-K\|\mathbf{\Phi}\|_E^2) \le -\|\int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}\|_{\mathbf{H}}^2$, that is, $\|\int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}\|_{\mathbf{H}}^2 \le K\int_\Pi \|\mathbf{\Phi}\|_E^2 \mu_0\,d\theta$. □

The information conveyed by Theorem 2.3 is that for p finite, $(2, p)$-bounded processes have in our nonstationary framework the role of WSS processes with absolutely continuous L^p-spectrum. We also note that the majorizability definition of stochastic measure could equally have been given via the Gram structure since $[\int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}, \int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}] \le \int_\Pi \mathbf{\Phi}\mathbf{\Phi}^*\,d\mu$ holds for all $\mathbf{\Phi} \in \mathbf{C}(\Pi)$ if and only if (2.1) holds for all $\mathbf{\Phi} \in \mathbf{C}(\Pi)$, and such is the case when and only when there exists a positive matrix measure $\mathbf{M}$ such that, for all $\mathbf{\Phi} \in \mathbf{C}(\Pi)$, $[\int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}, \int_\Pi \mathbf{\Phi}\,d\boldsymbol{\zeta}] \le \int_\Pi \mathbf{\Phi}\,d\mathbf{M}\,\mathbf{\Phi}^*$. A stochastic measure is $(2, p)$-bounded if (2.2) holds for some $\mu \in L^{p/(p-2)}(\Pi)$. Combining Theorems 1.5 and 2.3 gives:

THEOREM 2.4. *A process* $\mathbf{x}$ *is* $(2, p)$*-bounded if and only if there exists a* $(2, p)$*-bounded stochastic measure* $\boldsymbol{\zeta}$ *such that* $\mathbf{x}_n = \hat{\boldsymbol{\zeta}}(n)$ *for all* n *in* $\mathbb{Z}$.

3. Matrix bimeasure integration. Since for a $(2, p)$-bounded process $\mathbf{x}$, $\mathbf{x}_n = \int_{-\pi}^{\pi} e^{in\theta}\,d\boldsymbol{\zeta}(\theta)$, it has to be expected that $\mathbf{R}(n, m) = [\mathbf{x}_n, \mathbf{x}_m]$ also has a Fourier integral representation, namely, we would like to write $\mathbf{R}(n, m) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{in\theta}e^{-im\psi}\,d[\boldsymbol{\zeta}(\theta), \boldsymbol{\zeta}(\psi)]$. This double integral must, however, be interpreted cautiously and in general is not a Lebesgue integral.

DEFINITION 3.1. A *matrix bimeasure* $\boldsymbol{\beta}$ is a $\mathbb{C}^{k\times k}$-valued function on the Cartesian product $\mathscr{B}(\Pi) \times \mathscr{B}(\Pi)$ such that $\boldsymbol{\beta}(E, \cdot\,)$ and $\boldsymbol{\beta}(\cdot\,, F)$ are σ-additive for all E, F in $\mathscr{B}(\Pi)$.

A matrix bimeasure is *positive definite* (p.d.) whenever

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \mathbf{A}_i \boldsymbol{\beta}(E_i, E_j)\mathbf{A}_j^* \ge 0$$

for all $N \in \mathbb{N}^*$, $E_1, \dots, E_N \in \mathscr{B}(\Pi)$, $\mathbf{A}_1, \dots, \mathbf{A}_N \in \mathbb{C}^{k\times k}$, equivalently, whenever $\sum_{i=1}^N \sum_{j=1}^N \mathbf{z}_i \boldsymbol{\beta}(E_i, E_j)\mathbf{z}_j^* \ge 0$ for all $N \in \mathbb{N}^*$, $E_1, \dots, E_N \in \mathscr{B}(\Pi)$, $\mathbf{z}_1, \dots, \mathbf{z}_N \in \mathbb{C}^k$. It is easily verified that the set function $\boldsymbol{\beta}(\cdot\,, \cdot\,) = [\boldsymbol{\zeta}(\cdot), \boldsymbol{\zeta}(\cdot)]$ induced by the stochastic measure $\boldsymbol{\zeta}$ is a p.d. matrix bimeasure and indeed, by the usual Gaussian process argument, any p.d. matrix bimeasure is given that way. As alluded by the induced case, a matrix bimeasure is not necessarily of bounded Vitali variation and thus does not extend to a matrix measure on

$\mathscr{B}(\Pi) \otimes \mathscr{B}(\Pi)$. However, another notion of variation is finite. This result follows readily from the corresponding complex valued case as given by Ylinen (1978).

THEOREM 3.2. *A matrix bimeasure* $\boldsymbol{\beta}$ *has finite Fréchet variation, that is,* $\|\boldsymbol{\beta}\|(\Pi, \Pi) < \infty$ *where for each* E, F *in* $\mathscr{B}(\Pi)$,

$$\|\boldsymbol{\beta}\|(E, F) = \sup\left\{ \left\| \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{A}_i \boldsymbol{\beta}(E_i, E_j) \mathbf{B}_j^* \right\|_E, \forall N \in \mathbb{N}^*, \right.$$

$$\forall \{E_i\}_{i=1}^N, \{F_j\}_{j=1}^N \textit{ Borel partitions of } E \textit{ and } F,$$

$$\left. \forall \{\mathbf{A}_i\}_{i=1}^N, \{\mathbf{B}_j\}_{j=1}^N, \mathbf{A}_i, \mathbf{B}_j \in \mathbb{C}^{k \times k}, \|\mathbf{A}_i\|_E \le 1, \|\mathbf{B}_j\|_E \le 1 \right\}.$$

In order to obtain a Fourier representation for the covariance of a $(2, p)$-bounded process, the integral of matrix valued functions with respect to a matrix bimeasure has to be defined. Our definition of integrability relies on two results. The first one is due to Robertson and Rosenberg (1968), and asserts that if $\mathbf{M}$: $\mathscr{B}(\Pi) \to \mathbb{C}^{k \times k}$ is a matrix measure and if $\boldsymbol{\Phi}$: $\Pi \to \mathbb{C}^{k \times k}$ is left $\mathbf{M}$-integrable, then the set function $\mathbf{N}$: $\mathscr{B}(\Pi) \to \mathbb{C}^{k \times k}$, $E \to \mathbf{N}(E) = \int_E \boldsymbol{\Phi}\, d\mathbf{M}$, is also matrix measure [in all of this section, integration is taken in the sense of Robertson and Rosenberg (see Preliminaries)]. The second result is just the multivariate version of a result of Ylinen and states that if $\boldsymbol{\Phi}$: $\Pi \to \mathbb{C}^{k \times k}$ is left $\boldsymbol{\beta}(\cdot, F)$-integrable for all $F \in \mathscr{B}(\Pi)$, the set function ${}_{\boldsymbol{\Phi}}\boldsymbol{\beta}(E, \cdot)$: $\mathscr{B}(\Pi) \to \mathbb{C}^{k \times k}$, $F \to {}_{\boldsymbol{\Phi}}\boldsymbol{\beta}(E, F) = \int_E \boldsymbol{\Phi}(\cdot)\, d\boldsymbol{\beta}(\cdot, F)$ is a matrix measure for each $E \in \mathscr{B}(\Pi)$. Combining these, we see that ${}_{\boldsymbol{\Phi}}\boldsymbol{\beta}(\cdot, \cdot)$ and $\boldsymbol{\beta}_{\boldsymbol{\Psi}}(\cdot, \cdot)$ are actually matrix bimeasures. Of course, we can replace left integrability by right integrability in both results.

DEFINITION 3.3. A pair of matrix functions $(\boldsymbol{\Phi}, \boldsymbol{\Psi})$ is said to be integrable with respect to the matrix bimeasure $\boldsymbol{\beta}$ ($\boldsymbol{\beta}$-integrable for short) if the following two conditions hold:

(i) $\boldsymbol{\Phi}$ is left $\boldsymbol{\beta}(\cdot, F)$-integrable for all $F \in \mathscr{B}(\Pi)$ and $\boldsymbol{\Psi}$ is right $\boldsymbol{\beta}(E, \cdot)$-integrable for all $E \in \mathscr{B}(\Pi)$.

(ii) $\boldsymbol{\Phi}$ is left $\boldsymbol{\beta}_{\boldsymbol{\Psi}}(\cdot, F)$-integrable for all $F \in \mathscr{B}(\Pi)$ and $\boldsymbol{\Psi}$ is right ${}_{\boldsymbol{\Phi}}\boldsymbol{\beta}(E, \cdot)$-integrable for all $E \in \mathscr{B}(\Pi)$.

For complex valued functions, the integration theory with respect to bimeasures was initiated by Morse and Transue and our integral is a restricted matricial Morse–Transue integral. Under weaker conditions than the ones of Definition 3.5, an extra Fubini type property is usually required to ensure β-integrability; this is unnecessary here since the following result holds [see Houdré (1989) for the univariate proof, which can easily be adapted, and the references therein for more information on the bimeasure integral].

THEOREM 3.4. *Let the pair* $(\mathbf{\Phi}, \mathbf{\Psi})$ *be* $\boldsymbol{\beta}$*-integrable. Then*

$$\int_E \mathbf{\Phi}(\cdot)\, d\boldsymbol{\beta}_{\mathbf{\Psi}}(\cdot, F) = \int_F \mathbf{\Psi}(\cdot)\, d_{\mathbf{\Phi}}\boldsymbol{\beta}(E, \cdot)$$

for all $E, F \in \mathscr{B}(\Pi)$, *and the common value is denoted by* $\int_E \int_F \mathbf{\Phi}\, d\boldsymbol{\beta}\, \mathbf{\Psi}$.

As given by Definition 3.3 and Theorem 3.4, the matrix bimeasure integral shares familiar properties, such as bilinearity in $(\mathbf{\Phi}, \mathbf{\Psi})$, dominated convergence, etc. Since the matrix bimeasures encountered in the rest of this study are induced by stochastic measures, the stochastic integral results suffice for our purpose. This is also shown by our next two results. The proof of the first one follows by matricial adaption of arguments in Houdré (1989), while the second can be obtained with the help of Theorem 1.5.

THEOREM 3.5. *Let* $\boldsymbol{\beta}$ *be an induced matrix bimeasure, that is, let* $\boldsymbol{\beta}(\cdot, \cdot) = [\boldsymbol{\zeta}(\cdot), \boldsymbol{\zeta}(\cdot)]$ *for some stochastic measure* $\boldsymbol{\zeta}$. *A pair* $(\mathbf{\Phi}, \mathbf{\Phi}^*)$ *is* $\boldsymbol{\beta}$*-integrable if and only if* $\mathbf{\Phi}$ *is* $\boldsymbol{\zeta}$*-integrable; in either case and for all* $E \in \mathscr{B}(\Pi)$, $[\int_E \mathbf{\Phi}\, d\boldsymbol{\zeta}, \int_E \mathbf{\Phi}\, d\boldsymbol{\zeta}] = \int_E \int_E \mathbf{\Phi}\, d\boldsymbol{\beta}\, \mathbf{\Phi}^*$.

COROLLARY 3.6. *A process* $\mathbf{x}$ *is* $(2, \infty)$*-bounded if and only if there exists a (unique positive definite) matrix bimeasure* $\boldsymbol{\beta}$ *such that* $[\mathbf{x}_n, \mathbf{x}_m] = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{in\theta}\, d\boldsymbol{\beta}(\theta, \psi) e^{-im\psi}$, $n, m \in \mathbb{Z}$.

When $\boldsymbol{\beta}$ is induced by the stochastic measure $\boldsymbol{\zeta}$ of a $(2, \infty)$-bounded process $\mathbf{x}$, $[\mathbf{\Phi}, \mathbf{\Psi}] = \int_\Pi \int_\Pi \mathbf{\Phi}\, d\boldsymbol{\beta}\, \mathbf{\Psi}^*$ defines a matricial semi-inner product, which becomes a Gram product after identification of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ such that $[\mathbf{\Phi} - \mathbf{\Psi}, \mathbf{\Phi} - \mathbf{\Psi}] = 0$. The associated inner product space $\mathbf{L}^2(\boldsymbol{\beta}) = \{\mathbf{\Phi}\colon \Pi \to \mathbb{C}^{k \times k}\colon (\mathbf{\Phi}, \mathbf{\Phi}) = \operatorname{tr}[\mathbf{\Phi}, \mathbf{\Phi}] < +\infty\}$ is called the *spectral domain* of the process $\mathbf{x}$. The WSS case corresponds to $\boldsymbol{\beta}(E, F) = 0$, whenever $E \cap F = \phi$, and $\boldsymbol{\beta}$ uniquely determines a positive matrix measure supported on the diagonal of $\Pi \times \Pi$.

REMARK 3.7. From Theorem 3.5 and Corollary 3.6, we have $\mathbf{L}^2(\boldsymbol{\beta}) = \mathbf{L}^1(\boldsymbol{\zeta})$, thus $\mathbf{L}^2(\boldsymbol{\beta})$ is complete if and only if $\mathbf{L}^1(\boldsymbol{\zeta})$ is complete. In general $\mathbf{L}^2(\boldsymbol{\beta})$ is not a Hilbert space, even when $\boldsymbol{\beta}$ is a measure on Π^2 [see Miamee and Salehi (1989) for a univariate counterexample]. This important difference between stationary and $(2, \infty)$-bounded processes does not represent, for prediction purposes, an impossible drawback. The completeness of $\mathbf{L}^2(\boldsymbol{\beta})$ is only indispensable in obtaining autoregressive predictors and this problem has only been partially solved, even for stationary processes.

We complete this section by studying the decomposition of matrix bimeasures.

DEFINITION 3.8. A matrix bimeasure $\boldsymbol{\beta}$ is said to be *continuous* if $\boldsymbol{\beta}(E, F) = 0$ for all finite subsets E, F of Π and *discrete* if there exist two increasing sequences $\{E_n\}$ and $\{F_n\}$ of finite subsets of Π such that

$$\lim_{n\to\infty} \|\boldsymbol{\beta} - \boldsymbol{\beta}_n\|(\Pi, \Pi) = 0,$$

where $\boldsymbol{\beta}_n$ is the restriction of $\boldsymbol{\beta}$ to (E_n, F_n), that is, $\boldsymbol{\beta}_n(G, H) = \boldsymbol{\beta}(G \cap E_n, H \cap F_n)$ for all G, H in $\mathscr{B}(\Pi)$.

THEOREM 3.9. *A positive definite matrix bimeasure $\boldsymbol{\beta}$ admits a unique decomposition $\boldsymbol{\beta} = \boldsymbol{\beta}_c + \boldsymbol{\beta}_d$, where $\boldsymbol{\beta}_c$ is a continuous positive definite matrix bimeasure and where $\boldsymbol{\beta}_d$ is a discrete positive definite matrix bimeasure.*

PROOF. Let μ be any of the finite positive dominating measures in Theorem 2.3 ($\boldsymbol{\beta}$ is p.d., hence induced by a stochastic measure) and let $E = \{\theta \in \Pi\colon \mu(\theta) > 0\}$. E is at most countable, and $E = \bigcup E_n$ where $E_1 \subset E_2 \subset \cdots \subset E_n \subset \cdots$ are finite sets. Let $\boldsymbol{\beta}_d$ (resp., $\boldsymbol{\beta}_n$) be the restriction of $\boldsymbol{\beta}$ to (E, E) [resp., to (E_n, E_n)], then $\boldsymbol{\beta}_d$ and the $\boldsymbol{\beta}_n$ are p.d. We now show that $\boldsymbol{\beta}_d$ is discrete. Let $\{G_i\}$ to be a Borel partition of Π, let $\boldsymbol{\Phi} = \sum_{i=1}^N \mathbf{A}_i \chi_{G_i}, \mathbf{A}_i \in \mathbb{C}^{k\times k}$, $\|\mathbf{A}_i\|_E \le 1$, and let E_n^c denote the complement of E_n in E. Then, by bilinearity,

$$\begin{aligned}(3.1)\qquad &\left\|\sum_{i=1}^N \sum_{j=1}^N \mathbf{A}_i \boldsymbol{\beta}_d(G_i, G_j)\mathbf{A}_j^* - \sum_{i=1}^N \sum_{j=1}^N \mathbf{A}_i \boldsymbol{\beta}_n(G_i, G_j)\mathbf{A}_j^*\right\|_E \\ &= \left\|\int_\Pi\int_\Pi \boldsymbol{\Phi}\chi_{E_n}\, d\boldsymbol{\beta}\, \chi_{E_n^c}\boldsymbol{\Phi}^* + \int_\Pi\int_\Pi \boldsymbol{\Phi}\chi_{E_n^c}\, d\boldsymbol{\beta}\, \chi_{E_n}\boldsymbol{\Phi}^* \right. \\ &\qquad \left. + \int_\Pi\int_\Pi \boldsymbol{\Phi}\chi_{E_n^c}\, d\boldsymbol{\beta}\, \chi_{E_n^c}\boldsymbol{\Phi}^*\right\|_E.\end{aligned}$$

By the p.d. assumption, the triangle and Cauchy–Schwarz inequalities, (3.1) is majorized by

$$\begin{aligned}&2\left(\left\|\int_\Pi\int_\Pi \boldsymbol{\Phi}\chi_{E_n^c}\, d\boldsymbol{\beta}\, \chi_{E_n^c}\boldsymbol{\Phi}^*\right\|_E \left\|\int_\Pi\int_\Pi \boldsymbol{\Phi}\chi_{E_n}\, d\boldsymbol{\beta}\, \chi_{E_n}\boldsymbol{\Phi}^*\right\|_E\right)^{1/2} \\ &\qquad + \left\|\int_\Pi\int_\Pi \boldsymbol{\Phi}\chi_{E_n^c}\, d\boldsymbol{\beta}\, \chi_{E_n^c}\, \boldsymbol{\Phi}^*\right\|_E \\ &\qquad \le 2\left(\int_{E_n^c} \|\boldsymbol{\Phi}\|_E^2\, d\mu \int_{E_n} \|\boldsymbol{\Phi}\|_E^2\, d\mu\right)^{1/2} + \left(\int_{E_n^c} \|\boldsymbol{\Phi}\|_E^2\, d\mu\right) \quad \text{(Theorem 2.3)} \\ &\qquad \le 2\big(\mu(E_n^c)\mu(\Pi)\big)^{1/2} + \mu(E_n^c) \quad (\text{since } \|\boldsymbol{\Phi}\|_E \le 1).\end{aligned}$$

Finally, since $\boldsymbol{\beta}$ is p.d., the Euclidean norm for matrices is submultiplicative (modulo the order k), hence

$$\lim_{n\to\infty} \|\boldsymbol{\beta}_d - \boldsymbol{\beta}_n\|(\Pi, \Pi) \le k \lim\nolimits_{n\to\infty}\left(2\big(\mu(E_n^c)\mu(\Pi)\big)^{1/4} + \big(\mu(E_n^c)\big)^{1/2}\right)^2 = 0,$$

since $\mu(\cdot E_n^c) \to 0$ as $n \to +\infty$. $\boldsymbol{\beta}_d$ is thus discrete. Let $\boldsymbol{\beta}_c = \boldsymbol{\beta} - \boldsymbol{\beta}_d$, let F_1 and F_2 be two finite sets and let E be as above. Then,

$$\begin{aligned}
\boldsymbol{\beta}_c(F_1, F_2) &= \boldsymbol{\beta}(F_1, F_2) - \boldsymbol{\beta}_d(F_1, F_2) \\
&= \boldsymbol{\beta}(E \cap F_1, E \cap F_2) + \boldsymbol{\beta}(E^c \cap F_1, E \cap F_2) \\
&\quad + \boldsymbol{\beta}(E \cap F_1, E^c \cap F_2) + \boldsymbol{\beta}(E^c \cap F_1, E^c \cap F_2) - \boldsymbol{\beta}_d(F_1, F_2) \\
&= \boldsymbol{\beta}(E^c \cap F, E \cap F) + \boldsymbol{\beta}(E \cap F, E^c \cap F) + \boldsymbol{\beta}(E^c \cap F, E^c \cap F),
\end{aligned}$$

since $\boldsymbol{\beta}_d$ is the restriction of $\boldsymbol{\beta}$ to (E, E). Now, the discrete part of μ is supported on E, hence $\mu(E^c \cap F_1) = \mu(E^c \cap F_2) = 0$ and by Theorem 2.3, $\boldsymbol{\beta}(E^c \cap F, E \cap F) = \boldsymbol{\beta}(E \cap F, E^c \cap F) = \boldsymbol{\beta}(E^c \cap F, E^c \cap F) = 0$ (again, $\boldsymbol{\beta}$ is p.d.), hence $\boldsymbol{\beta}_c$ is continuous. For $G, H \in \mathscr{B}(\Pi)$ since $\boldsymbol{\beta}_c$ is continuous, $\boldsymbol{\beta}_c(G \cap E, H \cap E^c) = \boldsymbol{\beta}_c(G \cap E^c, H \cap E) = \boldsymbol{\beta}_c(G \cap E, H \cap E) = 0$, hence

$$\begin{aligned}
\boldsymbol{\beta}_c(G_i, G_j) &= \boldsymbol{\beta}_c(G_i \cap E^c, G_j \cap E^c) \\
&= \boldsymbol{\beta}(G_i \cap E^c, G_j \cap E^c) - \boldsymbol{\beta}_d(G_i \cap E^c, G_j \cap E^c) \\
&= \boldsymbol{\beta}(G_i \cap E^c, G_j \cap E^c),
\end{aligned}$$

$G_i, G_j \in \mathscr{B}(\Pi)$, and $\boldsymbol{\beta}_c$ is p.d. The uniqueness of the decomposition follows readily from Definition 3.8 and the theorem is complete. □

The decomposition of scalar bimeasures into continuous and discrete parts is due to Saeki (1976), but the above proof is directly inspired by the work of Graham and Schreiber (1984). In fact, these authors do not require any p.d. assumption, and such is also true for generic matrix bimeasures. In sharp contrast to measures, in general, $\boldsymbol{\beta}_c$ does not admit a further decomposition into a $\mathbf{L}^1(\Pi)$ matrix function and what could be called a singular continuous matrix bimeasure (Graham and Schreiber have a counterexample in the scalar case). However, in Section 5 we obtain a decomposition result for $\boldsymbol{\beta}_c$ which plays a role similar to the Lebesgue decomposition in the case of measures.

A p.d. matrix bimeasure $\boldsymbol{\beta}$ is *$(2, p)$-bounded* if it is the spectral matrix bimeasure of a $(2, p)$-bounded process, equivalently, if $\boldsymbol{\beta}$ satisfies (2.2). From Theorem 2.3, we have:

COROLLARY 3.10. *For $2 \le p < +\infty$, a $(2, p)$-bounded p.d. matrix bimeasure is continuous.*

4. Time domain prediction. Associated with a process $\mathbf{x}$ defined on $\mathbf{H}$ is a class of spaces carrying the information about this process: *The present and the past subspaces*: $\mathbf{H}_n^{\mathbf{x}} = \mathfrak{S}\{\mathbf{x}_m\colon m \le n,\ n, m \in \mathbb{Z}\} \subset \mathbf{H}$, *the remote past*: $\mathbf{H}_{-\infty}^{\mathbf{x}} = \bigcap_{n \in \mathbb{Z}} \mathbf{H}_n^{\mathbf{x}} \subset \mathbf{H}$ and *the time domain*: $\mathbf{H}_{+\infty}^{\mathbf{x}} = \mathfrak{S}\{\mathbf{x}_m\colon m \subset \mathbb{Z}\} \subset \mathbf{H}$, where $\mathfrak{S}\{\cdot\}$ denotes the closure (in the norm of $\mathbf{H}$) of the linear span (with respect to $k \times k$ complex matrix coefficients) of the corresponding set. In this framework, the orthogonal projection theorem has the following form.

LEMMA 4.1. *Let* $\mathbf{f} \in \mathbf{H}$ *and let* $\mathbf{K}$ *be a subspace of* $\mathbf{H}$, *then there exists a unique* $\hat{\mathbf{f}}$ *in* $\mathbf{H}$, *called the orthogonal projection of* $\mathbf{f}$ *onto* $\mathbf{K}$ *and denoted by* $(\mathbf{f}|\mathbf{K})$, *satisfying any of the following equivalent conditions*:

(i) $\hat{\mathbf{f}} \in \mathbf{K}$ *and* $\mathbf{f} - \hat{\mathbf{f}} \perp \mathbf{K}$.
(ii) $\hat{\mathbf{f}} \in \mathbf{K}$ *and* $[\mathbf{f} - \hat{\mathbf{f}}, \mathbf{f} - \hat{\mathbf{f}}] \leq [\mathbf{f} - \mathbf{g}, \mathbf{f} - \mathbf{g}]$ *for all* $\mathbf{g}$ *in* $\mathbf{K}$.

Since $\mathbf{K} = \otimes_{i=1}^{k} K$, *the ith component of* $\hat{\mathbf{f}}$ *is just the (ordinary) orthogonal projection of the ith component of* $\mathbf{f}$ *onto* K [*Wiener and Masani* (1957)], *and* (i) *and* (ii) *are equivalent to*

(iii) $\hat{\mathbf{f}} \in \mathbf{K}$ *and* $\|\mathbf{f} - \hat{\mathbf{f}}\|_{\mathbf{H}} \leq \|\mathbf{f} - \mathbf{g}\|_{\mathbf{H}}$ *for all* $\mathbf{g} \in \mathbf{K}$.

Since for each n, $\mathbf{H}_{n-1}$ is a subspace of $\mathbf{H}_{\mathbf{n}}$ (when no confusion seems possible we suppress the reference to the process, for example, write $\mathbf{H}_{\mathbf{n}}$ for $\mathbf{H}_{\mathbf{n}}^{\mathbf{x}}$), we have $\mathbf{x}_n = (\mathbf{x}_n|\mathbf{H}_{n-1}) + \mathbf{e}_n$. The process $\mathbf{e} = \{\mathbf{x}_n - (\mathbf{x}_n|\mathbf{H}_{n-1})\}_{n \in \mathbb{Z}}$, which is clearly orthogonal, that is, $[\mathbf{e}_n, \mathbf{e}_n] = \mathbf{\Sigma}_n \delta_{n,m}$, for all n, m in $\mathbb{Z}$, is called the *innovation (process)* of $\mathbf{x}$.

DEFINITION 4.2. A process $\mathbf{x}$ has *full rank* if $\mathbf{\Sigma}_n = [\mathbf{e}_n, \mathbf{e}_n]$ is invertible for all $n \in \mathbb{Z}$, *rank* q $(0 \leq q \leq k)$ if rank $\mathbf{\Sigma}_n = q$ for all $n \in \mathbb{Z}$ and *uniform full rank* whenever there exists a constant $C > 0$ such that $C\mathbf{I} \leq \mathbf{\Sigma}_n$ for all n in $\mathbb{Z}$.

A time-domain classification of the nature of a process is also possible. Recall that a process is *purely deterministic* if $\mathbf{H}_{-\infty} = \mathbf{H}_{+\infty}$, *purely nondeterministic* if $\mathbf{H}_{-\infty} = \{0\}$ and *nondeterministic* if $\mathbf{H}_{-\infty} \not\subseteq \mathbf{H}_{+\infty}$. In contrast to the WSS case, the above properties are not shift invariant and are only local in time. The importance of the time domain analysis is illustrated by the following well-known *Wold decomposition* due to Cramér (1961).

THEOREM 4.3. *A process* $\mathbf{x}$ *can be uniquely decomposed as* $\mathbf{x} = \mathbf{u} + \mathbf{v}$, *with the processes* $\mathbf{u} = \{\mathbf{u}_n\}_{n \in \mathbb{Z}}$ *and* $\mathbf{v} = \{\mathbf{v}_n\}_{n \in \mathbb{Z}}$ *having the following two properties*:

(i) $\mathbf{u}_n, \mathbf{v}_n \in \mathbf{H}_n^{\mathbf{x}}$ *and* $\mathbf{u}_n \perp \mathbf{v}_m$ *for all* $n, m \in \mathbb{Z}$.
(ii) *The process* $\mathbf{u}$ *is purely nondeterministic, while* $\mathbf{v}$ *is purely deterministic*.

In addition, $\mathbf{u}$ *has the property* (iii).

(iii) *For each* $n \in \mathbb{Z}$, $\mathbf{u}_n = \sum_{p=-\infty}^{n} \mathbf{A}_{n,p} \mathbf{e}_p$, *(in* $\mathbf{H}$*), where* $\mathbf{e}$ *is the innovation process*.

The processes $\mathbf{u}$ and $\mathbf{v}$ are given by $\mathbf{u}_n = (\mathbf{x}_n|\mathbf{H}_n^{\mathbf{x}} \ominus \mathbf{H}_{-\infty}^{\mathbf{x}})$ and $\mathbf{v}_n = (\mathbf{x}_n|\mathbf{H}_{-\infty}^{\mathbf{x}})$, where $\mathbf{H}_n^{\mathbf{x}} \ominus \mathbf{H}_{-\infty}^{\mathbf{x}}$ is the orthogonal complement of $\mathbf{H}_{-\infty}^{\mathbf{x}}$ with respect to $\mathbf{H}_n^{\mathbf{x}}$

and are called the *purely nondeterministic* and the *purely deterministic* components of $\mathbf{x}$. If $\sqrt{\boldsymbol{\Sigma}_p}$ denotes any square root of $\boldsymbol{\Sigma}_p$, that is, $\sqrt{\boldsymbol{\Sigma}_p}\sqrt{\boldsymbol{\Sigma}_p^*} = \boldsymbol{\Sigma}_p$, we have $\|\mathbf{u}_n\|_{\mathbf{H}}^2 = \sum_{p=-\infty}^{n} \|\mathbf{A}_{n,p}\sqrt{\boldsymbol{\Sigma}_p}\|_E^2 < +\infty$, $n \in \mathbb{Z}$. For all n and $p \leq n$, the coefficients $\mathbf{A}_{n,p}\boldsymbol{\Sigma}_p$ are uniquely determined by $\mathbf{A}_{n,p}\boldsymbol{\Sigma}_p = [\mathbf{u}_n, \mathbf{e}_p]$ and $\mathbf{A}_{n,n}\boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}_n\mathbf{A}_{n,n}^*$. Finally, the coefficients $\mathbf{A}_{n,p}$ are unique when and only when $\mathbf{x}$ has full rank, in which case $\mathbf{A}_{n,n} = \mathbf{I} = \mathbf{A}_{n,n}^*$, $n \in \mathbb{Z}$. If $\mathbf{x}$ has full rank, let $\boldsymbol{\varepsilon}_p = \sqrt{\boldsymbol{\Sigma}_p^{-1}}\,\mathbf{e}_p$, then the representation $\mathbf{x}_n = \sum_{p=-\infty}^{n} \mathbf{A}_{n,p}\sqrt{\boldsymbol{\Sigma}_p}\,\boldsymbol{\varepsilon}_p + \mathbf{v}_n$, where $[\boldsymbol{\varepsilon}_n, \boldsymbol{\varepsilon}_m] = \delta_{n,m}\mathbf{I}$, $n, m \in \mathbb{Z}$ is called the *normalized Wold decomposition* of $\mathbf{x}$. If $\mathbf{x}$ is WSS, then so are $\mathbf{e}$, $\mathbf{u}$ and $\mathbf{v}$. If $\mathbf{x}$ is $(2, p)$-bounded, so is the orthogonal process $\mathbf{e}$. Moreover, $\mathbf{u}, \mathbf{v}$ are also $(2, p)$-bounded.

THEOREM 4.4. *Let $\mathbf{x}$ be $(2, p)$-bounded, $p \leq +\infty$, with matrix bimeasure $\boldsymbol{\beta}_{\mathbf{x}}$ and let $\mathbf{u}$ and $\mathbf{v}$ be, respectively, its purely nondeterministic and purely deterministic components. Then $\mathbf{u}$ and $\mathbf{v}$ are $(2, p)$-bounded with matrix bimeasures $\boldsymbol{\beta}_{\mathbf{u}}$ and $\boldsymbol{\beta}_{\mathbf{v}}$ such that $\boldsymbol{\beta}_{\mathbf{x}} = \boldsymbol{\beta}_{\mathbf{u}} + \boldsymbol{\beta}_{\mathbf{v}}$.*

PROOF. Since the set of $(2, p)$-bounded processes is a $\mathbb{C}^{k\times k}$-module, it is enough to show that $\mathbf{u}$ or $\mathbf{v}$ is $(2, p)$-bounded. In the Wold decomposition, $\mathbf{v} = (\mathbf{x}|\mathbf{H}_{-\infty})$, hence if $\mathbf{P}_{-\infty}$ denotes the orthogonal projection operator from $\mathbf{H}$ onto $\mathbf{H}_{-\infty}^{\mathbf{x}}$, by Theorem 1.4(ii), we have

$$\mathbf{v}_n = \mathbf{P}_{-\infty}\mathbf{x}_n = \mathbf{P}_{-\infty}\int_{-\pi}^{\pi} e^{in\theta}\, d\boldsymbol{\zeta}_{\mathbf{x}}(\theta) = \int_{-\pi}^{\pi} e^{in\theta}\, d(\mathbf{P}_{-\infty}\circ\boldsymbol{\zeta}_{\mathbf{x}})(\theta).$$

Clearly, $\mathbf{P}_{-\infty}\circ\boldsymbol{\zeta}_{\mathbf{x}}$ is $(2, p)$-bounded and from Theorem 2.4 so is $\mathbf{v}$, with stochastic measure $\boldsymbol{\zeta}_{\mathbf{v}} = \mathbf{P}_{-\infty}\circ\boldsymbol{\zeta}_{\mathbf{x}}$. The orthogonality of $\mathbf{u}$ and $\mathbf{v}$ gives $\boldsymbol{\beta}_{\mathbf{x}} = \boldsymbol{\beta}_{\mathbf{u}} + \boldsymbol{\beta}_{\mathbf{v}}$. □

We wish now to characterize, among the orthogonal decompositions of the process $\mathbf{x}$, the one which is the Wold decomposition [see Masani (1966) for the stationary version]. First, for $\mathbf{x}$ $(2, p)$-bounded, we say that $\mathbf{x} = \mathbf{y} + \mathbf{z}$ is an *orthogonal decomposition* of $\mathbf{x}$ whenever $\mathbf{y}$ and $\mathbf{z}$ are also $(2, p)$-bounded with $\mathbf{y} \perp \mathbf{z}$, that is, $\mathbf{y}_n \perp \mathbf{z}_m$ for all n and m in $\mathbb{Z}$.

THEOREM 4.5. *Let $\mathbf{x} = \mathbf{y} + \mathbf{z}$ be an orthogonal decomposition of the $(2, p)$-bounded process $\mathbf{x}$. Then, $\mathbf{x} = \mathbf{y} + \mathbf{z}$ is the Wold decomposition of $\mathbf{x}$ if and only if $\mathbf{y}$ is purely nondeterministic, $\mathbf{z}$ is purely deterministic and $\mathbf{H}_n^{\mathbf{y}} \subset \mathbf{H}_n^{\mathbf{x}}$, for all n in $\mathbb{Z}$.*

PROOF. Theorems 4.3 and 4.4 give the only if part. For the if part, since $\mathbf{H}_n^{\mathbf{y}} \subset \mathbf{H}_n^{\mathbf{x}}$ and since $\mathbf{y} \perp \mathbf{z}$, we have for all n, $\mathbf{H}_n^{\mathbf{x}} = \mathbf{H}_n^{\mathbf{y}} \oplus \mathbf{H}_n^{\mathbf{z}}$, where $\oplus$ denotes the orthogonal direct sum. Taking intersections (the sum of the intersections is the intersection of the sums), we have $\mathbf{H}_{-\infty}^{\mathbf{x}} = \mathbf{H}_{-\infty}^{\mathbf{y}} \oplus \mathbf{H}_{-\infty}^{\mathbf{z}}$, and since $\mathbf{y}$ is purely nondeterministic $\mathbf{H}_{-\infty}^{\mathbf{x}} = \mathbf{H}_{-\infty}^{\mathbf{z}}$. Thus, $(\mathbf{x}|\mathbf{H}_{-\infty}^{\mathbf{x}}) = (\mathbf{z}|\mathbf{H}_{-\infty}^{\mathbf{z}}) = \mathbf{z}$, since $\mathbf{z}$ is purely deterministic. The result follows. □

From Lemma 4.1 and for any $m \geq 1$, the process $\hat{\mathbf{x}}^m = \{\hat{\mathbf{x}}_n^m = (\mathbf{x}_{n+m}|\mathbf{H}_n^{\mathbf{x}})\}_{n \in \mathbb{Z}}$, is uniquely defined and is called the (mth-*step ahead*) *prediction process* of $\mathbf{x}$ (based on its own past), while $\mathbf{x} - \hat{\mathbf{x}}^m = \{\mathbf{x}_n - \hat{\mathbf{x}}_n^m\}_{n \in \mathbb{Z}}$ is called the *prediction error process*. Now, as a direct consequence of the Wold decomposition, and since $\{\mathbf{e}_{n+1}, \dots, \mathbf{e}_{n+m}\} \perp \mathbf{H}_n^{\mathbf{x}}$ with $\mathbf{H}_n^{\mathbf{e}} \subset \mathbf{H}_n^{\mathbf{x}}$, we have $\hat{\mathbf{x}}_n^m = \sum_{p=-\infty}^{n} \mathbf{A}_{n+m,p}\mathbf{e}_p + \mathbf{v}_{n+m}$ and $\mathbf{x} - \hat{\mathbf{x}}^m = \{\sum_{p=n+1}^{n+m} \mathbf{A}_{n+m,p}\mathbf{e}_p\}_{n \in \mathbb{Z}}$ with covariance matrix sequence $\{\sum_{p=n+1}^{n+m} \mathbf{A}_{n+m,p}\mathbf{\Sigma}_p\mathbf{A}_{n+m,p}^*\}_{n \in \mathbb{Z}}$. Since this process $\mathbf{v}$ is purely deterministic, $\mathbf{v}_{n+m}$ is observed and can be predicted without error, that is, $\mathbf{v}_{n+m} \in \mathbf{H}_{-\infty}^{\mathbf{x}}$ and $(\mathbf{v}_{n+m}|\mathbf{H}_n^{\mathbf{x}}) = \mathbf{v}_{n+m}$. We thus have an elementary solution to the linear least squares estimation problem without any $(2, p)$-boundedness assumption. This is, however, unsatisfactory since we obtained the predictor $\hat{\mathbf{x}}^m$ in terms of $\mathbf{e}$ and not in terms of $\mathbf{x}$ itself. However, the use of spectral methods, as done in the rest of this work, will give such a representation. Before doing so, we state a last result due for $p = +\infty$ to Niemi and Truong-Van. Its proof is similar to the univariate proof given for $p < +\infty$ in Houdré (1990a).

THEOREM 4.6. *A process* $\mathbf{x}$ *defined on* $\mathbf{H}$ *is* $(2, p)$*-bounded*, $2 \leq p \leq +\infty$, *if and only if there exist a Hilbert module* $\mathbf{K}$ *containing* (*a submodule isometrically isomorphic to*) $\mathbf{H}$ *and a* $(2, p)$*-bounded*, $2 \leq p \leq +\infty$, *WSS process* $\mathbf{y}$ *defined on* $\mathbf{K}$ *such that* $\mathbf{x} = \mathbf{P}\mathbf{y}$, *that is*, $\mathbf{x}_n = \mathbf{P}\mathbf{y}_n$ *for all* n, *where* $\mathbf{P}$ *is the orthogonal projection from* $\mathbf{K}$ *onto* $\mathbf{H}$.

5. Spectral domain prediction. In this section, we classify the time domain properties of $(2, p)$-bounded processes by giving various characterizations and decompositions of the associated matrix bimeasures.

DEFINITION 5.1. A matrix bimeasure $\boldsymbol{\beta}$ is *nonanticipative* if for all $m > n$, $\hat{\boldsymbol{\beta}}(n, m) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{in\theta}\, d\boldsymbol{\beta}(\theta, \psi)e^{-im\psi} = 0$ and *anticipative* if for all $m < n$, $\hat{\boldsymbol{\beta}}(n, m) = 0$.

Respectively, we denote by $\mathbf{BM}$, $\mathbf{BM}_+$ and $\mathbf{BM}_-$ the sets of matrix bimeasures, nonanticipative matrix bimeasures and anticipative matrix bimeasures and by $\mathbf{BM^p}$, $\mathbf{BM^p_+}$ and $\mathbf{BM^p_-}$ the corresponding sets of $(2, p)$-bounded matrix bimeasures, $p < +\infty$.

DEFINITION 5.2. A matrix bimeasure $\boldsymbol{\beta}$ is *factorizable* if there exists a matrix bimeasure $\boldsymbol{\gamma}$ in $\mathbf{BM}_+$ such that $\hat{\boldsymbol{\beta}}(n, m) = \sum_{p=-\infty}^{\infty} \hat{\boldsymbol{\gamma}}(n, p)\hat{\boldsymbol{\gamma}}(m, p)^*$, for all n and m in $\mathbb{Z}$.

In the preceding definition, the summation extends only to $n \wedge m = \min(n, m)$ since $\boldsymbol{\gamma}$ is purely nonanticipative also, a factorizable $\boldsymbol{\beta}$ is necessarily p.d. with $\sum_{p=-\infty}^{\infty} \|\hat{\boldsymbol{\gamma}}(n, p)\|_E^2 \leq K$, for all n. The matrix bimeasure $\boldsymbol{\beta}$ is called the *smash product* of $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}^*$ and is denoted by $\boldsymbol{\gamma} \Phi \boldsymbol{\gamma}^*$. Our definition of factorizable matrix bimeasure involves involves only $\mathbb{C}^{k \times k}$-valued matrix bimeasures as opposed to $\mathbb{C}^{k\times q}$-valued ($q \le k$) ones. The latter case corresponds to processes of constant rank q; it has not been considered here in order not to overburden the notations [Rozanov (1958) has the WSS case]. We can characterize the purely nondeterministic $(2, p)$-bounded processes.

THEOREM 5.3. *A $(2,\infty)$-bounded process is purely nondeterministic if and only if its spectral matrix bimeasure is factorizable.*

PROOF. Let $\mathbf{x}$ be purely nondeterministic, then $\mathbf{x}_n = \sum_{p=-\infty}^{n} \mathbf{A}_{n,p}\mathbf{e}_p$ and $\sum_{p=-\infty}^{n} \|\mathbf{A}_{n,p}\sqrt{\mathbf{\Sigma}_p}\|_E^2 < +\infty$, where $\sqrt{\mathbf{\Sigma}_p}$ is any square root of $\mathbf{\Sigma}_p = [\mathbf{e}_p, \mathbf{e}_p]$. To prove the necessity part of the theorem, it is enough to show that the matricial (Schwartz) distribution whose Fourier coefficients are

$$\begin{Bmatrix} \mathbf{A}_{n,p}\sqrt{\mathbf{\Sigma}_p} & p \le n \\ 0 & p > n \end{Bmatrix},$$

is actually a matrix bimeasure, equivalently, that

$$\boldsymbol{\gamma} \sim \sum_{n=-\infty}^{+\infty} \sum_{p=-\infty}^{n} \mathbf{A}_{n,p}\sqrt{\mathbf{\Sigma}_p}\, e^{-in\theta} e^{ip\psi} \in \mathbf{BM}_+.$$

It is enough to show that for $\mathbf{P}(\theta) = \sum_{j=1}^{N} \mathbf{P}_j e^{-in_j\theta}$ and $\mathbf{Q}(\psi) = \sum_{p=1}^{N} \mathbf{Q}_p e^{in_p\psi}$, $\mathbf{P}_j, \mathbf{Q}_p \in \mathbb{C}^{k\times k}$, the function $\boldsymbol{\gamma}(\mathbf{P},\mathbf{Q}) = \sum_{j=1}^{N}\sum_{p=1}^{N} \mathbf{P}_j \mathbf{A}_{n_j,n_p}\sqrt{\mathbf{\Sigma}_{n_p}}\mathbf{Q}_p$ induces a bounded bi-$\mathbb{C}^{k\times k}$-linear matrix function on $\mathbf{C}(\Pi) \times \mathbf{C}(\Pi)$. But

$$(5.1) \quad \left\| \sum_{j=1}^{N}\sum_{p=1}^{N} \mathbf{P}_j \mathbf{A}_{n_j,n_p}\sqrt{\mathbf{\Sigma}_{n_p}}\mathbf{Q}_p \right\|_E^2 \le \sum_{p=1}^{N} \left\| \sum_{j=1}^{N} \mathbf{P}_j \mathbf{A}_{n_j,n_p}\sqrt{\mathbf{\Sigma}_{n_p}} \right\|_E^2 \sum_{p=1}^{N} \|\mathbf{Q}_p\|_E^2$$

and

$$(5.2) \quad \sum_{p=1}^{N} \left\| \sum_{j=1}^{N} \mathbf{P}_j \mathbf{A}_{n_j,n_p}\sqrt{\mathbf{\Sigma}_{n_p}} \right\|_E^2 = \operatorname{tr} \sum_{j=1}^{N}\sum_{q=1}^{N} \mathbf{P}_j \left(\sum_{p=1}^{N} \mathbf{A}_{n_j,n_p}\mathbf{\Sigma}_{n_p}\mathbf{A}^*_{n_q,n_p} \right) \mathbf{P}_q^*.$$

Since $\hat{\boldsymbol{\beta}}(n,m) = \sum_{p=-\infty}^{n\wedge m} \mathbf{A}_{n,p}\mathbf{\Sigma}_p\mathbf{A}^*_{m,p}$, with $\boldsymbol{\beta}$ a p.d. matrix bimeasure, (5.1) and (5.2) give

$$(5.3) \qquad \|\boldsymbol{\gamma}(\mathbf{P},\mathbf{Q})\|_E^2 \le K\|\mathbf{P}\|_\infty^2\|\mathbf{Q}\|_2^2 \le K\|\mathbf{P}\|_\infty^2\|\mathbf{Q}\|_\infty^2.$$

For the converse, let $\mathbf{w} \subset \mathbf{H}$ be an orthonormal process (a Gram–Schmidt orthogonalization method will give such a process). Since $\boldsymbol{\gamma} \in \mathbf{BM}_+$, $\mathbf{y}_n = \sum_{p=-\infty}^{n} \hat{\boldsymbol{\gamma}}(n,p)\mathbf{w}_p \in \mathbf{H}$ and $[\mathbf{y}_n, \mathbf{y}_m] = \sum_{p=-\infty}^{n\wedge m} \hat{\boldsymbol{\gamma}}(n,p)\hat{\boldsymbol{\gamma}}(m,p)^* = \hat{\boldsymbol{\beta}}(n,m) = [\mathbf{x}_n, \mathbf{x}_m]$, $n, m \in \mathbb{Z}$. Thus, $\mathbf{x}$ and $\mathbf{y}$ have the same second moments and $\mathbf{H}_n^{\mathbf{x}} \simeq \mathbf{H}_n^{\mathbf{y}}$ (the two spaces are isomorphic). Since $\mathbf{w}$ is orthogonal, $\mathbf{H}_{-\infty}^{\mathbf{w}} = \{0\}$, but $\mathbf{H}_{-\infty}^{\mathbf{x}} \simeq \mathbf{H}_{-\infty}^{\mathbf{y}} \subset \mathbf{H}_{-\infty}^{\mathbf{w}} = \{0\}$ and $\mathbf{x}$ is purely nondeterministic. □

REMARK 5.4. In the above proof, $\boldsymbol{\gamma}$ (hence the smash product of $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}^*$) has only been defined implicitly, namely, we proved the existence of a bimeasure $\boldsymbol{\gamma}$ such that $\hat{\boldsymbol{\beta}}(n,m) = \sum_{p=-\infty}^{+\infty} \hat{\boldsymbol{\gamma}}(n,p)\hat{\boldsymbol{\gamma}}^*(p,m)$, where $\hat{\boldsymbol{\gamma}}^*$ is the unique matrix bimeasure such that $\hat{\gamma}^*(p,m) = \hat{\gamma}(m,p)^*$. This is without loss of generality, since it is not difficult to prove that $\boldsymbol{\gamma}$ is the weak * limit of the sequence of matrix bimeasures $\{\sum_{q=-n}^{n}\sum_{p=-\infty}^{q}\mathbf{A}_{q,p}\sqrt{\boldsymbol{\Sigma}_p}e^{-iq\theta}e^{ip\psi}\,d\theta\,d\psi\}_n$ [see Houdré (1988) for more details]. More generally, given two matrix bimeasures $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$, the unique matrix bimeasure $\boldsymbol{\lambda}$ (when it exists) such that $\hat{\boldsymbol{\lambda}}(n,m) = \sum_{p=-\infty}^{+\infty}\hat{\boldsymbol{\rho}}(n,p)\hat{\boldsymbol{\sigma}}(p,m)$, is called the smash product of $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ and is denoted by $\boldsymbol{\lambda} = \boldsymbol{\rho} \oplus \boldsymbol{\sigma}$. The resulting $\boldsymbol{\lambda}$ can also be obtained as a weak * limit. The same comments apply if given $\boldsymbol{\lambda}$ there exist $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ such that $\hat{\boldsymbol{\lambda}}(n,m) = \sum_{p=-\infty}^{+\infty}\hat{\boldsymbol{\rho}}(n,p)\hat{\boldsymbol{\sigma}}(p,m)$ as well as to defining the smash product of a matrix bimeasure $\boldsymbol{\lambda}$ and of a stochastic measure $\boldsymbol{\zeta}$, that is, $\boldsymbol{\lambda} \oplus \boldsymbol{\zeta}$. This will be used without any further comment in the rest of this work.

When $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ are concentrated on the diagonal of Π^2, equivalently, when $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ are $\mathbf{L}^2(\Pi)$-valued functions, the smash product is the ordinary multiplication. This can be seen as follows: $\hat{\boldsymbol{\rho}}(n,p) = \hat{\boldsymbol{\rho}}(n-p)$ and $\hat{\boldsymbol{\sigma}}(n,p) = \hat{\boldsymbol{\sigma}}(n-p)$, hence $\sum_{p=-\infty}^{+\infty}\hat{\boldsymbol{\rho}}(n-p)\hat{\boldsymbol{\sigma}}(p-m) = \widehat{\boldsymbol{\rho\sigma}}(n-m)$, that is, $\boldsymbol{\rho} \oplus \boldsymbol{\sigma} = \boldsymbol{\rho\sigma}$. Heuristically, this can also be written in the following way: $\boldsymbol{\rho} \oplus \boldsymbol{\sigma}(\theta,\psi) \sim \int_{-\pi}^{\pi}\boldsymbol{\rho}(\theta,t)\boldsymbol{\sigma}(t,\psi)\,dt \sim \int_{-\pi}^{\pi}\boldsymbol{\rho}(\theta,t)\delta(\theta-t)\boldsymbol{\sigma}(t,\psi)\delta(t-\psi)\,dt \sim \boldsymbol{\rho}(\theta)\boldsymbol{\sigma}(\theta)$. Specializing Theorem 5.3 to WSS processes and denoting by $\mathbf{L}^2_+(\Pi)$ the set matrix valued functions $\boldsymbol{\gamma} \in \mathbf{L}^2(\Pi)$ such that $\hat{\boldsymbol{\gamma}}(n) = 0$ for $n < 0$, we recover a classical result of Rozanov (1958): A WSS process with positive matrix measure $\boldsymbol{\beta}$ is purely nondeterministic if and only if there exists $\boldsymbol{\gamma}$ in $\mathbf{L}^2_+(\Pi)$ such that $d\boldsymbol{\beta} = \boldsymbol{\gamma}\boldsymbol{\gamma}^*\,d\theta$.

As indicated in the remarks following Theorem 3.7, a (matrix) bimeasure does not have a Lebesgue decomposition. However, another type of decomposition is possible.

THEOREM 5.5. *A positive definite matrix bimeasure* $\boldsymbol{\beta}$ *has a unique decomposition* $\boldsymbol{\beta} = \boldsymbol{\beta}_{pc} + \boldsymbol{\beta}_{sc} + \boldsymbol{\beta}_d$, *where* $\boldsymbol{\beta}_{pc}$, $\boldsymbol{\beta}_{sc}$ *and* $\boldsymbol{\beta}_{\mathbf{d}}$ *are positive definite matrix bimeasures such that* $\boldsymbol{\beta}_{pc}$ *is factorizable,* $\boldsymbol{\beta}_{sc}$ *is continuous and not factorizable and* $\boldsymbol{\beta}_d$ *is discrete.*

PROOF. Since $\boldsymbol{\beta}$ is a p.d. matrix bimeasure, $\boldsymbol{\beta}$ is induced by a stochastic measure $\boldsymbol{\zeta}$. Let $\mathbf{x}_n = \int_{-\pi}^{\pi}e^{in\theta}\,d\boldsymbol{\zeta}(\boldsymbol{\theta})$. Then $\mathbf{x} = \{\mathbf{x}_n\}_{n\in\mathbb{Z}}$ is a $(2,\infty)$-bounded process such that $[\mathbf{x}_n, \mathbf{x}_m] = \hat{\boldsymbol{\beta}}(n,m)$. From Theorem 4.4 and 5.3, $\boldsymbol{\beta} = \boldsymbol{\beta}_{\mathbf{u}} + \boldsymbol{\beta}_{\mathbf{v}} = \boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^* + \boldsymbol{\beta}_{\mathbf{v}}$ with $\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$ and $\boldsymbol{\beta}_{\mathbf{v}}$ p.d. matrix bimeasures. But again, by Theorem 5.3, $\boldsymbol{\beta}_{\mathbf{v}} = \boldsymbol{\beta}_{\mathbf{v}c} + \boldsymbol{\beta}_{\mathbf{v}d}$, uniquely, with $\boldsymbol{\beta}_{\mathbf{v}c}$ and $\boldsymbol{\beta}_{\mathbf{v}d}$ positive definite and $\boldsymbol{\beta}_{\mathbf{v}c}$ not factorizable, since $\mathbf{v}$ is purely deterministic. Let $\boldsymbol{\beta}_{pc} = \boldsymbol{\beta}_u = \boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$, $\boldsymbol{\beta}_{sc} = \boldsymbol{\beta}_{\mathbf{v}c}$ and $\boldsymbol{\beta}_d = \boldsymbol{\beta}_{\mathbf{v}d}$, then $\boldsymbol{\beta} = \boldsymbol{\beta}_{pc} + \boldsymbol{\beta}_{sc} + \boldsymbol{\beta}_d$, uniquely. □

$\boldsymbol{\beta}_{pc}$, $\boldsymbol{\beta}_{sc}$ and $\boldsymbol{\beta}_d$ are, respectively, called the *purely continuous*, *singular continuous* and *discrete component* of $\boldsymbol{\beta}$. For (2, p)-bounded processes $p < +\infty$, $\boldsymbol{\beta} = \boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^* + \boldsymbol{\beta}_{sc}$, with $\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$ and $\boldsymbol{\beta}_{sc}$ in $\mathbf{BM^p}$, and hence is continuous

(Theorem 4.4 and Corollary 3.10). For multivariate WSS processes, $\boldsymbol{\beta} = \boldsymbol{\gamma\gamma}^* + \boldsymbol{\beta}_{sc} + \boldsymbol{\beta}_d$ [Cramér (1940)] and $\boldsymbol{\gamma\gamma}^*$ is again continuous. Although it is reasonable to expect that for general $(2, \infty)$-bounded processes $\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$ is also continuous, we could not prove it. Nevertheless, in view of the $(2, p)$-bounded $(p < +\infty)$ and WSS cases, we will keep the terminology purely continuous for $\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$.

REMARK 5.6. Our method of proof for Theorem 5.5 relies on both bimeasure theoretic arguments and on the Wold decomposition. It is of interest to find a pure bimeasure theoretic proof of Theorem 5.5. This should give new insights on the decomposition of (matrix) bimeasures. If such a proof is obtained, it is not difficult to see that there is concordance, for processes of constant rank, between the decomposition of Theorem 5.5 and the Wold decomposition. This, of course, parallels the WSS constant rank case.

To close this prediction in the spectral domain section, we relate the prediction error to the spectral matrix bimeasure. Let $\boldsymbol{\beta}$ be a spectral bimeasure whose factorizable part is $\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$. We denote by $\boldsymbol{\gamma}_n$ the "nth *Fourier marginal*" of $\boldsymbol{\gamma}$, that is, $\boldsymbol{\gamma}_n(\psi) = \sum_{p=+\infty}^{n} \hat{\boldsymbol{\gamma}}(n,p) e^{ip\psi}$.

THEOREM 5.7. *Let* $\mathbf{x}$ *be a* $(2, \infty)$*-bounded full rank process, then* $\log|\det \boldsymbol{\gamma}_n| \in L^1(\Pi)$, $n \in \mathbb{Z}$.

PROOF. Let $\boldsymbol{\Sigma}_n > 0$, then $\boldsymbol{\gamma}_n(\psi) = e^{in\psi} \sum_{p=0}^{+\infty} \mathbf{A}_{n,n-p} \sqrt{\boldsymbol{\Sigma}_{n-p}}\, e^{-ip\psi}$. Since $\tilde{\boldsymbol{\gamma}}_n(\psi) = \sum_{p=0}^{+\infty} \hat{\boldsymbol{\gamma}}(n, n-p) e^{-ip\psi} \in \mathbf{L}^2_+(\Pi)$, $|\det \tilde{\boldsymbol{\gamma}}_n|^{1/k} = |\det \boldsymbol{\gamma}_n|^{1/k} \in L^2_+(\Pi)$. If $\tilde{\boldsymbol{\gamma}}_{\mathbf{n}}(z)$ denotes the analytic extension of $\tilde{\boldsymbol{\gamma}}_n(\psi)$ inside the unit disc, Jensen's inequalities with $\mathbf{A}_{n,n} = \mathbf{I}$ give

$$-\infty < \tfrac{1}{2} \log \det \boldsymbol{\Sigma}_n = \log|\det \tilde{\boldsymbol{\gamma}}_n(0)| \le \int_{-\pi}^{\pi} \log|\det \boldsymbol{\gamma}_n(\psi)|\, d\psi$$

$$= \tfrac{1}{2} \int_{-\pi}^{\pi} \log \det \boldsymbol{\gamma}_n(\psi) \boldsymbol{\gamma}_n^*(\psi)\, d\psi \le \tfrac{1}{2} \log \det \int_{-\pi}^{\pi} \boldsymbol{\gamma}_n(\psi) \boldsymbol{\gamma}_n^*(\psi)\, d\psi < +\infty. \quad \square$$

Theorem 5.7 only partially recovers a WSS result of Wiener and Masani which was the basis for our proof. It is our belief that the converse in the above theorem holds, in other words, that the $\tilde{\boldsymbol{\gamma}}_n$'s are outer matrix functions for all n. We could not prove it in full generality. If true, it is then not difficult to see, in the equivalent minimizing problem, that the contribution from $\boldsymbol{\beta}_{sc}$ and $\boldsymbol{\beta}_d$ are both null. Hence,

$$\exp \int_{-\pi}^{\pi} \log \det \boldsymbol{\gamma}_n(\psi) \boldsymbol{\gamma}_n^*(\psi)\, d\psi$$

$$= \inf_{\mathbf{P}} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \det\left(e^{in\theta} - \mathbf{P}_n(\theta)\right) d\boldsymbol{\beta}(\theta, \psi) \left(e^{in\psi} - \mathbf{P}_n(\psi)\right)^*.$$

6. Spectral and autoregressive predictors. For WSS processes a solution to the prediction problem, per se, involves a spectral factorization and an inversion formula. Such is the case for $(2,\infty)$-bounded processes. We first investigate conditions under which inversion is possible.

THEOREM 6.1. *Let* $\mathbf{x}$ *be a purely nondeterministic* $(2,\infty)$*-bounded process with spectral matrix bimeasure* $\boldsymbol{\beta}$. *Then,* $\boldsymbol{\beta}$ *is invertible in* **BM** *if and only if there exists a constant* $K_1 > 0$ *such that*

$$(6.1)\qquad K_1\|\boldsymbol{\Phi}\|_\infty \le \left(\operatorname{tr}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\boldsymbol{\Phi}(\theta)\,d\boldsymbol{\beta}(\theta,\psi)\boldsymbol{\Phi}^*(\psi)\right)^{1/2}$$

for all $\boldsymbol{\Phi}$ *in* $\mathbf{C}(\Pi)$.

PROOF. Since $\mathbf{x}$ is purely nondeterministic, $d\boldsymbol{\beta} = d\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$, with $\boldsymbol{\gamma}$ purely nonanticipative. Let

$$\mathbf{T}\colon \mathbf{P}(\theta) = \sum_{j=1}^{N}\mathbf{P}_j e^{in_j\theta} \to (d/d\psi)\int_{-\pi}^{\pi}\mathbf{P}(\theta)\,d\boldsymbol{\gamma}(\theta,\psi)$$
$$= \sum_{j=1}^{N}\mathbf{P}_j\sum_{p=-\infty}^{n_j}\boldsymbol{\gamma}(n_j,p)e^{ip\psi},$$

where $d/d\psi$ denotes Radon–Nikodym derivative. Then $\mathbf{T}$ is well defined, $\mathbb{C}^{k\times k}$-linear and

$$\int_{-\pi}^{\pi}\operatorname{tr}\frac{d}{d\psi}\int_{-\pi}^{\pi}\mathbf{P}(\theta)\,d\boldsymbol{\gamma}(\theta,\psi)\frac{d}{d\psi}\int_{-\pi}^{\pi}\mathbf{P}(\theta)\,d\boldsymbol{\gamma}(\theta,\psi)^*\,d\psi$$
$$= \operatorname{tr}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\mathbf{P}(\theta)\,d\boldsymbol{\gamma}\oplus\boldsymbol{\gamma}^*(\theta,\psi)\mathbf{P}^*(\psi)$$
$$= \operatorname{tr}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\mathbf{P}(\theta)\,d\boldsymbol{\beta}(\theta,\psi)\mathbf{P}^*(\psi) \le K_2\|\mathbf{P}\|_\infty^2.$$

$\mathbf{T}$ can be extended to a bounded $\mathbb{C}^{k\times k}$-linear operator from $\mathbf{C}(\Pi)$ to $\mathscr{R}\subset\mathbf{L}^2(\Pi)$. Readily, $\boldsymbol{\beta}$ is invertible in **BM** if and only if $\mathbf{T}$ is invertible (as bounded operator), namely, if and only if $\|\mathbf{T}(\mathbf{P})\|_2 \ge K_1\|\mathbf{P}\|_\infty$ and if $\mathscr{R}$ is dense in $\mathbf{L}^2(\Pi)$. So we just have to show that $\mathscr{R}$ is dense in $\mathbf{L}^2(\Pi)$. If $\overline{\mathscr{R}} \ne \mathbf{L}^2(\Pi)$, then by Lemma 4.1 there exists $\boldsymbol{\Phi}\in\mathbf{L}^2(\Pi)$ such that $\boldsymbol{\Phi}\ne 0$ and $\boldsymbol{\Phi}\perp\overline{\mathscr{R}}$. In particular, for any matrix trigonometric polynomial $\mathbf{P}$ we have $\int_{-\pi}^{\pi}(\mathbf{TP})(\theta)\boldsymbol{\Phi}^*(\theta)\,d\theta = 0$. $\mathbf{C}(\Pi)$ is dense in $\mathbf{L}^2(\Pi)$, hence there exists $\tilde{\boldsymbol{\Phi}}\in\mathbf{C}(\Pi)$ such that $\|\boldsymbol{\Phi}-\tilde{\boldsymbol{\Phi}}\|_2 \le \varepsilon$. It thus follows that

$$\left\|\int_{-\pi}^{\pi}(\mathbf{TP})(\theta)\boldsymbol{\Phi}^*(\theta)\,d\theta - \int_{-\pi}^{\pi}(\mathbf{TP})(\theta)\tilde{\boldsymbol{\Phi}}^*(\theta)\,d\theta\right\|_E \le \varepsilon\|\mathbf{T}\|\|\mathbf{P}\|_\infty$$

and $\int_{-\pi}^{\pi}(\mathbf{TP})(\theta)\mathbf{\Phi}^*(\theta)\,d\theta = 0$ if and only if $\int_{-\pi}^{\pi}(\mathbf{TP})(\theta)\tilde{\mathbf{\Phi}}^*(\theta)\,d\theta = 0$. With $\mathbf{P}(\cdot) = e^{in\cdot}$, $n \in \mathbb{Z}$, we get $\sum_{p=-\infty}^{n}\gamma(n,p)\hat{\tilde{\mathbf{\Phi}}}(p)^* = 0$, for all n, and $\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}e^{in\theta}\,d\boldsymbol{\beta}(\theta,\psi)\tilde{\mathbf{\Phi}}^*(\psi) = 0$, also for all n. Since $\tilde{\mathbf{\Phi}} \in \mathbf{C}(\Pi)$ and since $\boldsymbol{\beta}$ is p.d. and belongs to **BM**, dominated convergence and the density of matrix trigonometric polynominals (for vector measures) give

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\tilde{\mathbf{\Phi}}(\theta)\,d\boldsymbol{\beta}(\theta,\psi)\tilde{\mathbf{\Phi}}^*(\psi) = 0;$$

hence $\tilde{\mathbf{\Phi}} = 0$ and $\mathbf{\Phi} = 0$ (a.s. Lebesgue). □

The crucial information in Theorem 6.1 is that $\boldsymbol{\beta}^{-1}$ is a $(2,\infty)$-bounded bimeasure with moreover $(\boldsymbol{\beta}^{-1})^* = (\boldsymbol{\beta}^*)^{-1} \in \mathbf{BM}$. If (6.1) does not hold, this is generally not true, even for simple processes. If $\mathbf{w}$ is an orthogonal process such that $[\mathbf{w}_n, \mathbf{w}_n] = (1/(|n|+1))\mathbf{I}$, then its spectral matrix bimeasure $\boldsymbol{\beta}_{\mathbf{w}}$ has an algebraic inverse which is *not* a matrix bimeasure. For $p < +\infty$, the inversion problem has a different solution. By essentially the same techniques as above, the following can be shown: If the left-hand side of (6.1) is replaced by $K_1(\int_{-\pi}^{\pi}\|\mathbf{\Phi}(\theta)\|_E^r\,d(\theta)^{1/r}$, $1 \le r \le 2$, $\boldsymbol{\beta}$ becomes invertible in $\mathbf{BM}^{\mathbf{r}}$. For $r > 2$, inversion in $\mathbf{BM}^{\mathbf{r}}$ is impossible, unless $r = +\infty$ in which case $\boldsymbol{\beta}$ is invertible in **BM**. Spectral factorization and invertibility can now be combined in a single result.

COROLLARY 6.2. *Let* $\mathbf{x}$ *be a purely nondeterministic* $(2,\infty)$*-bounded process with spectral matrix bimeasure* $\boldsymbol{\beta}$. *Then,* $d\boldsymbol{\beta} = d\boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^*$ *with* $\boldsymbol{\gamma}$ *and* $\boldsymbol{\gamma}^{-1}$ *in* $\mathbf{BM}_+$ *if and only if* $\boldsymbol{\beta}$ *is invertible in* **BM**. *Moreover, the factorization is unique up to the equivalence under unitary diagonal matrix bimeasures.*

PROOF. From Theorem 5.3 and 6.1, we just have to prove the uniqueness. Let $\boldsymbol{\alpha} \oplus \boldsymbol{\alpha}^*$ be another factorization. First, $\hat{\boldsymbol{\sigma}}(n,m) = \widehat{\boldsymbol{\alpha}^{-1} \oplus \boldsymbol{\gamma}}(n,m) = \sum_{p=m}^{n}\hat{\boldsymbol{\alpha}}^{-1}(n,p)\hat{\boldsymbol{\gamma}}(p,m)$ for $m \le n$, and 0 for $m > n$, since $\boldsymbol{\alpha}^{-1}$ and $\boldsymbol{\gamma} \in \mathbf{BM}_+$. But $\boldsymbol{\sigma} \oplus \boldsymbol{\sigma}^* = \boldsymbol{\alpha}^{-1} \oplus \boldsymbol{\gamma} \oplus \boldsymbol{\gamma}^* \oplus \boldsymbol{\alpha}^{-1*} = \boldsymbol{\alpha}^{-1} \oplus \boldsymbol{\beta} \oplus \boldsymbol{\alpha}^{-1} = \boldsymbol{\alpha}^{-1} \oplus \boldsymbol{\alpha} \oplus \boldsymbol{\alpha}^* \oplus \boldsymbol{\alpha}^{-1*} = \boldsymbol{\delta} = \boldsymbol{\sigma}^* \oplus \boldsymbol{\sigma}$, hence $\boldsymbol{\sigma}^* = \boldsymbol{\sigma}^{-1}$. But $\hat{\boldsymbol{\sigma}}^*(n,m) = 0$ for $m \ge n$ while since $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}^{-1} \in \mathbf{BM}_+$, $\hat{\boldsymbol{\sigma}}^{-1}(n,m) = 0$ for $m \le n$. Thus, $\hat{\boldsymbol{\sigma}}(n,m) = 0$ for $n \ne m$, that is, σ is diagonal. Finally, $\hat{\boldsymbol{\sigma}}(n,n)\hat{\boldsymbol{\sigma}}(n,n)^* = \hat{\boldsymbol{\sigma}}(n,n)^*\hat{\boldsymbol{\sigma}}(n,n) = \mathbf{I}$ give the unitary property. □

A similar corollary holds for $\boldsymbol{\beta}$ dominated (above and below) by p and r norms, for which inversion is possible, with a slight difference: $\boldsymbol{\beta}$ is factorizable (from Theorem 4.6, for $p < +\infty$, $\mathbf{x}$ has a stationary dilation $\mathbf{y}$ with $\mathbf{H}_{-\infty}^{\mathbf{x}} \subset \mathbf{H}_{-\infty}^{\mathbf{y}}$, but the spectrum of $\mathbf{y} \in \mathbf{L}^{\mathbf{p}/(\mathbf{p}-\mathbf{2})}(\Pi)$ so $\mathbf{H}_{-\infty}^{\mathbf{y}} = \{0\}$ and the factorizability). Rissanen and Barbosa (1969) obtained for bounded linear operators, that is, in the (2, 2)-bounded case, a sufficient condition for invertible spectral factorization, which was an early motivation for our work.

Combining the above results we find *an explicit spectral formula* for the predictor $\hat{\mathbf{x}}$. A last notation: We denote the nonanticipative truncation operation by $[\cdot]_+$, that is, for a matricial Schwartz distribution

$$\boldsymbol{\chi}(\theta,\psi) \sim \sum_{n=-\infty}^{+\infty}\sum_{p=-\infty}^{+\infty} \mathbf{A}_{n,p}e^{-in\theta}e^{ip\psi},$$

$$[\boldsymbol{\chi}(\theta,\psi)]_+ \sim \sum_{n=-\infty}^{+\infty}\sum_{p=-\infty}^{n} \mathbf{A}_{n,p}e^{-in\theta}e^{ip\psi}.$$

THEOREM 6.3. *Let* $\mathbf{x}$ *be a purely nondeterministic* $(2,\infty)$*-bounded process with spectral stochastic measure* $\boldsymbol{\zeta}_{\mathbf{x}}$, *and let its spectral matrix bimeasure* $\boldsymbol{\beta}_{\mathbf{x}}$ *satisfy the condition* (6.1). *Then, for each* m *in* $\mathbb{N}^*$, $\hat{\mathbf{x}}^m = \{\hat{\mathbf{x}}_n^m = (\mathbf{x}_{n+m}|\mathbf{H}_n^{\mathbf{x}})\}_{n\in\mathbb{Z}}$ *is* $(2,\infty)$*-bounded and uniquely given for each* n *by* $\hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi}e^{in\theta}\,d\boldsymbol{\zeta}_{\hat{\mathbf{x}}^m}(\theta)$, *where* $d\boldsymbol{\zeta}_{\hat{\mathbf{x}}^m} = d[e^{im\cdot}\boldsymbol{\gamma}]_+ \,\Phi\, \boldsymbol{\gamma}^{-1}\,\Phi\,\boldsymbol{\zeta}_{\mathbf{x}}$.

PROOF. First, a $(2,\infty)$-bounded process which satisfies (6.1) has uniform full rank since for $\mathbf{P}(\theta) = e^{in\theta}\sum_{p=1}^{N}\mathbf{P}_p e^{-ip\theta}$,

$$\begin{aligned}[\mathbf{e}_n,\mathbf{e}_n] &= \inf_{\mathbf{P}}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(e^{in\theta}-\mathbf{P}(\theta))\,d\boldsymbol{\beta}(\theta,\psi)(e^{in\psi}-\mathbf{P}(\psi))^*\\ &\geq K_1\inf_{\mathbf{P}}\int_{-\pi}^{\pi}(\mathbf{I}-\mathbf{P}(\theta))(\mathbf{I}-\mathbf{P}(\theta))^*\,d\theta\\ &= K_1\inf\left(\mathbf{I}+\sum_{p=1}^{N}\mathbf{P}_p\mathbf{P}_p^*\right) = K_1\mathbf{I}.\end{aligned}$$

Since $\mathbf{x}$ is purely nondeterministic, $\boldsymbol{\beta}_{\mathbf{x}} = \boldsymbol{\gamma}\,\Phi\,\boldsymbol{\gamma}^*$, $\boldsymbol{\gamma}\in\mathbf{BM}_+$ and

$$(e^{im\cdot}\boldsymbol{\gamma})(\theta,\psi) \sim \sum_{n=-\infty}^{+\infty}\sum_{p=-\infty}^{+\infty}\mathbf{A}_{n,p}\sqrt{\boldsymbol{\Sigma}_p}\,e^{-i(n-m)\theta}e^{ip\psi}\in\mathbf{BM},$$

we thus have

$$[e^{im\cdot}\boldsymbol{\gamma}]_+(\theta,\psi) \sim \sum_{n=-\infty}^{+\infty}\sum_{p=-\infty}^{n}\mathbf{A}_{n+m,p}\sqrt{\boldsymbol{\Sigma}_p}\,e^{-in\theta}e^{ip\psi}\in\mathbf{BM}_+$$

and $[e^{im\cdot}\boldsymbol{\gamma}]_+\,\Phi\,[e^{im\cdot}\boldsymbol{\gamma}]_+^*\in\mathbf{BM}$. Since $\mathbf{x}$ has full rank and from the remark preceding Theorem 4.6, $\hat{\mathbf{x}}_n^m = \sum_{p=-\infty}^{n}\mathbf{A}_{n+m,p}\sqrt{\boldsymbol{\Sigma}_p}\,\boldsymbol{\varepsilon}_p$ and $\hat{\mathbf{x}}$ is $(2,\infty)$-bounded [(1.1) is satisfied] so $\hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi}e^{in\theta}\,d\boldsymbol{\zeta}_{\hat{\mathbf{x}}^m}(\theta)$. $\sum_{p=-\infty}^{n}\mathbf{A}_{n+m,p}\sqrt{\boldsymbol{\Sigma}_p}\,e^{ip\theta}\in\mathbf{L}^2(\Pi)$, with $\boldsymbol{\varepsilon}$ orthonormal, so interchanging sum and integral we get

$$\hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi}\sum_{p=-\infty}^{n}\mathbf{A}_{n+m,p}\sqrt{\boldsymbol{\Sigma}_p}\,e^{ip\theta}\,d\boldsymbol{\zeta}_{\boldsymbol{\varepsilon}} = \int_{-\pi}^{\pi}e^{in\theta}\,d\big([e^{im\cdot}\boldsymbol{\gamma}]_+\,\Phi\,\boldsymbol{\zeta}_{\boldsymbol{\varepsilon}}\big)(\theta),$$

where $[e^{im\cdot}\boldsymbol{\gamma}]_+\,\Phi\,\boldsymbol{\zeta}_{\boldsymbol{\varepsilon}}$ is well defined because $[e^{im\cdot}\boldsymbol{\gamma}]_+\in\mathbf{BM}_+$. Similarly, $\mathbf{x}_n = \int_{-\pi}^{\pi}e^{in\theta}\,d\boldsymbol{\zeta}_{\mathbf{x}}(\theta) = \int_{-\pi}^{\pi}e^{in\theta}\,d(\boldsymbol{\gamma}\,\Phi\,\boldsymbol{\zeta}_{\boldsymbol{\varepsilon}})(\theta)$. The uniqueness of the Fourier transform gives $\zeta_{\hat{\mathbf{x}}^m} = [e^{im\cdot}\gamma]_+ \Phi\, \zeta_\varepsilon$ and $\zeta_{\mathbf{x}} = \gamma \Phi\, \zeta_\varepsilon$. Finally $\beta_{\mathbf{x}}$, equivalently γ, is invertible in **BM** and $[e^{im\cdot}\gamma]_+ \Phi\, \gamma^{-1} \Phi\, \zeta_{\mathbf{x}}$ is well defined with $\zeta_{\hat{\mathbf{x}}^m} = [e^{im\cdot}\gamma]_+ \Phi\, \gamma^{-1} \Phi\, \zeta_{\mathbf{x}}$. □

In the above result, we have $[\hat{\mathbf{x}}_n^m, \hat{\mathbf{x}}_p^m] = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{in\theta}\, d\beta^m(\theta,\psi) e^{-ip\psi}$, where $d\beta^m = d[e^{im\cdot}\gamma]_+ \Phi [e^{im\cdot}\gamma]_+^*$ and $\mathbf{x}_n - \hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi} e^{in\theta}\, d[(\delta - [e^{im\cdot}\gamma]_+ \Phi\, \gamma^{-1}) \Phi\, \zeta_{\mathbf{x}}](\theta)$. For β invertible in **BM** or $\mathbf{BM^r}$, $1 \le r \le 2$ and $p < +\infty$, a similar result holds, moreover $\hat{\mathbf{x}} \in \mathscr{M}^{\mathbf{p}}$.

We close this work by studying the important theoretical and practical problem of obtaining for $\hat{\mathbf{x}}$ an autoregressive series representation. This relies, first, on the possibility of transfering methods from the spectral domain of the process to its time domain. As is well known, this transfer is given by the Kolmogorov isomorphism, which asserts that the correspondence $\Phi \to \int_{-\pi}^{\pi} \Phi(\theta)\, d\zeta(\theta)$ is an isomorphism from $\mathbf{L}^2(\beta)$ onto $\mathbf{H}_{+\infty}$ if and only if $\mathbf{L}^2(\beta)$ is complete [see Truong-Van (1981) and Remark 3.7]. Hence, with a completeness assumption, the problem of obtaining autoregressive representations reduces to one of convergence of Fourier series in $\mathbf{L}^2(\beta)$.

A natural requirement to solve this Fourier series problem is again a boundedwise condition à la (6.1), namely, that there exists a constant $K > 0$ such that

$$K\|\Phi\|_2 \le \left(\operatorname{tr} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \Phi(\theta)\, d\beta(\theta,\psi) \Phi^*(\psi)\right)^{1/2} \tag{6.2}$$

for all Φ in $\mathbf{C}(\Pi)$. For $(2,2)$-bounded processes, (6.2) also ensures that $\mathbf{L}^2(\beta) = \mathbf{L}^2(\Pi)$ hence the completeness, while in the $(2,2)$-bounded WSS case it translates to β and $\beta^{-1} \in \mathbf{L}^\infty(\Pi)$, which is Wiener and Masani's original boundedness condition.

THEOREM 6.4. *Let* $\mathbf{x}$ *be a* $(2, p)$*-bounded process,* $1 \le p \le 2$, *whose spectral domain is complete and whose spectral matrix bimeasure* β *satisfies* (6.2). *Then,* $\hat{\mathbf{x}}_n^m = \sum_{s=-\infty}^{n} \hat{\gamma}^m(n,s)\mathbf{x}_s$, *where* $\gamma^m = [e^{im\cdot}\gamma]_+ \Phi\, \gamma^{-1}$ *and where the convergence is in* $\mathbf{H}_{+\infty}$. *Equivalently,* $\hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi} \kappa_n^m(\theta)\, d\zeta_{\mathbf{x}}(\theta)$, *where* $\kappa_n^m(\theta) = \sum_{s=-\infty}^{n} \hat{\gamma}^m(n,s) e^{is\theta}$ *in* $\mathbf{L}^2(\beta)$.

PROOF. First, $\mathbf{x}$ is purely nondeterministic (see the comments after Corollary 6.2) and $\beta = \gamma \Phi\, \gamma^*$; moreover, by Kolmogorov's isomorphism, the convergence of the two infinite series is equivalent. Since $\hat{\mathbf{x}} \in \mathbf{H}_{+\infty}$, we have $\hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi} \kappa_n^m(\theta)\, d\zeta_{\mathbf{x}}(\theta)$ for some κ_n^m in $\mathbf{L}^2(\beta)$. By Theorem 6.3 and the uniqueness of the Fourier transform, we also have

$$\kappa_n^m(\psi) = (d/d\psi) \int_{-\pi}^{\pi} d\big([e^{im\cdot}\gamma]_+ \Phi\, \gamma^{-1}\big)(\theta,\psi),$$

with by (6.2) and $(2,p)$-boundedness, $|\int_{-\pi}^{\pi} \mathbf{P}\kappa_n^m\, d\theta| \le K\|\mathbf{P}\|_p$, for all matrix trigonometric polynomial $\mathbf{P}$, hence as in the proof of Theorem 1.2, $\kappa_n^m \in \mathbf{L}^{\mathbf{p}/\mathbf{p}-1}(\Pi) \subset \mathbf{L}^1(\Pi)$. In particular, $\hat{\kappa}_n^m(s)$, the sth Fourier coefficient of κ_n^m, is well defined and equal to $\hat{\gamma}^m(n,s)$. Finally, since for $1 \le p \le 2$, $\mathbf{L}^{\mathbf{p}/\mathbf{p}-\mathbf{1}}(\Pi) \subset \mathbf{L}^{\mathbf{p}}(\Pi)$, we get

$$\begin{aligned}
&\left\| \hat{\mathbf{x}}_n^m - \sum_{s=-N}^{n} \hat{\gamma}^m(n,s)\mathbf{x}_s \right\|_{\mathbf{H}}^2 \\
&\quad = \operatorname{tr} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \left(\boldsymbol{\kappa}_n^m(\theta) - \sum_{s=-N}^{n} \hat{\gamma}^m(n,s)e^{is\theta} \right) d\boldsymbol{\beta}_{\mathbf{x}}(\theta,\psi) \\
&\quad\quad \times \left(\boldsymbol{\kappa}_n^m(\theta) - \sum_{s=-N}^{n} \hat{\gamma}^m(n,s)e^{is\theta} \right)^* \\
&\quad \le K \left(\int_{-\pi}^{\pi} \left\| \boldsymbol{\kappa}_n^m(\theta) - \sum_{s=-N}^{n} \hat{\gamma}^m(n,s)e^{is\theta} \right\|_E^p d\theta \right)^{2/p} \\
&\quad \le K \left(\int_{-\pi}^{\pi} \left\| \boldsymbol{\kappa}_n^m(\theta) - \sum_{s=-N}^{n} \hat{\gamma}^m(n,s)e^{is\theta} \right\|_E^{p/p-1} d\theta \right)^{2(p-1)/p}.
\end{aligned}$$

Fourier series of $\mathbf{L}^{\mathbf{p}/\mathbf{p}-\mathbf{1}}(\Pi)$ functions converge in $\mathbf{L}^{\mathbf{p}/\mathbf{p}-\mathbf{1}}(\Pi)$, hence the result. □

When applied to WSS processes, the proof of Theorem 6.4 provides a new set of conditions to have autoregressive predictors, that is, a physically realizable set of filters [Masani (1960) also has a different set of conditions].

COROLLARY 6.5. *Let* $\mathbf{x}$ *be a WSS process with* $\boldsymbol{\beta}$ *in* $\mathbf{L}^{\mathbf{p}}(\Pi)$, $2 \le p \le +\infty$, *and* $\boldsymbol{\beta}^{-1}$ *in* $\mathbf{L}^{+\infty}(\Pi)$. *For all* n,

$$\hat{\mathbf{x}}_n^m = \int_{-\pi}^{\pi} e^{in\theta}\left[e^{im\theta}\boldsymbol{\gamma}(\theta)\right]_+ \boldsymbol{\gamma}^{-1}(\theta)\, d\boldsymbol{\zeta}_{\mathbf{x}}(\theta) = \sum_{s=0}^{\infty} \widehat{\left[e^{im\cdot}\boldsymbol{\gamma}\right]_+ \boldsymbol{\gamma}^{-1}}(s)\mathbf{x}_{n-s},$$

in $\mathbf{H}_{+\infty}$.

Clearly, a nonanticipative autoregressive convolution representation exists when and only when the Fourier series of $[e^{im\cdot}\boldsymbol{\gamma}]_+\boldsymbol{\gamma}^{-1}$ converges in $\mathbf{L}^2(\boldsymbol{\beta})$. The particular form of $[e^{im\cdot}\boldsymbol{\gamma}]_+\boldsymbol{\gamma}^{-1}$ makes the conditions of Corollary 6.5 independent of the Helson and Szegö (1960) characterization of measures for which the exponentials form a basis [see Houdré (1988)] for counterexamples in the case $p = 2$. If in (6.2), $\|\boldsymbol{\Phi}\|_2$ is replaced by $\|\boldsymbol{\Phi}\|_r$, $1 \le r < 2$, a result similar to Theorem 6.4 holds. However, the computations involved become quite cumbersome and this will not be stated. The problem of estimating a process based on the past of another process admits also a similar solution using the techniques developed here.

7. Some questions and problems. Many questions remain unanswered, and many problems untouched, in both theory and applications. This is illustrated by the following nonexhaustive set of problems and questions which have only been partially raised and solved, when not simply ignored. Foremost among these is the prospective study of the relationship between analytic functions of two complex variables and Fourier series of bimeasures. This in turn calls for a better understanding of the decomposition of bimeasures and the possibility of obtaining a Szegö-type formula for the prediction error. Such questions lead also to problems such as minimality, interpolability, subordination, characterization of constant rank processes and so forth. The study of the shift operator (when it exists) for $(2, p)$-bounded processes is another topic on its own. On the filtering scene, an algorithmic analysis and its comparison to Kalman's approach is desirable [this has been initiated in Houdré (1990b)]. Finally, the problem of finding minimal conditions to obtain autoregressive predictors is still open, even in the stationary case.

REFERENCES

BARTLE, R. G. (1956). A general bilinear vector integral. *Studia Math.* **15** 337–352.

BARTLE, R. G., DUNFORD, N. and SCHWARTZ, J. T. (1955). Weak compactness and vector measures. *Canad. J. Math.* **7** 289–305.

BOCHNER, S. (1956). Stationarity, boundedness, almost periodicity of random valued functions. In *Proc. Third Berkeley Symp. Math. Statist. Probab.* **2** 7–27. Univ. California Press, Berkeley.

CRAMÉR, H. (1940). On the theory of stationary random processes. *Ann. of Math.* (*2*) **41** 215–230.

CRAMÉR, H. (1961). On some classes of nonstationary stochastic processes. In *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **2** 55–77. Univ. California Press, Berkeley.

DUNFORD, N. and SCHWARTZ, J. T. (1957). *Linear Operators, Part I: General Theory.* Interscience, New York.

GRAHAM, C. C. and SCHREIBER, B. M. (1984). Bimeasure algebras on LCA groups. *Pacific J. Math.* **115** 91–127.

GROTHENDIECK, A. (1953). Sur les applications linéaires faiblement compactes d'espaces du type $C(K)$. *Canad. J. Math.* **5** 129–173.

HELSON, H. and LOWDENSLAGER, D. (1959). Prediction theory and Fourier series in several variables, I. *Acta Math.* **99** 165–202.

HELSON, H. and LOWDENSLAGER, D. (1961). Prediction theory and Fourier series in several variables, II. *Acta Math.* **106** 175–213.

HELSON, H. and SZEGÖ G. (1960). A problem in prediction theory. *Ann. Mat. Pura Appl.* (4) **51** 107–138.

HOUDRÉ, C. (1988). Nonstationary processes, system theory and prediction. Ph.D. dissertation, McGill Univ.

HOUDRÉ, C. (1989). A vector bimeasure integral with some applications. Technical Report 214, Center for Stochastic Processes, Univ. North Carolina.

HOUDRÉ, C. (1990a). Harmonizability, V-boundedness, $(2, p)$-boundedness of stochastic processes. *Probab. Theory Related Fields* **84** 39–54.

HOUDRÉ, C. (1990b). Factorization algorithms and non-stationary Wiener filtering. *Stochastics.* To appear.

KAILATH, T. (1974). A view of three decades of linear filtering theory. *IEEE Trans. Inform. Theory* **IT-20** 146–181.

KALMAN, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82** 34–45.

KOLMOGOROV, A. N. (1941). Stationary sequences in Hilbert space. *Bull. Math. Univ. Moscow* **2** (6) 1–40. (In Russian.)

MASANI, P. (1960). The prediction theory of multivariate stochastic processes, III. *Acta Math.* **104** 141–162.

MASANI, P. (1966). Recent trends in multivariate prediction theory. In *Multivariate Analysis* (P. R. Krishnaiah, ed.) 351–382. Academic, New York.

MASANI, P. (1968). Orthogonally scattered measures. *Adv. in Math.* **2** 61–117.

MIAMEE, A. G. and SALEHI, H. (1989). An example of a harmonizable process whose spectral domain is not complete. RM–502. *Scand. J. Statist.* To appear.

NIEMI, H. (1975a). Stochastic processes as Fourier transforms of stochastic measures. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **591** 1–47.

NIEMI, H (1975b). On stationary dilations and the linear prediction of certain stochastic processes. *Soc. Sci. Fenn. Comment. Phys.-Math.* **45** 111–130.

NIEMI, H. (1977). On orthogonally scattered dilations of bounded vector measures. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **593** 43–52.

PIETSCH, A. (1969). *P*-majorisierbare vektorwertige Masse. *Wiss. Z. Friedrich-Schiller-Univ., Jena Math.-Naturwiss.* **18** 243–247.

RISSANEN, J. and BARBOSA, L. (1969). A factorization problem and the problem of predicting non-stationary vector-valued stochastic processes. *Z. Wahrsch. Verw. Gebiete* **12** 255–266.

ROBERTSON, J. B. and ROSENBERG, M. (1968). The decomposition of matrix-valued measures. *Michigan Math. J.* **15** 353–368.

ROSENBERG, M. (1964). The square integrability of matrix-valued functions with respect to a non-negative hermitian measure. *Duke Math. J.* **31** 291–298.

ROSENBERG, M. (1982). Quasi-isometric dilations of operator-valued measures and Grothendieck's inequality. *Pacific J. Math.* **103** 135–161.

ROZANOV, YU. A. (1958). The spectral theory of multidimensional, stationary random processes with discrete time. *Uspekhi Mat. Nauk.* **13** 93–142. (In Russian.)

ROZANOV, YU. A. (1967). *Stationary Random Processes*. Holden–Day, San Francisco.

SAEKI, S. (1976). Tensor products of $C(X)$-spaces and their conjugate spaces. *J. Math. Soc. Japan* **28** 33–47.

TRUONG-VAN, B. (1981). Une généralization du théorème de Kolmogorov–Aronszajn, processus *V*-bornés *q*-dimensionnels: Domaine spectral, dilatations stationnaires. *Ann. Inst. H. Poincaré. Probab. Statist.* **17** 31–49.

WIENER, N. (1942). *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. Technology Press, New York. [Reissued (1949) by Wiley, New York. Originally issued as a Nat. Defense Research Council Report.]

WIENER, N. (1981). *Collected Works* (P. Masani, ed.) **3**. MIT Press, Cambridge, Mass.

WIENER, N. and MASANI, P. (1957). The prediction theory of multivariate stochastic processes, I. *Acta Math.* **98** 111–150.

WIENER, N. and MASANI, P. (1958). The prediction theory of multivariate stochastic processes, II. *Acta Math.* **99** 93–137.

YLINEN, K. (1978). On vector bimeasures. *Ann. Mat. Pura Appl. (4)* **117** 115–138.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742