Stochastic LTI Processes and Their Invariance Properties: a Behavioral Perspective ⋆

Giacomo Baggio * **Rodolphe Sepulchre** **

* *Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B 35131 Padova, Italy. (e-mail: giacomo.baggio@studenti.unipd.it).*
** *Department of Engineering, University of Cambridge Trumpington Street, Cambridge CB2 1PZ, UK (e-mail:).*

Abstract: In this paper, we revisit the definition of linear time-invariant (LTI) stochastic process within a behavioral systems framework. Building on Willems (2013), we derive a canonical representation of a LTI stochastic process and a physically grounded notion of interconnection between independent stochastic processes. Furthermore, we analyze the invariance properties of LTI processes and we translate them into equivalent properties in the space of rational discrete-time spectral densities.

Keywords: Stochastic modelling, Stochastic systems, Behavioral theory, Spectral densities.

1. INTRODUCTION

Stochastic models play a crucial role in elucidating many areas of the natural and engineering sciences. Indeed, mathematical models of stochastic phenomena are widely used in many branches of physics (Van Kampen, 2007), engineering (Åström, 1970; Lindquist and Picci, 2015), economics (Dupačová et al., 2002), and biology (Wilkinson, 2011). Their importance relies on the fact that many—if not any—physical phenomena are unavoidably corrupted by "noise". This inherent source of stochasticity sometimes can not be neglected in the mathematical modelling procedure, in the sense that it represents an essential part of the model. In this paper, we address the problem of modelling stochastic phenomena from an open systems viewpoint. A system is *open* if it can interact with its external environment by means of some "unmodelled" features. As opposed to closed systems, open systems are amenable to *interconnection.* The term interconnection here is intended in the most general sense, i.e. as *variable sharing* between systems (Willems, 2007). In this setting, a theory of open systems is de facto a theory of interconnected systems. When dealing with deterministic linear time-invariant (LTI, for short) systems, an elegant and mature theory of open systems is provided by the theory of behaviors (Willems and Polderman, 1997). Recently, a generalization of the latter theory to a stochastic framework has been proposed in Willems (2013). In that paper the emphasis is put on *static* stochastic systems.

Building on Willems (2013), in this paper, we present an extension of the notion of open stochastic system to the dynamical case. We then specialize it to the case of LTI discrete-time processes. In particular, we analyze: (i) a representation of such processes in terms of a linear model, (ii) in what sense and under which conditions these processes can be interconnected, and (iii) the invariance properties inherited by these processes. With reference to the latter point, a new invariance property in the space of rational multivariate discrete-time spectral densities is discussed.

Paper structure. The rest of the paper is organized as follows. After reviewing in §2 some results of behavioral theory, in §3 we introduce the definition of stochastic process in a behavioral flavor. In §4 we focus on the linear-time invariant case. §5 is devoted to present and analyze interconnection between LTI processes. In §6 we study the invariance properties of LTI processes. Finally, §7 collects some concluding remarks and future research directions.

Notation. Throughout the paper, we denote by $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ the set of integer, real, and complex numbers, respectively. Given a set $\mathbb{A}$, A^c will denote the complement of A w.r.t. $\mathbb{A}$, while $\emptyset$ the empty set. We let $\mathbb{R}^{n\times m}$ denote the set of real-valued $n\times m$ matrices and $A^\top \in \mathbb{R}^{m\times n}$ denote the transpose of $A \in \mathbb{R}^{n\times m}$. The symbols $\mathbb{R}[z,z^{-1}]^{n\times m}$ and $\mathbb{R}(z)^{n\times m}$ stand for the set of $n\times m$ Laurent polynomial matrices and the set of $n\times m$ rational matrices with real coefficients in the indeterminate z, respectively. The normal rank of a rational matrix $A(z)\in\mathbb{R}(z)^{n\times m}$ is defined as the rank of $A(z)$ almost everywhere in $z\in\mathbb{C}$ and will be denoted by $\mathrm{rk}(A)$. Given $A(z)\in\mathbb{R}(z)^{n\times m}$, we let $A^*(z) := A^\top(1/z)$ and, for square $A(z)$'s of full normal rank, $A^{-*}(z) := [A^\top(1/z)]^{-1}$. We recall that a Laurent unimodular matrix is a square Laurent polynomial matrix whose inverse is also Laurent polynomial or, equivalently, whose determinant is a non-zero monomial λz^k, $\lambda\in\mathbb{R}\setminus\{0\}$, $k\in\mathbb{Z}$. We denote by $\mathbb{U}[z,z^{-1}]^{n\times n}$ the group of n-dimensional Laurent unimodular matrices

⋆ Sponsor and financial support acknowledgment goes here.

and by $\mathcal{S}_+(\mathbf{T})^{n\times n}$ the set of $n\times n$ matrix-valued functions which are positive definite on the unit circle $\mathbf{T} := \{z \in \mathbb{C} : |z| = 1\}$, i.e. $n\times n$ discrete-time coercive spectral densities. $\mathcal{S}_{\mathrm{rat}}^{n\times n} \subset \mathcal{S}_+(\mathbf{T})^{n\times n}$ will denote the set of $n \times n$ *rational* discrete-time coercive spectral densities. Finally, we shall suppose the reader to be acquainted with some elementary notions of probability theory, e.g. the definitions of σ-algebra, (smallest) σ-algebra generated by a collection of sets, Borel σ-algebra, probability measure; notions that can be found in any standard textbook of probability or measure theory, e.g. Billingsley (1986).

2. BACKGROUND ON BEHAVIORAL THEORY

In this preliminary section, we quickly review some basic notions and results of behavioral theory. We refer the reader to the seminal papers Willems (1986, 1989, 1991) and to the monograph Willems and Polderman (1997) for a comprehensive treatment on the subject.

In the theory of behaviors, a dynamical system is defined as a triple $\Sigma := (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, where $\mathbb{T}$ is the set of times over which the system evolves (*time axis*), $\mathbb{W}$ is the set over which the variables of the signals being modelled take values (*signal space*), and $\mathfrak{B}$ is a subset of $\mathbb{W}^{\mathbb{T}}$ (i.e. the set of all maps from $\mathbb{T}$ to $\mathbb{W}$, also called *universum*) in which all the admissible system trajectories live (the *behavior* of the system). The dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is *linear* if $\mathbb{W}$ is a vector space and $\mathfrak{B}$ a linear subspace of $\mathbb{W}^{\mathbb{T}}$. Σ is said to be *time-invariant* if $\mathbb{T}$ is closed under addition and $\sigma^t\mathfrak{B} \subseteq \mathfrak{B}$ for all $t \in \mathbb{T}$, where σ denotes the backward shift operator defined as $(\sigma f)(t') := f(t'+1)$. In this paper we mainly focus on n-dimensional, real-valued, *discrete-time* systems. Hence we set $\mathbb{T} := \mathbb{Z}$ and $\mathbb{W} := \mathbb{R}^n$. As a consequence, the behavior of the system becomes a subset of $(\mathbb{R}^n)^{\mathbb{Z}}$ (the set of maps from $\mathbb{Z}$ to $\mathbb{R}^n$), i.e. a family of n-dimensional, real-valued, discrete-time sequences. A linear time-invariant system may be described by an Auto-Regressive (AR) model [1]

$$R_\ell w(t+\ell) + R_{\ell+1} w(t+\ell+1) + \cdots + R_L w(t+L) = 0,$$

for all $t \in \mathbb{Z}$, where $L, \ell \in \mathbb{Z}$, $L > \ell$. The Laurent polynomial matrix $R(z) := R_\ell z^\ell + R_{\ell+1} z^{\ell+1} + \cdots + R_L z^L \in \mathbb{R}[z, z^{-1}]^{n\times p}$ defines an operator in the shift σ, $R\colon (\mathbb{R}^p)^{\mathbb{Z}} \to (\mathbb{R}^n)^{\mathbb{Z}}$ which allows to rewrite the previous expression as

$$R(\sigma)w(t) = 0. \tag{1}$$

The behavior of the LTI system is then given by

$$\mathrm{Ker}_\infty R := \{w \in (\mathbb{R}^p)^{\mathbb{Z}} \,:\, R(\sigma)w = 0\}. \tag{2}$$

Equation (1) is known as the *kernel representation* of a LTI behavior. A LTI behavior $\mathfrak{B} \subseteq (\mathbb{R}^n)^{\mathbb{Z}}$ is said to be *complete* if for all $w(t) \in \mathfrak{B}$, $w|_{[t_1,t_2]} \in \mathfrak{B}|_{[t_1,t_2]}\ \forall\, t_1, t_2 \in \mathbb{Z}$, $t_1 \le t_2$, where $w|_{[t_1,t_2]}$ and $\mathfrak{B}|_{[t_1,t_2]}$ denote the restriction of $w(t)$ and $\mathfrak{B}$, respectively, to the time interval $[t_1, t_2]$. A fundamental result in behavioral theory states that every LTI complete behavior admits a kernel representation (1).

To conclude, consider two LTI complete behaviors $\mathfrak{B}_1$, $\mathfrak{B}_2$ with kernel representations described by $R_1(z)$, $R_2(z) \in \mathbb{R}[z, z^{-1}]^{n\times p}$, respectively. We say that the two behaviors $\mathfrak{B}_1$, $\mathfrak{B}_2$ are *equivalent* if $R_1(z) = U(z)R_2(z)$ with $U(z) \in \mathbb{U}[z, z^{-1}]^{n\times n}$ being a Laurent unimodular matrix. Hence, with reference to the kernel representation, every behavior is uniquely determined by its kernel matrix up to a unimodular transformation acting on the left.

[1] We remark that the behavior is always *deterministic*, i.e. it is composed by deterministic trajectories. Therefore, in this section, the term AR has not to be intended in relation to a stochastic framework.

3. FROM DETERMINISTIC TO STOCHASTIC BEHAVIORS

We begin by recalling that a *probability space*—or *stochastic system*, using the terminology of Willems (2013)—is defined as a triple $(\mathbb{V}, \mathscr{E}, P)$, where $\mathbb{V}$ is the outcome space, $\mathscr{E}$ is a σ-algebra of events, and $P\colon \mathscr{E} \to [0,1]$ is a probability measure which assigns to each event in $\mathscr{E}$ a value in the interval $[0,1]$.

Consider now a deterministic system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$. Σ can be regarded as a very special probability space. As a matter of fact, Σ coincides with the probability space $(\mathbb{W}^{\mathbb{T}}, \{\emptyset, \mathbb{W}^{\mathbb{T}}, \mathfrak{B}, \mathfrak{B}^c\}, P_\Sigma)$ in which P_Σ is chosen such that $P_\Sigma(\mathfrak{B}) = 1$. Since the previous definition unavoidably involves the probability measure P_Σ (and specifically the constraint $P_\Sigma(\mathfrak{B}) = 1$), we could clearly have used a σ-algebra of events richer than $\{\emptyset, \mathbb{W}^{\mathbb{T}}, \mathfrak{B}, \mathfrak{B}^c\}$. However the latter seems to be a more natural choice due to the fact that it is the most "parsimonious" σ-algebra, that is, the smallest possible σ-algebra of events containing the deterministic behavior $\mathfrak{B}$.

Now assume that some source of stochasticity is added to the deterministic system Σ (for instance, some additive noise acting on the trajectories of $\mathfrak{B}$), then we expect that the newly generated (stochastic) system will possess a richer σ-algebra of events—indeed the noise modifies, and, more precisely, enlarges the space of admissible trajectories of the system—and, as a consequence, a new probability measure. Furthermore, by adding more and more sets to our event space, we are, in a sense, moving more and more away from the class of deterministic systems. Loosely speaking, the cardinality of the σ-algebra of events can be considered as a measure of the "degree" of stochasticity of the system.

From this intuitive description, we can see that the σ-algebra of events plays an important role in the mathematical model of a stochastic system, perhaps as important as the probability measure associated to the system. When dealing with static systems, i.e. systems which do not evolve in time, this is exactly the point raised in Willems (2013). One of the aims of the present paper is to show that this is also true in the dynamical case. To this end, we first revisit the definition of a stochastic process in the spirit of behavioral theory.

Definition 1. A stochastic process is a quadruple $\Sigma := (\mathbb{T}, \mathbb{W}, \mathscr{E}, P)$, where

(1) $\mathbb{T}$ is the time axis,
(2) $\mathbb{W}$ is the signal space, i.e. the set in which the variables whose (noisy) time evolution is modelled take on their values,
(3) $\mathscr{E}$ is a σ-algebra of subsets of $\mathbb{W}^{\mathbb{T}}$ with elements called events,
(4) $P\colon \mathscr{E} \to [0,1]$ is the probability measure defined on the σ-algebra of events.

Remark 2. With reference to the above definition, we observe that:

(i) A stochastic process is a probability space where the outcome space is given by $\mathbb{W}^{\mathbb{T}}$. Two important classes of stochastic processes are obtained by selecting $\mathbb{T} = \mathbb{R}$, in which case the outcome space is the space of functions $f\colon \mathbb{R} \to \mathbb{W}$, and $\mathbb{T} = \mathbb{Z}$, in which case the outcome space is the space of sequences $\{f_t\}_{t\in\mathbb{Z}}$ taking values on $\mathbb{W}$. Intuitively, we can think of a stochastic process as a system described by a collection of "behaviors" consisting of sets contained in the σ-algebra $\mathscr{E}$, where to each "behavior" is assigned, through P, a probability of being selected.

(ii) The standard definition of stochastic process is a family of random variables (i.e. measurable functions) $\{f_t\}_{t\in\mathbb{T}}$ defined on some probability space and parametrized by an index $t \in \mathbb{T}$, which usually represents time. By specifying the finite-dimensional probability distributions of the family $\{f_t\}_{t\in\mathbb{T}}$ it is then possible to characterize the infinite-dimensional distributions of the process (by virtue of Kolmogorov existence theorem (Billingsley, 1986, Thm. 36.1)). Our definition of a stochastic process is essentially equivalent to the latter one but formulated in terms of σ-algebras of events defined on the (usually infinite-dimensional) space of trajectories $\mathbb{W}^{\mathbb{T}}$. From this point of view, in Definition 1 emphasis is put on the event space itself rather than on the variables that generate that space.

$\diamondsuit$

For the rest of the paper, we restrict our attention to the class of n-dimensional real-valued discrete-time stochastic processes, i.e. we set $\mathbb{T} = \mathbb{Z}$ and $\mathbb{W} = \mathbb{R}^n$. Notice that, in this setting, we can identify two particular subclasses of stochastic processes, namely the subclass of deterministic dynamical systems whose σ-algebra of events is given by $\{\emptyset, (\mathbb{R}^n)^{\mathbb{Z}}, \mathfrak{B}, \mathfrak{B}^c\}$ with $\mathfrak{B} \subset (\mathbb{R}^n)^{\mathbb{Z}}$, and the subclass of stochastic processes whose σ-algebra of events is given by the Borel σ-algebra generated by the open sets of $(\mathbb{R}^n)^{\mathbb{Z}}$ equipped with the product topology (i.e. the topology of pointwise convergence), which we will denote by $\mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}})$. Since the Borel σ-algebra of $(\mathbb{R}^n)^{\mathbb{Z}}$ coincides with the σ-algebra containing all the "non-pathological" subsets of $(\mathbb{R}^n)^{\mathbb{Z}}$, we can think of these two subclasses as two extremes in the space of stochastic processes.

Remark 3. It is worth noting that when dealing with continuous-time stochastic processes ($\mathbb{T} = \mathbb{R}$) the Borel σ-algebra generated by the open sets of the product space $(\mathbb{R}^n)^{\mathbb{R}}$ equipped with the product topology proves often to be inadequate for describing the events of the process (Billingsley, 1986, Ch. 7). Indeed, for instance, it can be shown that the many "interesting" sets of functions, e.g. the set of continuous functions, are not contained in $\mathscr{B}((\mathbb{R}^n)^{\mathbb{R}})$. To overcome this issue, other types of σ-algebras can be considered in place of $\mathscr{B}((\mathbb{R}^n)^{\mathbb{R}})$, for instance the Borel σ-algebra generated by the open sets in the space of continuous functions equipped the topology of uniform convergence on compact sub-intervals. $\diamondsuit$

4. LINEAR TIME-INVARIANT STOCHASTIC PROCESSES

In this section we introduce the notion of linear time-invariant (discrete-time) stochastic process. For these processes the "coarseness" of the σ-algebra of events runs along the subspace of trajectories defined by a deterministic LTI behavior. We then discuss a canonical representation for these systems.

Definition 4. The stochastic process $\Sigma := (\mathbb{Z}, \mathbb{R}^n, \mathscr{E}, P)$ is said to be *linear* and *time-invariant* (LTI, for short) if there exists a linear and time-invariant behavior $\mathfrak{L} \subset (\mathbb{R}^n)^{\mathbb{Z}}$ such that the events are the Borel subsets of the quotient space $(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}$, i.e. $\mathscr{E} := \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L})$, and P is a Borel probability measure on the same quotient space, i.e. $P\colon \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}) \to [0, 1]$.

Observe that $\mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L})$ is a well-defined Borel σ-algebra. Indeed, it coincides exactly with the Borel σ-algebra generated by the open sets $A/\mathfrak{L}$ (open sets of the quotient topology), with A an open set of the topological vector space $(\mathbb{R}^n)^{\mathbb{Z}}$ equipped with the product topology. Moreover, using the terminology introduced in Willems (2013), we call $\mathfrak{L}$ the *fiber* of the LTI stochastic process.

Definition 4 can be intuitively interpreted as follows: given any (Borel) subset $\bar{E} \subset (\mathbb{R}^n)^{\mathbb{Z}}$, which consists of a subset of trajectories in $(\mathbb{R}^n)^{\mathbb{Z}}$, if the stochastic process Σ is LTI with fiber $\mathfrak{L}$ then the subset E generated by adding to $\bar{E}$ the trajectories belonging to the LTI behavior $\mathfrak{L}$ is an event of Σ. Loosely speaking, an event is a collection of subsets in $(\mathbb{R}^n)^{\mathbb{Z}}$ with trajectories "parallel" to the LTI behavior $\mathfrak{L}$ (see Fig. 1 for a pictorial example).

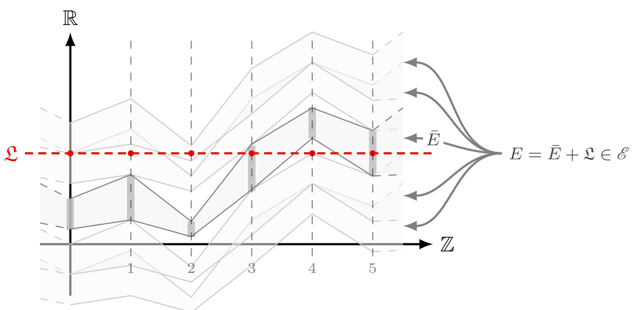

Fig. 1. In a LTI stochastic process an event $E \in \mathscr{E}$ corresponds to a fixed event $\bar{E}$ plus all the subsets "shifted" by the fiber $\mathfrak{L}$ (in this figure $\mathfrak{L}$ is made of constant trajectories, i.e. trajectories belonging to the set $\{w \in \mathbb{R}^{\mathbb{Z}} : w(t+1) = w(t),\ \forall t \in \mathbb{Z}\}$).

Throughout the paper, we will often make the assumption that the fibers are *complete* LTI behaviors, this means that all the trajectories belonging to the fiber admit a finite dimensional characterization, that is, they are uniquely determined by their restrictions over all possible finite time intervals. Under this assumption, the fiber of a LTI stochastic process can always be represented by means of the kernel of a Laurent polynomial matrix, as recalled in §2. LTI processes characterized by a complete fiber admit a canonical representation, called *kernel representation* by analogy with the deterministic case.

Theorem 5. A stochastic process $\Sigma = (\mathbb{Z}, \mathbb{R}^n, \mathscr{E}, P)$ is a LTI stochastic process with fiber $\mathfrak{L}$ being a complete LTI behavior if and only if Σ can be described by a stochastic sequence $w(\cdot)$ satisfying for all $t \in \mathbb{Z}$

$$R(\sigma)w(t) = e(t), \tag{3}$$

where $R(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$ is of full row normal rank, i.e. $\mathrm{rk}(R) = m$, and $e(\cdot)$ describes the stochastic process $\Sigma_e := (\mathbb{Z}, \mathbb{R}^m, \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}}), P_e)$.

Proof. See Appendix A.

With reference to the previous result, we remark that if $E \in \mathscr{E}$ is an event of Σ then its probability measure P is defined through $e(\cdot)$ by

$$P(E) := P_e(R[E]),$$

being $R[E] \in \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}})$ the image of E under $R(\sigma)$.

Furthermore, if we consider the restriction of a LTI process to a finite set of time indices, say $I := \{t_1, t_2, \ldots, t_n\}$, $t_i \in \mathbb{Z}$, $n > 0$, then we obtain a (static) stochastic system described by the triple $\Sigma|_I := ((\mathbb{R}^n)^{|I|}, \mathscr{E}|_I, P|_I)$, where $|I|$ is the cardinality of the set I, $\mathscr{E}|_I := \mathscr{B}((\mathbb{R}^n)^{|I|}/\mathfrak{L}|_I)$ with $\mathfrak{L}|_I$ the restriction of the complete LTI behavior to the time set I, and $P|_I$ a restricted probability measure defined for all $E \in \mathscr{E}|_I$ as

$$P|_I(E) := P\left(\bigcup_{G_i\in\pi^{-1}[E]} G_i\right),$$

being $\pi^{-1}[E]$ the pre-image of E under the canonical projection $\pi\colon (\mathbb{R}^n)^{\mathbb{Z}} \to (\mathbb{R}^n)^{|I|}$, $\{f_t\}_{t\in\mathbb{Z}} \mapsto \{f_{t_1}, f_{t_2}, \ldots, f_{t_n}\}$. Since in general the restriction $\mathfrak{L}|_I$ returns a non-empty linear finite-dimensional subspace of $(\mathbb{R}^n)^{|I|}$, we note that $\Sigma|_I$ does not, in general, describe a *classical* random vector, where for classical we mean a random vector characterized by a Borel σ-algebra of events on $(\mathbb{R}^n)^{|I|}$, as in (Willems, 2013, Def. 2).

5. INTERCONNECTION OF STOCHASTIC PROCESSES

Interconnection is a property characterizing open systems, i.e. systems which are allowed to interact with their environment. With reference to the mathematical model of a deterministic dynamical system, this interaction can take place if some variables of the system are left unmodelled (Willems, 2007). In this section we present an extension of the definition of interconnection between deterministic dynamical systems which applies to stochastic processes. After introducing some general definitions, we will focus on the discrete-time LTI case.

As in the deterministic case, interconnection of two stochastic processes can be thought of as *variable sharing* between the two processes. In other words, interconnection between two processes is obtained by simply imposing an equality constraint on the variables describing the stochastic laws of the two processes (Fig. 2).

In the deterministic case, given two dynamical systems $\Sigma_1 = (\mathbb{W}, \mathbb{T}, \mathfrak{B}_1)$ and $\Sigma_2 = (\mathbb{W}, \mathbb{T}, \mathfrak{B}_2)$ having the same time axis and signal space, the interconnection between Σ_1 and Σ_2 is defined as the deterministic system $\Sigma_1 \wedge \Sigma_2 := (\mathbb{W}, \mathbb{T}, \mathfrak{B}_1 \cap \mathfrak{B}_2)$ (Willems, 2007). In the stochastic case, the definition of interconnection we are going to present is similar to the latter one if we replace the role of the deterministic behaviors with the σ-algebras of events of the processes (which indeed represent collections of admissible "behaviors" of the processes). However, in this case, a problem arises. As a matter of fact, since interconnection of stochastic processes also involves the probability laws defined on the processes, a natural compatibility condition between the two to-be-interconnected processes has to be fulfilled. This natural condition states that the probability measure defined on the interconnected process must be consistent, in a sense explained below, with the probability measures defined on the original processes. This condition was introduced in Willems (2013) with reference to (static) stochastic systems under the name of *complementarity*. In the following definition, we adapt the notion of complementarity to the case of stochastic processes.

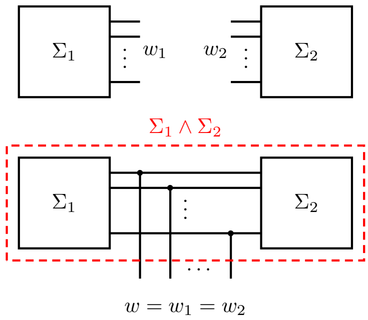

Fig. 2. Interconnection of stochastic processes $\Sigma_1 = (\mathbb{T}, \mathbb{W}, \mathscr{E}_1, P_1)$ and $\Sigma_2 = (\mathbb{T}, \mathbb{W}, \mathscr{E}_2, P_2)$.

Definition 6. Two stochastic processes $\Sigma_1 = (\mathbb{T}, \mathbb{W}, \mathscr{E}_1, P_1)$ and $\Sigma_2 = (\mathbb{T}, \mathbb{W}, \mathscr{E}_2, P_2)$ are said to be *complementary* if for all E_1, $E_1' \in \mathscr{E}_1$ and E_2, $E_2' \in \mathscr{E}_2$ such that $E_1 \cap E_2 = E_1' \cap E_2'$ it holds

$$P_1(E_1)P_2(E_2) = P_1(E_1')P_2(E_2').$$

Moreover, the two σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ are said to be *complementary* if for all non-empty sets E_1, $E_1' \in \mathscr{E}_1$ and E_2, $E_2' \in \mathscr{E}_2$ such that $E_1 \cap E_2 = E_1' \cap E_2'$ it holds

$$E_1 = E_1' \text{ and } E_2 = E_2'.$$

Remark 7. The notion of complementarity between σ-algebra of the processes is weaker than the notion of complementarity of processes. Indeed, the former represents only a sufficient condition for complementarity of processes. However, working with complementarity of σ-algebra is usually easier since this notion does not involve the probability laws describing the processes, as pointed out also in Willems (2013). $\diamondsuit$

Under the assumption of complementarity between two stochastic processes, we arrive at a formal definition of interconnection.

Definition 8. Let $\Sigma_1 = (\mathbb{T}, \mathbb{W}, \mathscr{E}_1, P_1)$ and $\Sigma_2 = (\mathbb{T}, \mathbb{W}, \mathscr{E}_2, P_2)$ be two independent [2] and complementary stochastic processes. The *interconnection* of Σ_1 and Σ_2 is defined as the stochastic process

$$\Sigma_1 \wedge \Sigma_2 := (\mathbb{T}, \mathbb{W}, \mathscr{E}, P),$$

where $\mathscr{E}$ is the σ-algebra generated by $\mathscr{E}_1 \cup \mathscr{E}_2$ and P is defined on the sets $\{E_1 \cap E_2 : E_1 \in \mathscr{E}_1, E_2 \in \mathscr{E}_2\}$ by

$$P(E_1 \cap E_2) := P_1(E_1)P_2(E_1)$$

[2] We say that two stochastic processes are (stochastically) independent if their σ-algebras of events are so, with respect to any joint probability measure.

and extended to all of $\mathscr{E}$ by virtue of the Hahn-Kolmogorov extension theorem (also known as Carathéodory extension theorem or simply extension theorem, see e.g. (Billingsley, 1986, Ch. 3)). [3]

Remark 9. It is worth pointing out that interconnection between stochastic processes, as given in Definition 8, differs from the classical notion of *coupling* between stochastic processes. As a matter of fact, even in the static case, coupling of two stochastic systems $\Sigma_1 = (\mathbb{W}, \mathscr{E}_1, P_1)$ and $\Sigma_2 = (\mathbb{W}, \mathscr{E}_2, P_2)$ requires the construction of a new stochastic system with signal space $\mathbb{W} \times \mathbb{W}$, σ-algebra generated by the sets $\{E_1 \times E_2 \,:\, E_1 \in \mathscr{E}_1, E_2 \in \mathscr{E}_2\}$, and probability measure having prescribed marginal distributions, see e.g. Lindvall (2002). On the other hand, interconnection between Σ_1 and Σ_2 means that a new σ-algebra is constructed on the *same* signal space shared by the two to-be-interconnected systems, that is $\mathbb{W}$. More precisely, the events which lie in the intersection between the two σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ generate the σ-algebra of the interconnected system. From this viewpoint, coupling appears more similar to *juxtaposition* of stochastic processes than interconnection, where for juxtaposition we mean that starting from two processes described by stochastic laws w_1 and w_2 we construct a new process described by (w_1, w_2), as in Willems (2013). $\diamondsuit$

We now restrict the attention to the class of LTI stochastic processes. In this case, the fiber of the process is given by a LTI behavior. If we add the further assumption that the fiber is described by a complete LTI behavior, then, by virtue of Theorem 5, the process admits a kernel representation. For this class of stochastic processes, it is possible to derive a condition on the kernel matrices which is equivalent to complementary of σ-algebras of events of the processes.

Theorem 10. Consider two independent LTI complete stochastic processes $\Sigma_1 := (\mathbb{Z}, \mathbb{R}^n, \mathscr{E}_1, P_1)$ and $\Sigma_2 := (\mathbb{Z}, \mathbb{R}^n, \mathscr{E}_2, P_2)$ described by fibers $\mathfrak{L}_1 := \mathrm{Ker}_\infty R_1$ and $\mathfrak{L}_2 := \mathrm{Ker}_\infty R_2$, for suitable Laurent polynomial matrices $R_1(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$, $R_2(z) \in \mathbb{R}[z, z^{-1}]^{p\times n}$ with $\mathrm{rk}(R_1) = m$ and $\mathrm{rk}(R_2) = p$. The two σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ are complementary if and only if it holds

$$\mathrm{rk}\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = m + p. \tag{4}$$

In this case, the fiber of the interconnected process $\Sigma_1 \wedge \Sigma_2$ is given by

$$\mathfrak{L}_{1\wedge 2} := \mathfrak{L}_1 \cap \mathfrak{L}_2 = \mathrm{Ker}_\infty \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}. \tag{5}$$

Proof. See Appendix A.

Example 11. As a simple example of interconnection, consider two LTI processes $\Sigma_1 := (\mathbb{Z}, \mathbb{R}^2, \mathscr{E}_1, P_1)$ and $\Sigma_2 := (\mathbb{Z}, \mathbb{R}^2, \mathscr{E}_2, P_2)$ described by kernel representations

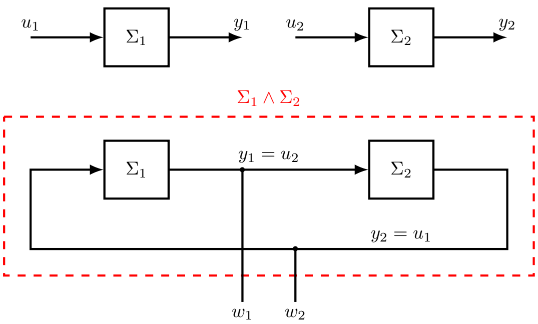

Fig. 3. Interconnection of LTI stochastic processes Σ_1 and Σ_2 of Example 11 in an input-output representation.

$$[\sigma + a_1 \;\; \sigma + b_1]\, w_1(t) = e_1(t), \quad a_1, b_1 \in \mathbb{R},$$
$$[\sigma + b_2 \;\; \sigma + a_2]\, w_2(t) = e_2(t), \quad a_2, b_2 \in \mathbb{R},$$

respectively. Furthermore, assume that $e_1(\cdot)$ and $e_2(\cdot)$ describe stochastically independent processes $\Sigma_{e_1} = (\mathbb{Z}, \mathbb{R}, \mathscr{B}(\mathbb{R}^{\mathbb{Z}}), P_{e_1})$ and $\Sigma_{e_2} = (\mathbb{Z}, \mathbb{R}, \mathscr{B}(\mathbb{R}^{\mathbb{Z}}), P_{e_2})$, respectively. If we partition the variables w_1 and w_2 in a "input-output" form $w_1 := [u_1 \; y_1]^\top$ and $w_2 := [y_2 \; u_2]^\top$, then the two LTI processes Σ_1 and Σ_2 can be regarded as two noisy input/output LTI systems (see also Fig. 3). The two σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ are complementary if and only if

$$\mathrm{rk}\begin{bmatrix} z + a_1 & z + b_1 \\ z + b_2 & z + a_2 \end{bmatrix} = 2,$$

or, equivalently, if and only if $a_1 + a_2 \neq b_1 + b_2$ and $a_1 a_2 \neq b_1 b_2$. If the latter conditions are met, the interconnected process $\Sigma_1 \wedge \Sigma_2$ is a well-defined LTI process described by the laws of the stochastic sequence $w(t) := [w_1(t) \; w_2(t)]^\top$ satisfying

$$\begin{bmatrix} \sigma + a_1 & \sigma + b_1 \\ \sigma + b_2 & \sigma + a_2 \end{bmatrix} w(t) = \begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix}.$$

$\diamondsuit$

To conclude this section, we present a straightforward corollary of Theorem 10 which gives a characterization of interconnected processes with a "full" σ-algebra of events.

Corollary 12. Consider the two Laurent polynomial matrices $R_1(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$ and $R_2(z) \in \mathbb{R}[z, z^{-1}]^{p\times n}$ describing the LTI processes defined in Theorem 10. The σ-algebra of the interconnected system $\Sigma_1 \wedge \Sigma_2$ is given by $\mathscr{E} = \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}})$ if and only if $R := [R_1^\top \; R_2^\top]^\top \in \mathbb{U}[z, z^{-1}]^{n\times n}$, i.e. R is a unimodular Laurent polynomial matrix.

6. INVARIANCE PROPERTIES OF LTI STOCHASTIC PROCESSES

[Giacomo: This section has to be revised!]

In the previous sections, we have pointed out that in the definition of stochastic process a crucial role is played by the event space $\mathscr{E}$. For LTI stochastic processes, the structure of the event space is characterized by its *fiber*, i.e. by the subspace of trajectories described by a LTI behavior. Let us restrict the attention to fibers defined by complete behaviors and consider two stochastic processes in their kernel representations

$$\Sigma_1: \quad R_1(\sigma) w_1(t) = e_1(t), \tag{6}$$
$$\Sigma_2: \quad R_2(\sigma) w_2(t) = e_2(t). \tag{7}$$

[3] The Hahn-Kolmogorov theorem gives conditions under which a function $\mu\colon \mathscr{A} \to [0,1]$ defined on an algebra $\mathscr{A}$ of subsets of Ω can be extended to a unique bona fide probability measure on the σ-algebra generated by $\mathscr{A}$. These conditions are: (i) $\mu(\Omega) = 1$, (ii) countably additivity, i.e. $\mu\left(\bigcup_{i=1}^\infty A_i\right) = \sum_{i=1}^\infty \mu(A_i)$ for any countable disjoint family of subsets $\{A_i\}_{i=1}^\infty$, $A_i \in \mathscr{A}$, such that $\bigcup_{i=1}^\infty A_i \in \mathscr{A}$. (A function satisfying these two requirements is called a *pre-measure* on $\mathscr{A}$.) In our case, the theorem applies to P since the latter is defined through the product of two bona fide probability measures.

with $R_1(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$, $R_2(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$ with $\mathrm{rk}(R_1) = \mathrm{rk}(R_2) = m$ and $e_1 \sim (\mathbb{Z}, (\mathbb{R}^m)^{\mathbb{Z}}, \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}}), P_{e_1})$ and $e_2 \sim (\mathbb{Z}, (\mathbb{R}^m)^{\mathbb{Z}}, \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}}), P_{e_2})$. If the two processes happen to have the same fiber, then $\mathrm{Ker}_\infty R_1$ and $\mathrm{Ker}_\infty R_2$ are equivalent behaviors. This in turn implies that R_1 and R_2 are connected by a unimodular transformation acting on the left, i.e. $R_1(z) = U(z)R_2(z)$ with $U(z) \in \mathbb{U}[z, z^{-1}]^{m\times m}$.

Thus, similarly to the deterministic case, we say that two LTI stochastic processes described by complete fibers are *equivalent* if their fibers are equivalent behaviors, i.e. if their kernel polynomial matrices are identical up to unimodular transformations acting on the left.

We now investigate what this unimodular invariance property entails when applied to spectral densities of stochastic processes. To this extent, assume that e_1, e_2 are white noise processes, and w_1, w_2 stochastic processes in the classical sense, i.e. equipped with Borel σ-algebras of events. [4] For $n = m$, the spectral densities of w_1 and w_2 in (6)-(7) are given, respectively, by

$$\Phi_1(z) := R_1^{-1}(z)R_1^{-*}(z), \quad \Phi_2(z) := R_2^{-1}(z)R_2^{-*}(z).$$

The fact that Σ_1 and Σ_2 are equivalent translates into

$$\Phi_1(z) = R_2^{-1}(z)V(z)V^*(z)R_2^{-*}(z).$$

for a suitable unimodular matrix $V(z) \in \mathbb{U}[z, z^{-1}]^{n\times n}$. We can generalize this observation to arbitrary rational discrete-time coercive spectral densities as follows. Consider two rational spectral densities $\Phi_1, \Phi_2 \in \mathcal{S}_{\mathrm{rat}}^{n\times n}$ and let $W_2 \in \mathbb{R}(z)^{n\times n}$ be a given minimal spectral factor of Φ_2, say its minimum-phase spectral factor. We say that two spectral densities $\Phi_1, \Phi_2 \in \mathcal{S}_{\mathrm{rat}}^{n\times n}$ are *unimodular equivalent* if $\Phi_1 = W_2VV^*W_2^*$ with $V(z) \in \mathbb{U}[z, z^{-1}]^{n\times n}$.

First, we remark that unimodular equivalence is a property of spectral densities since it does not depend on the chosen minimal spectral factor. As a matter of fact, if Φ_1 and Φ_2 are unimodular equivalent w.r.t. the minimum-phase spectral factor W_2, then they are so w.r.t. any other spectral factor of Φ_2, as long as it is minimal.

Second, we observe that in the scalar case, unimodular Laurent transformations take the form $u(z) := \alpha z^k$, where $\alpha \in \mathbb{R} \setminus \{0\}$ and $k \in \mathbb{Z}$. In this case, it easy to see that unimodular invariance reduces to a scaling invariance property between spectral densities. Namely, let $\Phi_1, \Phi_2 \in \mathcal{S}_{\mathrm{rat}}^{1\times 1}$, Φ_1 and Φ_2 are unimodular equivalent if and only if

$$\Phi_1 = \alpha\Phi_2, \quad \alpha \in \mathbb{R} \setminus \{0\}.$$

We believe that unimodular invariance property between rational spectra can be exploited in order to define a "natural" projective metric in the space $\mathcal{S}_{\mathrm{rat}}^{n\times n}$. In particular, a metric that satisfies unimodular invariance property in the scalar case has been proposed in Martin (2000). Martin's metric has received a lot of attention both from a theoretical and application-oriented viewpoint, see e.g. (De Cock and De Moor, 2002; Chaudhry and Vidal, 2009). However a multivariate extension of this metric is still missing. We think that unimodular invariance property of spectra might be enlightening in this regard. This topic will be the subject of future investigation.

[4] In fact, we can always go back to the "classical" case, by enriching the σ-algebra of events and suitably re-defining the probability over the new σ-algebra.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we addressed the problem of modelling stochastic dynamical systems from a behavioral perspective. We focused on LTI processes and we analyzed their interconnection and invariance properties. Building on this analysis, new invariance property, unimodular invariance, in the space of rational spectral densities has been derived. We believe that this property might prove to be useful for defining a metric in the latter space. This is one of the most compelling directions for future research.

REFERENCES

Åström, K.J. (1970). *Introduction to stochastic control theory*. Academic Press, New York and London.

Billingsley, P. (1986). *Probability and measure*. Wiley series in Probability and Statistics. John Wiley & Sons, 2nd edition.

Bourbaki, N. (2003). *Topological vector spaces: Chapters 1–5*. Springer-Verlag Berlin Heidelberg.

Chaudhry, R. and Vidal, R. (2009). Recognition of visual dynamical processes: Theory, kernels and experimental evaluation. Technical report, Johns Hopkins University, Baltimore, MD, USA.

De Cock, K. and De Moor, B. (2002). Subspace angles between ARMA models. *Systems & Control Letters*, 46(4), 265–270.

Dupačová, J., Štěpán, J., and Hurt, J. (2002). *Stochastic modeling in economics and finance*, volume 75 of *Applied Optimization*. Springer US.

Kechris, A. (2012). *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer New York.

Lindquist, A. and Picci, G. (2015). *Linear Stochastic Systems*, volume 1 of *Series in Contemporary Mathematics*. Springer-Verlag Berlin Heidelberg.

Lindvall, T. (2002). *Lectures on the coupling method*. Wiley series in probability and mathematical statistics. Wiley, New York.

Martin, R.J. (2000). A metric for ARMA processes. *Signal Processing, IEEE Transactions on*, 48(4), 1164–1170.

Van Kampen, N. (2007). *Stochastic processes in physics and chemistry*. Elsevier, 3rd edition.

Wilkinson, D.J. (2011). *Stochastic modelling for systems biology*. CRC press.

Willems, J.C. (1986). From time series to linear system – part I. finite dimensional linear time invariant systems. *Automatica*, 22(5), 561–580.

Willems, J.C. (1989). Models for dynamics. In *Dynamics reported*, 171–269. Springer.

Willems, J.C. (1991). Paradigms and puzzles in the theory of dynamical systems. *Automatic Control, IEEE Transactions on*, 36(3), 259–294.

Willems, J.C. (2007). The behavioral approach to open and interconnected systems. *Control Systems Magazine, IEEE*, 27(6), 46–99.

Willems, J.C. (2013). Open stochastic systems. *Automatic Control, IEEE Transactions on*, 58(2), 406–421.

Willems, J.C. and Polderman, J.W. (1997). *Introduction to mathematical systems theory: a behavioral approach*, volume 26 of *Texts in Applied Mathematics*. Springer-Verlag New York.

Appendix A. EXTENDED PROOFS

In this Appendix, we present the proofs of Theorem 5 and Theorem 10 of the main text.

Proof. [Theorem 5]. *"If"*: Assume that the stochastic process Σ is described by the stochastic law of the sequence $w(\cdot)$ satisfying (3). We first recall some facts concerning the topological vector space $(\mathbb{R}^n)^{\mathbb{Z}}$ and polynomial operators in the shift, which can be found in (Willems, 1989, §4). The space of time series $(\mathbb{R}^n)^{\mathbb{Z}}$ when equipped with the product topology is a completely metrizable and separable (i.e. Polish) topological vector space. Also, $\mathfrak{L} := \mathrm{Ker}_\infty R$ is a closed and linear subspace of $(\mathbb{R}^n)^{\mathbb{Z}}$. The polynomial operator in the shift $R(\sigma)$ is a linear, continuous, and surjective (since of full row normal rank) operator from $(\mathbb{R}^n)^{\mathbb{Z}}$ to $(\mathbb{R}^m)^{\mathbb{Z}}$. Consider now the quotient space $(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}$. Since $\mathfrak{L}$ is a closed and linear subspace of $(\mathbb{R}^n)^{\mathbb{Z}}$, this is again a Polish space, as it is separable (as every quotient space of a separable space) and completely metrizable w.r.t. the induced quotient topology (see e.g. (Bourbaki, 2003, Ch.1 §3.2)). Hence by taking the restriction of $R(\sigma)$ to the quotient space $(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}$, i.e. $R|_{(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}}$, we obtain a continuous and bijective operator between Polish spaces. From this fact it follows from (Kechris, 2012, Thm. 15.1) that $R|_{(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}}$ is a Borel isomorphism, i.e. both $R|_{(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}}$ and its inverse are Borel measurable. This implies that to each event set $E_e \in \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}})$ with associated probability $P_e(E_e)$ corresponds one and only one $E := R^{-1}[E_e] \in \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L})$ with probability $P(E) := P_e(E_e)$, where $R^{-1}[A]$ denotes the pre-image of A under $R(\sigma)$. Hence $w(\cdot)$ defines the LTI stochastic process $(\mathbb{Z}, \mathbb{R}^n, \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}), P)$, with $\mathfrak{L} = \mathrm{Ker}_\infty R$ being a complete LTI behavior.

"Only if": Assume now that $\mathfrak{L}$ is a complete LTI behavior and $\Sigma = (\mathbb{Z}, \mathbb{R}^n, \mathscr{E}, P)$ a LTI process with fiber $\mathfrak{L}$. Since $\mathfrak{L}$ is complete, there exists a Laurent polynomial matrix $R(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$, $\mathrm{rk}(R) = m \leq n$, such that $\mathfrak{L} = \mathrm{Ker}_\infty R$ (Willems, 1989, §4). As before, $R(\sigma)$ restricted to $(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}$ is a Borel isomorphism between $(\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}$, equipped with the quotient topology, and $(\mathbb{R}^m)^{\mathbb{Z}}$, equipped with the product topology. Consider

$$R(\sigma)w(t) = e(t),$$

where $e(\cdot)$ describes a stochastic process with signal space $\mathbb{R}^m$, σ-algebra $\mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}})$, and probability P_e such that $P_e(R[E]) := P(E)$ for all $E \in \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L})$, being $R[E] \in \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}})$ the image of E under R. From this construction it follows that the LTI stochastic process $((\mathbb{R}^n)^{\mathbb{Z}}, \mathscr{B}((\mathbb{R}^n)^{\mathbb{Z}}/\mathfrak{L}), P)$ is described by the stochastic law of $w(\cdot)$. This concludes the proof. □

Proof. [Theorem 10]. *"If"*: Assume that the normal rank condition in (4) holds. Firstly, observe that this implies that $m + p \leq n$, otherwise equality in (4) can not be attained. Secondly, as noticed in the proof of Theorem 5, the operator R_1, when restricted to the domain $(\mathbb{R}^n)^{\mathbb{Z}}/\mathrm{Ker}_\infty R_1$, is a Borel isomorphism between topological spaces $(\mathbb{R}^n)^{\mathbb{Z}}/\mathrm{Ker}_\infty R_1$ and $(\mathbb{R}^m)^{\mathbb{Z}}$. A similar result holds for R_2. This implies that R_1 (R_2, respectively) establishes a one-to-one correspondence between Borel sets in $(\mathbb{R}^m)^{\mathbb{Z}}$ ($(\mathbb{R}^p)^{\mathbb{Z}}$) and Borel sets in $(\mathbb{R}^n)^{\mathbb{Z}}/\mathrm{Ker}_\infty R_1$ ($(\mathbb{R}^n)^{\mathbb{Z}}/\mathrm{Ker}_\infty R_2$). Therefore every Borel set $\bar{E}_1 \in \mathscr{B}((\mathbb{R}^m)^{\mathbb{Z}})$ and $\bar{E}_2 \in \mathscr{B}((\mathbb{R}^p)^{\mathbb{Z}})$ uniquely determines events $E_1 := R_1^{-1}[\bar{E}_1] \in \mathscr{E}_1$ and $E_2 := R_2^{-1}[\bar{E}_2] \in \mathscr{E}_2$, respectively. Now, since (4) holds, we have that the polynomial operator in the shift

$$R(\sigma) := \begin{bmatrix} R_1(\sigma) \\ R_2(\sigma) \end{bmatrix}$$

is a linear, continuous, and *surjective* operator from $(\mathbb{R}^n)^{\mathbb{Z}}$ to $(\mathbb{R}^{m+p})^{\mathbb{Z}}$ (Willems, 1989, §4). Therefore, from

(i) surjectivity of R, and
(ii) the fact that $R_1|_{(\mathbb{R}^n)^{\mathbb{Z}}/\mathrm{Ker}_\infty R_1}$ and $R_2|_{(\mathbb{R}^n)^{\mathbb{Z}}/\mathrm{Ker}_\infty R_2}$ are Borel isomorphisms,

it follows that, for any non-empty event $E_1 \in \mathscr{E}_1$ and $E_2 \in \mathscr{E}_2$, the intersection $E_1 \cap E_2$ uniquely determines the set E_1 and E_2. (In particular, $E_1 \cap E_2$ is non-empty, if E_1 and/or E_2 are so.) This in turn implies that $\mathscr{E}_1$ and $\mathscr{E}_2$ are complementary σ-algebras.

"Only if": We prove the contrapositive, that is, if (4) does not hold then $\mathscr{E}_1$ and $\mathscr{E}_2$ are not complementary. Assume that (4) does not hold. Then the polynomial operator in the shift $R(\sigma) := [R_1(\sigma)^\top \ R_2(\sigma)^\top]^\top$ is not surjective, since the rows of $R(z)$ are linear dependent for every $z \in \mathbb{C}\backslash\{0\}$. Hence there exist two non-empty sets $E_1 \in \mathscr{E}_1$ and $E_2 \in \mathscr{E}_2$ whose intersection is the empty set. This in turn implies $E_1 \cap E_2^c = E_1 = E_1 \cap (\mathbb{R}^n)^{\mathbb{Z}}$. Therefore it follows that $\mathscr{E}_1$ and $\mathscr{E}_2$ are not complementary. □