Preprints of the 16th IFAC Symposium on System Identification
The International Federation of Automatic Control
Brussels, Belgium. July 11-13, 2012

Interconnection of Stochastic Systems

Jan C. Willems *

* *ESAT/SISTA, KU Leuven, B-3001 Leuven, Belgium*
Jan.Willems@esat.kuleuven.be, http://www.esat.kuleuven.be/∼jwillems

Abstract: We address the interconnection of stochastic systems. A stochastic system is defined as a probability triple. The specification of the set of events is an essential part of a stochastic model. Models often require a coarse event sigma-algebra. A stochastic system is linear if the events are cylinders with fibers parallel to a linear subspace of a vector space. Two stochastic systems can be interconnected if they are complementary. We discuss aspects of the identification problem from this vantage point.

Keywords: Interconnection, stochastic system, complementarity, linearity, system identification.

1. INTRODUCTION

One of the central aspects of systems thinking is the possibility of combining systems and viewing a complex system as an architecture of interconnected subsystems. This feature is important in all aspects of systems and control, in modeling, in system identification, in analysis, and in synthesis. In [1] and [2] we have discussed 'tearing, zooming, and linking' modeling procedures for deterministic systems. In the present section we deal with the composition of stochastic systems in an informal way. In Section 3 we formalize interconnection.

A convenient way to visualize systems is by block diagrams. Figure 1 shows a pictorial representation of a system as a black box with terminals. The variables w that are relevant in the

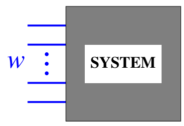

Fig. 1. A system as a black box

model are shown as associated with terminals. In some applications (as electrical circuits, and some mechanical, thermal, and hydraulic systems) these terminals can be taken literally, while for other applications they should be thought of as virtual terminals. For example, if $w \in \mathbb{R}^{\mathtt{n}}$, we may think of each of the terminals as corresponding to one of the components of $w = (w_1, w_2, \ldots, w_{\mathtt{n}})$. The black box indicates that the variables on the terminals are related, for example through the laws of a stochastic system.

One way of combining systems is by interconnection. We start with two systems with variables w_1 and w_2 respectively, and obtain a new system with variables w, as shown in Figure 2. The interconnection imposes variable sharing, $w_1 = w_2 = w$. Interconnection can also be viewed as an operation on the terminal variables of a single system. We start with a system with variables w_1 and w_2, and obtain a new system with variables w by setting $w = w_1 = w_2$, as shown in Figure 3. The basic idea of interconnection is variable sharing, in the sense explained in [1] and [2] for deterministic systems. Series, parallel, and feedback interconnections are readily seen to be special cases. We formalize interconnection of stochastic systems in Section 3.

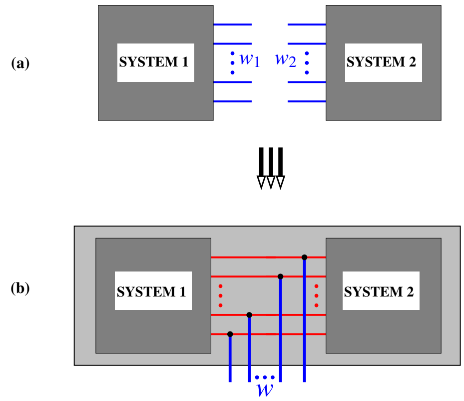

Fig. 2. Interconnection of systems

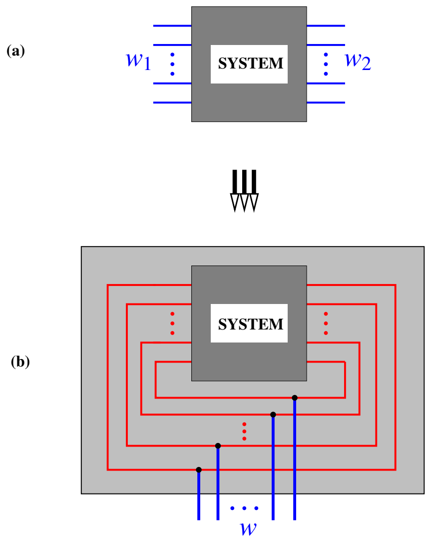

Fig. 3. Interconnection of terminals

By combining the operations explained above, it is possible to obtain complex interconnected systems from simpler subsystems. We have discussed so far the combination of two systems. These operations are readily extended sequentially to more than two systems, and therefore to complex architectures of interconnected systems.

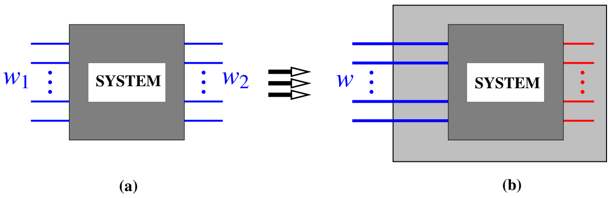

Fig. 4. Elimination of variables

2. STOCHASTIC SYSTEMS

In this section we introduce the concept of a stochastic system, which is nothing else than a probability space as put forward in orthodox mathematical probability theory.

Definition 1: A *stochastic system* is a triple $(\mathbb{W}, \mathscr{E}, P)$ with

- $\mathbb{W}$ a non-empty set, the *outcome space*, with elements of $\mathbb{W}$ called *outcomes*,
- $\mathscr{E}$ a σ-algebra of subsets of $\mathbb{W}$, with elements of $\mathscr{E}$ called *events*,
- P a *probability* measure on $\mathscr{E}$.

The intuitive background underlying this definition is as follows. Assume that we have a stochastic phenomenon that we wish to model. The phenomenon produces variables in the outcome space $\mathbb{W}$. The aim of the model is to specify (i) the subsets of the outcome space to which a probability is assigned and (ii) the numerical value of the probability (in the sense of relative frequency, degree of belief, or whatever interpretation of probability is relevant in the application at hand) that the outcomes belong to such a subset. The set in which the outcomes take on their value is the outcome space $\mathbb{W}$. The set of events $\mathscr{E}$ consists of those subsets of $\mathbb{W}$ to which the model assigns a probability. The probability that the outcomes belong to the set $E \in \mathscr{E}$ is $P(E)$.

Two important special cases are obtained as follows. We refer to these special cases as *classical* stochastic systems.

- A special case is $(\mathbb{W}, 2^{\mathbb{W}}, P)$ with $\mathbb{W}$ a countable set. P can then be specified by giving the probability p of the individual outcomes, $p : \mathbb{W} \to [0,1]$, and defining P by $P(E) = \sum_{e \in E} p(e)$. In this case, every subset of $\mathbb{W}$ is assumed to be an event, and P is completely determined by the probability of the singletons.
- Another special case is $(\mathbb{R}^{\mathtt{n}}, \mathscr{B}(\mathbb{R}^{\mathtt{n}}), P)$ with $\mathscr{B}(\mathbb{R}^{\mathtt{n}})$ the Borel σ-algebra. P can then be specified by a probability distribution on $\mathbb{R}^{\mathtt{n}}$, or, if the distribution is sufficiently smooth, by the probability density function $p : \mathbb{R}^{\mathtt{n}} \to [0, \infty)$ leading to $P(E) = \int_E p(x)\, dx$.

For a classical stochastic system 'essentially every' subset of $\mathbb{W}$ is an event and is therefore assigned a probability. We formalize the second special case as a definition.

Definition 2: The stochastic system $(\mathbb{R}^{\mathtt{n}}, \mathscr{B}(\mathbb{R}^{\mathtt{n}}), P)$ is called a *classical* $\mathtt{n}$*-dimensional random vector*.

Deterministic systems emerge as special cases of stochastic systems, as they should.

Definition 3: The stochastic system $(\mathbb{W}, \mathscr{E}, P)$ is said to be *deterministic* if $\mathscr{E} = \{\emptyset, \mathbb{B}, \mathbb{B}^{complement}, \mathbb{W}\}$ and $P(\mathbb{B}) = 1$. $\mathbb{B}$ is called the *behavior* of the deterministic system.

For a deterministic system, we can state only that outcomes belong to $\mathbb{B}$ with probability one, and to $\mathbb{B}^{\text{complement}}$ with probability zero. Deterministic and classical stochastic systems are extremes of a spectrum ranging from systems with very coarse to systems with very rich σ-algebras.

3. INTERCONNECTION

In this section we discuss interconnection formally. We start by considering the situation discussed in Figure 2 with the assumption that the two interconnected systems are stochastically independent. Note that interconnection comes down to imposing two distinct probabilistic laws on the same set of variables. *Is it possible to define one law which respects both laws?* As we shall see, this is indeed possible, provided a regularity condition, called 'complementarity', is satisfied.

Definition 4: Two σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ on a set $\mathbb{W}$ are said to be *complementary* if for all nonempty sets $E_1, E_1' \in \mathscr{E}_1, E_2, E_2' \in \mathscr{E}_2$ there holds

$$[E_1 \cap E_2 = E_1' \cap E_2'] \Rightarrow [E_1 = E_1' \text{ and } E_2 = E_2'].$$

The stochastic systems $\Sigma_1 = (\mathbb{W}, \mathscr{E}_1, P_1)$ and $\Sigma_2 = (\mathbb{W}, \mathscr{E}_2, P_2)$ are said to be *complementary* if for all $E_1, E_1' \in \mathscr{E}_1$ and $E_2, E_2' \in \mathscr{E}_2$ there holds

$$[E_1 \cap E_2 = E_1' \cap E_2'] \Rightarrow [P_1(E_1)P_2(E_2) = P_1(E_1')P_2(E_2')].$$

In words, complementarity of stochastic systems requires that the intersection of two events, one from each of the σ-algebras, determines the product of the probabilities of the intersecting events uniquely, while complementarity of the σ-algebras requires that the intersection of two sets, one from each of the σ-algebras, determines the intersecting sets uniquely.

Note that

$$[\mathscr{E}_1, \mathscr{E}_2 \text{ complementary}] \Rightarrow [\mathscr{E}_1 \cap \mathscr{E}_2 = \{\emptyset, \mathbb{W}\}]$$

$$[\mathscr{E}_1, \mathscr{E}_2 \text{ complementary}, E_1 \in \mathscr{E}_1, E_2 \in \mathscr{E}_2, \text{ and } E_1 \cap E_2 = \emptyset] \Rightarrow [E_1 = \emptyset \text{ or } E_2 = \emptyset].$$

Complementarity of two stochastic systems is implied by complementarity of the associated σ-algebras. In order to see this, let $E_1, E_1' \in \mathscr{E}_1, E_2, E_2' \in \mathscr{E}_2$. On the one hand, if the sets E_1, E_1', E_2, E_2' are all non-empty and $\mathscr{E}_1, \mathscr{E}_2$ are complementary, then $E_1 \cap E_2 = E_1' \cap E_2'$ implies $E_1 = E_1'$ and $E_2 = E_2'$, and, therefore, $P_1(E_1)P_2(E_2) = P_1(E_1')P_2(E_2')$. On the other hand, assume that at least one of the sets E_1, E_1', E_2, E_2', say E_1, is empty. Then $E_1 \cap E_2 = E_1' \cap E_2'$ implies $E_1' \cap E_2' = \emptyset$. Complementarity of $\mathscr{E}_1, \mathscr{E}_2$ therefore implies that either $E_1' = \emptyset$, or $E_2' = \emptyset$. Consequently, also in this case $E_1 \cap E_2 = E_1' \cap E_2'$ implies $P_1(E_1)P_2(E_2) = 0 = P_1(E_1')P_2(E_2')$.

It is easy to construct examples involving zero probability events that show that complementarity of two stochastic systems does not imply complementarity of the associated σ-algebras. Complementarity of the event σ-algebras is a more primitive condition that is convenient for proving complementarity of stochastic systems.

Definition 5: Let $\Sigma_1 = (\mathbb{W}, \mathscr{E}_1, P_1)$ and $\Sigma_2 = (\mathbb{W}, \mathscr{E}_2, P_2)$ be complementary stochastic systems. Then the *interconnection* of Σ_1 and Σ_2, assumed stochastically independent, denoted by $\Sigma_1 \wedge \Sigma_2$, is defined as the stochastic system

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

with $\mathscr{E}$ the σ-algebra generated by $\mathscr{E}_1 \cup \mathscr{E}_2$, and the probability P defined through 'rectangles' $\{E_1 \cap E_2 \mid E_1 \in \mathscr{E}_1, E_2 \in \mathscr{E}_2\}$ by

$$P(E_1 \cap E_2) := P_1(E_1)P_2(E_2) \text{ for } E_1 \in \mathscr{E}_1, E_2 \in \mathscr{E}_2.$$

The definition of the probability P for rectangles uses complementarity in an essential way. $\mathscr{E}$ is in fact the σ-algebra generated by these rectangles. It is readily seen that the class of subsets of $\mathbb{W}$ that consist of the union of a finite number of disjoint rectangles forms an algebra of sets, that is, a class of subsets of $\mathbb{W}$ that is closed under taking the complement, intersection, and union. The probability of rectangles defines the probability on the subsets of $\mathbb{W}$ that consist of a union of a finite number of disjoint rectangles. By the Hahn-Kolmogorov extension theorem, this leads to a unique probability measure P on $\mathscr{E}$, the σ-algebra generated by the rectangles. This construction of the probability measure P is completely analogous to the construction of a product measure.

The notions of interconnection of stochastic systems and of complementarity of stochastic systems and σ-algebras constitute the main original concepts of this paper, viewed as a contribution to mathematical probability theory. Obviously, there holds $\mathscr{E}_1, \mathscr{E}_2 \subseteq \mathscr{E}$. Also, for $E_1 \in \mathscr{E}_1$ and $E_2 \in \mathscr{E}_2$, we have $P(E_1) = P_1(E_1)$ and $P(E_2) = P_2(E_2)$. Hence interconnection refines the event σ-algebras and the probabilities. This implies in particular that Σ_1 and Σ_2 are unfalsified by $\Sigma_1 \wedge \Sigma_2$. The stochastic system $(\mathbb{W}, \mathscr{E}, P)$ is said to be *unfalsified* by $(\mathbb{W}, \mathscr{E}', P')$ if for all $E \in \mathscr{E} \cap \mathscr{E}'$ there holds $P(E) = P'(E)$.

Note that for $E_1 \in \mathscr{E}_1$ and $E_2 \in \mathscr{E}_2$, $P(E_1 \cap E_2) = P_1(E_1)P_2(E_2) = P(E_1 \cap \mathbb{W})P(\mathbb{W} \cap E_2) = P(E_1)P(E_2)$. Hence $\mathscr{E}_1$ and $\mathscr{E}_2$ are stochastically independent sub-σ-algebras of $\mathscr{E}$. This expresses that Σ_1 and Σ_2 model phenomena that are stochastically independent.

The deterministic systems $(\mathbb{W}, \mathscr{E}_1, P_1)$ and $(\mathbb{W}, \mathscr{E}_2, P_2)$ with behavior $\mathbb{B}_1$ and $\mathbb{B}_2$ respectively, are complementary if either $\mathbb{B}_1 = \mathbb{W}$, or if $\mathbb{B}_2 = \mathbb{W}$, or if $\mathbb{B}_1$ and $\mathbb{B}_2$ are both strict subsets of $\mathbb{W}$ and $\mathbb{B}_1 \cap \mathbb{B}_2 \neq \emptyset$. Their interconnection is equivalent to the deterministic system $(\mathbb{W}, \mathscr{E}, P)$ with behavior $\mathbb{B}_1 \cap \mathbb{B}_2$.

We illustrate interconnection by our two examples.

Example: The interconnected noisy resistor. Consider the interconnection of a noisy resistor and a voltage source with an internal resistance and thermal noise. This leads to the configu-

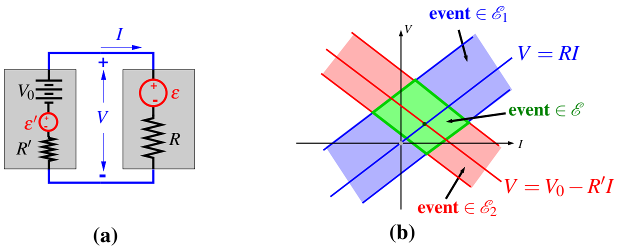

Fig. 5. Interconnection of noisy resistors

ration shown in Figure 5(a). System 1 corresponds to the noisy resistor and is described by $V = RI + \varepsilon$. System 2 correspond to the voltage source, and is described by $V = V_0 - R'I + \varepsilon'$ with V_0 a constant voltage, R' the internal resistance, and ε' a random variable independent of ε. Assume that ε' is gaussian, with zero mean and standard deviation σ'. A rectangular event of the interconnection is shown in Figure 5(b). It is easily seen that the corresponding σ-algebras are complementary if and only if $R + R' \neq 0$. The σ-algebra of the interconnected system is then the Borel σ-algebra on $\mathbb{R}^2$, and $\begin{bmatrix} V \\ I \end{bmatrix}$ is the classical 2-dimensional random vector governed by the equations

$$\begin{bmatrix} 1 & -R \\ 1 & R' \end{bmatrix} \begin{bmatrix} V \\ I \end{bmatrix} = \begin{bmatrix} \varepsilon \\ \varepsilon' + V_0 \end{bmatrix}.$$

Example: Equilibrium price/demand/supply. Consider first the deterministic price/demand and price/supply characteristics of an economic good. Assuming that these characteristics pertain to the same good imposes $\text{price}_1 = \text{price}_2$, while equilibrium imposes demand = supply. We view imposing these conditions as

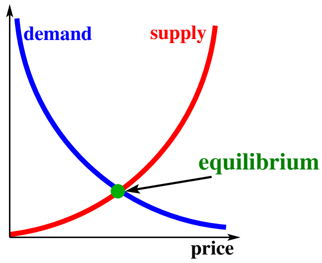

Fig. 6. Deterministic equilibrium price/demand/supply

interconnection. It is readily verified that the interconnection of the deterministic price/demand and price/supply systems yields the deterministic system with equilibrium behavior the intersection of the price/demand and price/supply characteristics as illustrated in Figure 6.

In the stochastic case, we start with the stochastic system $\Sigma_1 = ((0,\infty)^2, \mathscr{E}_1, P_1)$ that models the price/demand, and $\Sigma_2 = ((0,\infty)^2, \mathscr{E}_2, P_2)$ that models the price/supply. The elements of $\mathscr{E}_1$ and $\mathscr{E}_2$ are those to which a probability is assigned (see the discussion of Example 2 in Section 2). Interconnection of Σ_1 and Σ_2 means $p_1 = p_2 = p$ (expressing that the prices pertain to the same good), and $d = s$ (expressing the equilibrium condition demand = supply).

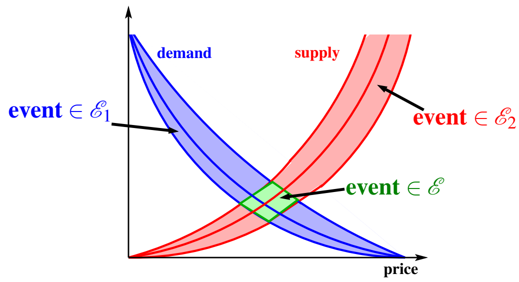

Fig. 7. Price/demand/supply event

Under reasonable conditions (related, for example, to the cardinality, shape, and monotonity of the price/demand and price/supply events) the associated σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ are complementary, and the interconnection σ-algebra consists of the Borel subsets of $(0,\infty)^2$. A rectangular event for the interconnected stochastic system is shown in Figure 7. The probability for the interconnected stochastic system follows the construction of Definition 6.

For interconnection of the stochastic system $\Sigma_1 = (\mathbb{W}, \mathscr{E}_1, P_1)$ with the deterministic system $\Sigma_2 = (\mathbb{W}, \mathscr{E}_2, P_2)$ with behavior $\mathbb{B}$, stochastic independence is trivially satisfied. Σ_1 and Σ_2 are then complementary if and only if

$$[\![E_1, E_1' \in \mathscr{E}_1, \text{ and } E_1 \cap \mathbb{B} = E_1' \cap \mathbb{B}]\!] \Rightarrow [\![P_1(E_1) = P_1(E_1')]\!].$$

Assuming complementarity, interconnection leads to the stochastic system that is equivalent to $(\mathbb{W},\mathscr{E},P)$ with $\mathscr{E} = \mathscr{E}_{\mathbb{B}} \cup \{\mathbb{B}^{\text{complement}},\mathbb{W}\}$, where $\mathscr{E}_{\mathbb{B}} = \{E_1 \cap \mathbb{B} \mid E_1 \in \mathscr{E}_1\}$. The probability P of the interconnection is given by $P(E) = P_1(E_1)$ with E_1 any element of $\mathscr{E}_1$ such that $E = E_1 \cap \mathbb{B}$. This implies that $P(\mathbb{B}) = 1$, and the probability in the interconnected system is therefore concentrated on $\mathbb{B}$.

We now consider the interconnection of terminals as shown in Figure 3. Before interconnection, we have the stochastic system $\Sigma = (\mathbb{W}\times\mathbb{W},\mathscr{E},P)$ with variables (w_1,w_2). Both w_1 and w_2 have their outcomes in $\mathbb{W}$, and these outcomes are coupled through $\mathscr{E}$ and P. The interconnection imposes $w_1 = w_2$ and we wish to consider the stochastic system that governs $w = w_1 = w_2$. This stochastic system a special case of the situation discussed in the previous paragraph with the behavior of the deterministic system defined by $\mathbb{B} = \{(w_1,w_2) \in \mathbb{W}\times\mathbb{W} \mid w_1 = w_2\}$. Complementarity requires that

$$[\![E_1,E_2 \in \mathscr{E} \text{ and } E_1 \cap \mathbb{B} = E_2 \cap \mathbb{B}]\!] \Rightarrow [\![P(E_1) = P(E_2)]\!].$$

Complementarity and interconnection yield the stochastic system $\Sigma' = (\mathbb{W},\mathscr{E}',P')$ with

$$[\![E' \in \mathscr{E}']\!] :\Leftrightarrow [\![\ \exists\, E \in \mathscr{E} \text{ such that } E' = \{(w,w) \mid (w,w) \in E\}]\!]$$
$$P'(E') = P(E).$$

System interconnection (see Figure 2) and terminal interconnection (see Figure 3) are closely related. However, terminal interconnection is more general, since it also deals with interconnection of systems that are not stochastically independent.

4. INTERCONNECTION OF LINEAR STOCHASTIC SYSTEMS

Definition 4: The $\mathtt{n}$-dimensional stochastic system $(\mathbb{R}^{\mathtt{n}},\mathscr{E},P)$ is said to be *linear* if there exists a linear subspace $\mathbb{L}$ of $\mathbb{R}^{\mathtt{n}}$ such that the events are the Borel subsets of the quotient space $\mathbb{R}^{\mathtt{n}}/\mathbb{L}$, and the probability is a Borel probability on $\mathbb{R}^{\mathtt{n}}/\mathbb{L}$. Note that $\mathbb{R}^{\mathtt{n}}/\mathbb{L}$ is a finite dimensional real vector space with, therefore, well-defined Borel sets. $\mathbb{R}^{\mathtt{n}}/\mathbb{L}$ has dimension = $\mathtt{n} - \mathtt{dimension}(\mathbb{L})$. $\mathbb{L}$ is called the *fiber* and $\mathtt{dimension}(\mathbb{L})$ *the number of degrees of freedom* of the linear stochastic system. The stochastic system $(\mathbb{R}^{\mathtt{n}},\mathscr{E},P)$ is said to be *gaussian* if it is linear and if the Borel probability on $\mathbb{R}^{\mathtt{n}}/\mathbb{L}$ is gaussian.

We consider a probability measure that is concentrated on a singleton to be gaussian. More generally, a gaussian probability measure may be concentrated on a linear variety.

The idea behind Definition 4 is illustrated in Figure 8(a). The

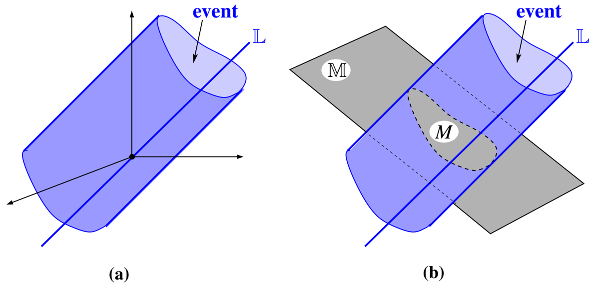

Fig. 8. Events for a linear system

events are cylinders in $\mathbb{R}^{\mathtt{n}}$ with rays parallel to the fiber $\mathbb{L}$. A linear stochastic system is a classical random vector if and only if $\mathbb{L} = \{0\}$. Every classical random vector with $\mathbb{W} = \mathbb{R}^{\mathtt{n}}$ defines a linear stochastic system. At the other extreme, when $\mathbb{L} = \mathbb{R}^{\mathtt{n}}$, the event set $\mathscr{E}$ becomes the trivial σ-algebra $\{\emptyset,\mathbb{R}^{\mathtt{n}}\}$.

A concrete way of thinking about a linear $\mathtt{n}$-dimensional stochastic system is in terms of two linear subspaces $\mathbb{L},\mathbb{M}$ of $\mathbb{R}^{\mathtt{n}}$ that are complementary, $\mathbb{L}\oplus\mathbb{M} = \mathbb{R}^{\mathtt{n}}$, and a Borel probability $P_{\mathbb{M}}$ on $\mathbb{M}$. Take as events the sets of the form

$$E = \{\bigcup_{w\in M} (w+\mathbb{L}) \mid M \text{ a Borel subset of } \mathbb{M}\}$$

(see Figure 8(b)) and $P(E)$ equal to $P_{\mathbb{M}}(M)$. A linear $\mathtt{n}$-dimensional stochastic system is thus parameterized by its linear fiber $\mathbb{L}$, a linear subspace $\mathbb{M}$ complementary to $\mathbb{L}$, and a Borel probability on $\mathbb{M}$.

Let $R \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$ be a matrix of full row rank (that is, $\mathtt{rank}(R) = \mathtt{p}$) and ε a classical $\mathtt{p}$-dimensional random vector with Borel probability P_ε. Consider the equation

$$Rw = \varepsilon \tag{1}$$

describing the stochastic laws of the vector $w \in \mathbb{R}^{\mathtt{n}}$. This equation defines the linear stochastic system $\Sigma = (\mathbb{R}^{\mathtt{n}},\mathscr{E},P)$ with

$$[\![E \in \mathscr{E}]\!] :\Leftrightarrow [\![E = R^{-1}(A) \text{ for some Borel subset } A \subseteq \mathbb{R}^{\mathtt{p}}]\!],$$

$$P\left(R^{-1}(A)\right) := P_\varepsilon(A).$$

R^{-1} denoted the pullback of R. The fiber of this linear stochastic system is $\mathtt{kernel}(R)$. The number of degrees of freedom equals $\mathtt{n}-\mathtt{p}$. We call (1) a *kernel representation* of Σ. Every $\mathtt{n}$-dimensional linear stochastic system admits a kernel representation. Note that (1) defines a gaussian stochastic system if and only if ε is gaussian. An $\mathtt{n}$-dimensional gaussian system with $\mathtt{n}-\mathtt{p}$ degrees of freedom represented by (1) is hence parameterized by the triple (R,m,S) with $R \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$ a matrix of full row rank, $m \in \mathbb{R}^{\mathtt{p}}$ the mean, and $S \in \mathbb{R}^{\mathtt{p}\times\mathtt{p}}, S = S^\top \succeq 0$, the covariance of ε. All triples (R,m,S) that define the same gaussian system are obtained by the transformation group

$$(R,m,S) \underset{U\in\mathbb{R}^{\mathtt{p}\times\mathtt{p}} \text{ nonsingular}}{\longmapsto} (UR,Um,USU^\top). \tag{2}$$

Theorem 1: Consider the linear $\mathtt{n}$-dimensional stochastic systems $\Sigma_1 = (\mathbb{R}^{\mathtt{n}},\mathscr{E}_1,P_1)$ and $\Sigma_2 = (\mathbb{R}^{\mathtt{n}},\mathscr{E}_2,P_2)$ with associated fibers $\mathbb{L}_1$ and $\mathbb{L}_2$. The σ-algebras $\mathscr{E}_1$ and $\mathscr{E}_2$ are complementary if and only if

$$\mathbb{L}_1 + \mathbb{L}_2 = \mathbb{R}^{\mathtt{n}}.$$

The proof is straightforward.

Consider the linear $\mathtt{n}$-dimensional stochastic systems Σ_1 and Σ_2 and assume that $\mathbb{L}_1+\mathbb{L}_2 = \mathbb{R}^{\mathtt{n}}$ is satisfied. Then the interconnected system $\Sigma_1\wedge\Sigma_2$ is again a linear $\mathtt{n}$-dimensional stochastic system. Its fiber is $\mathbb{L}_1\cap\mathbb{L}_2$. Hence $\Sigma_1\wedge\Sigma_2$ is a classical $\mathtt{n}$-dimensional random vector if and only if $\mathbb{L}_1\oplus\mathbb{L}_2 = \mathbb{R}^{\mathtt{n}}$. If Σ_1 and Σ_2 are gaussian, so is $\Sigma_1\wedge\Sigma_2$.

Let $R_1w = \varepsilon_1$ be kernel representation of Σ_1 and $R_2w = \varepsilon_2$ be a kernel representation of Σ_2 with R_1 and R_2 both of full row rank. Then $\mathbb{L}_1+\mathbb{L}_2 = \mathbb{R}^{\mathtt{n}}$ requires $\mathtt{kernel}(R_1)+\mathtt{kernel}(R_2) = \mathbb{R}^{\mathtt{n}}$, which is equivalent to requiring that $R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$ is also of full row rank. Assuming complementarity yields $\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} w = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$ as a kernel representation of the interconnection $\Sigma_1\wedge\Sigma_2$. Its fiber is $\mathtt{kernel}(R) = \mathtt{kernel}(R_1)\cap\mathtt{kernel}(R_2)$.

5. IDENTIFICATION

In this section we discuss some implications to the problem of building models from data of the view of stochastic systems and their interconnection that emerges from the previous sections. The question we deal with is system identification: *how can we recover the laws that govern a stochastic system from measurements?*

Consider the stochastic system $\Sigma = (\mathbb{W}, \mathscr{E}, P)$. Assume that outcomes, realizations of the variables $w \in \mathbb{W}$, of the phenomenon that is modeled by Σ are observed. The aim is to identify the model, that is $\mathscr{E}$ and P, from the observations. In order to generate these observations, experimental conditions need to be set up during the data collection process. The data do not emerge from the stochastic system all by itself, but from observing Σ in interaction with an environment (see Figure 9). One of the questions that arises is whether it is possible to disentangle from the data the laws of the stochastic system from the laws of the environment. We make a clear distinction between modeling a stochastic system from data and obtaining the statistical features of a random vector from samples. The latter problem consists of inferring the statistical laws by sampling a random vector in an experimental set-up, while the former problem requires in addition disentangling the laws of the system from the laws of the environment that was active while sampling.

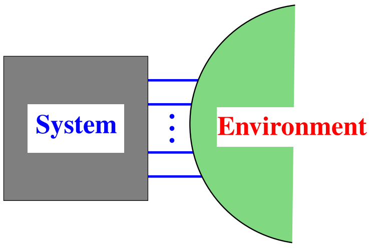

Fig. 9. Data collection

Let us illustrate this issue by means of the noisy resistor. The variables $\begin{bmatrix} V \\ I \end{bmatrix}$ are governed by $V = RI + \varepsilon$ and the identification problem consists in deducing the parameters of the model, that is R and σ, the standard deviation of ε, from measured voltage/current pairs. These measurements may be generated in various ways. One possibility is to fix the current by driving the noisy resistor by a constant current source and measure various realizations of the voltage. Another possibility is to fix the voltage by putting a constant voltage source across the noisy resistor and measure various realizations of the current. A third possibility is to terminate the noisy resistor by a voltage source with internal resistance and thermal noise as shown in Figure 5, and measure various realizations of the voltage/current pair. These terminations of the noisy resistor give rise to three data clouds, with completely different statistical features each, and from each of these data clouds we may attempt to deduce the parameters R and σ.

For the noisy resistor it may be reasonable to assume that the experimenter can control the environmental conditions that are active during data collection. On the other hand, in many situations, for instance in economics, in the social sciences, or in biology, the data are collected in a passive way, *in vivo*, so to speak. The problem of disentangling the laws of the system from the laws of the environment then becomes imperative. As an example, assume that we wish to identify the stochastic system that governs the price/demand of an economic good. We could attempt to deduce the laws of this stochastic system from observing various realizations of the variables (p, d). If these measurements are obtained under the equilibrium condition demand = supply, then, as shown in Section 3, under reasonable conditions, the data are realizations of a classical 2-dimensional random vector, and then the probability distribution of (p, d) depends not only on the stochastic price/demand system, but also on the stochastic price/supply system. The stochastic laws of the price/supply may also be unknown. Is it nevertheless possible to identify the stochastic price/demand system from the data?

In this paper we discuss only a very special case of the identification problem. We assume that the system to be identified is an $\mathtt{n}$-dimensional gaussian stochastic system. We further assume that the data are collected while the system is interconnected with another $\mathtt{n}$-dimensional gaussian stochastic system that is stochastically independent and complementary to the system to be identified, and such that the interconnected system is a classical random vector. As we have seen in Section 4, this classical random vector is also gaussian and we assume that from sampling, its mean and covariance matrix have been deduced. We assume therefore that the data consist of the mean and covariance of the probability distribution of the outcomes in the interconnected system.

Let $\mathbb{L} \subseteq \mathbb{R}^{\mathtt{n}}$ be the fiber of the gaussian system $\Sigma = (\mathbb{R}^{\mathtt{n}}, \mathscr{E}, P)$ to be identified and let $Rw = \varepsilon$ be a kernel representations of Σ. $R \in \mathbb{R}^{\mathtt{p} \times \mathtt{n}}$ is a matrix of full row rank and $\mathbb{L} = \mathtt{kernel}(R)$. Since Σ is assumed to be gaussian, ε is a classical gaussian $\mathtt{p}$-dimensional random vector. Let $m \in \mathbb{R}^{\mathtt{p}}$ be the mean and $S \in \mathbb{R}^{\mathtt{p} \times \mathtt{p}}, S = S^\top \succeq$ the covariance of ε. Let $\mathbb{L}' \subseteq \mathbb{R}^{\mathtt{n}}$ denote the fiber of the gaussian system $\Sigma' = (\mathbb{R}^{\mathtt{n}}, \mathscr{E}', P')$ that is interconnected with Σ during data collection. Assume that Σ and Σ' are stochastically independent and that $\mathbb{L} \oplus \mathbb{L}' = \mathbb{R}^{\mathtt{n}}$. Then, as shown in Section 4, the σ-algebras of Σ and Σ' are complementary and the interconnected system $\Sigma_{\text{observed}} = \Sigma \wedge \Sigma'$ is a classical $\mathtt{n}$-dimensional stochastic system $(\mathbb{R}^{\mathtt{n}}, \mathscr{B}(\mathbb{R}^{\mathtt{n}}), P_{\text{observed}})$ with P_{observed} a gaussian probability distribution on $\mathbb{R}^{\mathtt{n}}$. Let $\quad \in \mathbb{R}^{\mathtt{n}}$ be its the mean and $\Gamma \in \mathbb{R}^{\mathtt{n} \times \mathtt{n}}, \Gamma = \Gamma^\top \succeq 0$ its covariance.

Σ is unfalsified by Σ_{observed} if and only if

$$R \quad = m \quad \text{and} \quad R\Gamma R^\top = S.$$

The disentanglement question becomes: Is it possible to deduce from these equations Σ, that is (R, m, S) up to the equivalence (2), from Σ_{observed}, that is from $(\quad , \Gamma)$?

Let $R'w = \varepsilon'$ be a kernel representation of Σ'. $R' \in \mathbb{R}^{(\mathtt{n}-\mathtt{p}) \times \mathtt{n}}$ is a matrix of full row rank with $\mathtt{kernel}(R') = \mathbb{L}'$. Let $m' \in \mathbb{R}^{(\mathtt{n}-\mathtt{p})}$ be the mean and $S' \in \mathbb{R}^{(\mathtt{n}-\mathtt{p}) \times (\mathtt{n}-\mathtt{p})}, S' = S'^\top \succeq 0$ the covariance of ε'. Since Σ and Σ' are assumed to be stochastically independent, ε and ε' are independent. $\mathbb{L} \oplus \mathbb{L}' = \mathtt{kernel}(R) \oplus \mathtt{kernel}(R') = \mathbb{R}^{\mathtt{n}}$ implies that the matrix $\begin{bmatrix} R \\ R' \end{bmatrix} \in \mathbb{R}^{\mathtt{n} \times \mathtt{n}}$ is nonsingular. Hence

$$\begin{bmatrix} R \\ R' \end{bmatrix} w = \begin{bmatrix} \varepsilon \\ \varepsilon' \end{bmatrix}$$

is a kernel representation of $\Sigma_{\text{observed}} = \Sigma \wedge \Sigma'$. The mean $\quad$ and covariance Γ of $\Sigma_{\text{observed}} = \Sigma \wedge \Sigma'$ are related to the parameters R, m, S, R', m', S' of Σ and Σ' by

$$\begin{bmatrix} R \\ R' \end{bmatrix} \quad = \begin{bmatrix} m \\ m' \end{bmatrix}$$

$$\begin{bmatrix} R \\ R' \end{bmatrix} \Gamma \begin{bmatrix} R \\ R' \end{bmatrix}^\top = \begin{bmatrix} S & O_{\mathtt{p}\times(\mathtt{n}-\mathtt{p})} \\ O_{(\mathtt{n}-\mathtt{p})\times \mathtt{p}} & S' \end{bmatrix}. \qquad (3)$$

The following theorem shows the extent to which it is possible to deduce the parameters R, m, S, R', m', S' of Σ and Σ' from the parameters $\ ,\Gamma$ of $\Sigma_{\text{observed}} = \Sigma \wedge \Sigma'$.

Theorem 2: Let $\ \in \mathbb{R}^{\mathtt{n}}$ and $\Gamma \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}, \Gamma = \Gamma^\top \succeq 0$ be given. For every $R' \in \mathbb{R}^{(\mathtt{n}-\mathtt{p})\times\mathtt{n}}$ of full row rank, there exist

(1) $R \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$ with $\begin{bmatrix} R \\ R' \end{bmatrix} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ nonsingular,
(2) $m \in \mathbb{R}^{\mathtt{p}}$ and $m' \in \mathbb{R}^{\mathtt{n}-\mathtt{p}}$,
(3) $S \in \mathbb{R}^{\mathtt{p}\times\mathtt{p}}, S = S^\top \succeq$ and $S' \in \mathbb{R}^{(\mathtt{n}-\mathtt{p})\times(\mathtt{n}-\mathtt{p})}, S' = S'^\top \succeq 0$

such that (3) holds. If $R'\Gamma R'^\top \succ 0$, then R, m, S are uniquely determined by (3), up to the equivalence (2).

Proof: By choosing suitable bases, we can assume that $R' = [\ O_{(\mathtt{n}-\mathtt{p})\times\mathtt{p}} \quad I_{(\mathtt{n}-\mathtt{p})\times(\mathtt{n}-\mathtt{p})}\ ]$. Choose $R = [\ I_{\mathtt{p}\times\mathtt{p}} \quad -L\ ]$ with $L \in \mathbb{R}^{\mathtt{p}\times(\mathtt{n}-\mathtt{p})}$ to be determined. Clearly $\begin{bmatrix} R \\ R' \end{bmatrix} \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ is nonsingular. Partition $\ $ and Γ conformably to R', as

$$= \begin{bmatrix} \ _1 \\ \ _2 \end{bmatrix} \quad \text{and} \quad \Gamma = \begin{bmatrix} \Gamma_{1,1} & \Gamma_{1,2} \\ \Gamma_{2,1} & \Gamma_{2,2} \end{bmatrix}.$$

Equations (3) become

$$m = \ _1 + L \ _2, \ m' = \ _2,$$

$$S = \Gamma_{1,1} - \Gamma_{1,2}L^\top - L\Gamma_{2,1} + L\Gamma_{2,2}L^\top, \ \ S' = \Gamma_{2,2}, \ \ \Gamma_{1,2} = L\Gamma_{2,2}.$$

These equations define m, m', S, S', provided there exists L such that $\Gamma_{1,2} = L\Gamma_{2,2}$. $\Gamma \succeq 0$ implies that $\mathtt{kernel}(\Gamma_{2,2}) \subseteq \mathtt{kernel}(\Gamma_{1,2})$. Hence there indeed exists an L such that $\Gamma_{1,2} = L\Gamma_{2,2}$. Hence there exist then L, m, S, m', S' such that (3) holds. ■

Since $R'\Gamma R'^\top \succ 0$ corresponds to $\Gamma_{2,2} \succ 0$, this implies that the solution L is unique and given by $L = \Gamma_{1,2}\Gamma_{2,2}^{-1}$. Hence there then exist unique L, m, S, m', S' such that (3) holds.

The above theorem of course also holds with the roles of Σ and Σ' reversed. The theorem shows that without further assumptions on Σ or Σ', it is not possible to deduce the laws of Σ from the laws of Σ_{observed}. In fact, Σ being unfalsified from Σ_{observed} leaves the fiber of Σ completely unspecified. So, not only is Σ unidentifiable from Σ_{observed}, but the deterministic part of Σ, governed by $Rw = 0$, is left completely arbitrary. Without further structural information on the system or on the environment, it is not possible to recover the parameters of Σ from sampling. The theorem also implies that the parameters $\ ,\Gamma$ of Σ_{observed} together with the fiber $\mathbb{L}'$ of Σ' specify Σ and Σ' uniquely, provided $R'\Gamma R'^\top \succ 0$. The condition $R'\Gamma R'^\top \succ 0$ is called *sufficiency of excitation*. It requires that there is an adequate variety of experiments generated by the environment.

As a concrete example, consider gaussian linear regression. Partition w as $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ with $w_1 \in \mathbb{R}^{\mathtt{n}_1}$ and $w_2 \in \mathbb{R}^{\mathtt{n}_2}$. The stochastic system Σ_1 to be identified is described by

$$w_1 = Lw_2 + \varepsilon,$$

with $L \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ and with ε an $\mathtt{n}_1$-dimensional gaussian random vector with mean m and covariance S. The environment Σ_2 is described by

$$w_2 = \varepsilon',$$

with ε' an $\mathtt{n}_2$-dimensional gaussian random vector that is independent of ε. Denote the mean of ε' by m' and the covariance by S'. The form of this equation implies that the fiber of Σ_2 is known and equal to $\mathtt{image}\left(\begin{bmatrix} I_{\mathtt{n}_1\times\mathtt{n}_1} \\ 0_{\mathtt{n}_2\times\mathtt{n}_1} \end{bmatrix}\right)$. The above theorem guarantees therefore that in this case it is possible to identify parameters L, m, m', S, S' by sampling. Persistency of excitation means $S' \succ 0$. If this condition is satisfied then the parameters $L, m, m'S, S'$ are uniquely identifiable by sampling.

There are various further structural conditions that can be given on Σ or Σ' and that imply identifiability. For the noisy resistor terminated by a voltage source with internal resistance and thermal noise (see Figure 5), the following conditions are sufficient for identifiability of the noisy resistor parameters R and σ: either (i) knowledge of R', assuming $\sigma' > 0$, or (ii) knowledge of $V_0 \neq 0$.

For the economic example the full complexity of the identifiability question emerges. Sampling under equilibrium conditions does not lead to identification of the price/demand elasticity. A more elaborate controlled experiment is needed to entangle the price/demand and price/supply systems.

As we have already mentioned there are many applications in statistics in which one attempts to identify the stochastic laws governing a phenomenon involving two real variables. As we remarked, such a law often leads to a coarse σ-algebra. The important observation here is that data generation through sampling requires interconnection with another system, and therefore data collection involves *two* distinct random systems. One of these stochastic systems expresses the intrinsic random laws one is after, while the other expresses the features of the environment that happens to be acting during the data collection experiment. Disentangling these laws requires further structural assumptions on the experimental set-up.

ACKNOWLEDGEMENTS

I thank Tzvetan Ivanov for many enlightening discussions regarding this paper.

The SCD-SISTA research of the KU Leuven is supported by the projects. *Research Council KUL*: GOA/11/05 Ambiorics, GOA/10/09 MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC) en PFV/10/002 (OPTEC), IOF-SCORES4CHEM, several PhD, postdoc & fellowship grants. *Flemish Government: FWO:* PhD/postdoc grants, projects: G0226.06 (cooperative systems and optimization), G0321.06 (Tensors), G.0302.07 (SVM/Kernel), G.0320.08 (convex MPC), G.0558.08 (Robust MHE), G.0557.08 (Glycemia2), G.0588.09 (Brain-machine), Research communities (WOG: ICCoS, ANMMM, MLDM); G.0377.09 (Mechatronics MPC); *IWT*: PhD Grants, Eureka-Flite+, SBO LeCoPro, SBO Climaqs, SBO POM, O&O-Dsquare. *Belgian Federal Science Policy Office*: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011), IBBT. *EU*: ERNSI; FP7-HD-MPC (INFSO-ICT-223854), COST intelliCIS, FP7-EMBOCON (ICT-248940), FP7-SADCO (MC ITN-264735), ERC HIGHWIND (259 166). *Contract Research*: AMINAL, Helmholtz: viCERP, ACCM.

6. REFERENCES

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

[2] Willems, J.C. (2010), Terminals and ports, *IEEE Circuits and Systems Magazine*, 10, 8–16.