# Unknown Input Estimation Techniques in Networks and Applications to Open Channel Hydraulic Systems

PHD Thesis

**Author:** Siro Pillosu

**Advisor:** Prof. Elio Usai

Dottorato di ricerca in Ingegneria Industriale
Universitá degli Studi di Cagliari

March 2012

# INDEX

*To Dulcinea*

***Acknowledgements***

This work contains some results attained during my PhD (from 2009 to 2011) at the Department of Electrical and Electronic Engineering (DIEE) of the University of Cagliari.

The research presented here would not been possible without the unvaluable support and teachings of Elio Usai, the advisor of the present work, and Alessandro Pisano, co-worker and co-advisor. I would also like to thank prof. S.S. Stankovic for his precious advice during my period of study in Belgrade.

# Thesis' overview

In the Introduction, some motivations for the use of Unknown Input Observers (UIO) approach are discussed to state the framework within which this work is developed.

This thesis is divided in two fundamental parts, namely, the **Part I**, in which the theoretical background of the UIO, Consensus Algorithms and Decentralized Systems is discussed; after a collection of algorithms is presented. In the **Part II**, some important applicative problems are addressed and solved by means of the proposed approaches.

More specifically, as for the Part I, in *Chapter 1* the fundamentals regarding the matrix and graph theory are recalled. In the subsequent *Chapter 2* the attention is focused on the strong observability approach, and its main features are described. *Chapter 3* refers to the presentation of the Consensus algorithm, while in *Chapter 4* an estimation algorithm is recalled, which allows the estimation of the state in an "overlapped" system also in presence of Unknown Inputs (in *Chapter 5*), which are estimated as well.

In the Part II the estimation problems of flow ad infiltration, in open channel hydraulic system are solved, using a UIO approach(in *Chapters 6*). In *Chapters 7*, considering open channel hydraulic system, the UIO approach is used to solve a problem of fault detection and compensation.

# Introduction and Motivations

In recent years, motivated by a large amount of important practical problems, the estimation of uncertain systems has become an important subject of research. As a result, considerable progresses in estimation techniques, such as the introduction of consensus algorithm, unknown input observers (UIO) techniques and others, that explicitly account for an imprecise description of the model of the controlled plant guaranteeing the attainment of the relevant estimation objectives in the face of modeling error and/or parameter uncertainties, have been attained.

This work analyzes a quite recent development of unknown input observers (UIO) techniques, which is encountering a growing attention in the control research community.

The objective of this thesis is to survey the theoretical background of the unknown input observers and consensus based estimation, mainly developed in the last years, to present some new results, and to show that the UIO approach is an effective solution to the above-cited drawbacks and may constitute a good candidate for solving a wide range of important practical problems, like the estimation problem of unknown parameters in hydraulic networks, which is considered in last chapters of this thesis.

# Part I

# Theory

# 1 Graph Theory Background

## 1.1 Preliminaries

The main purpose of this chapter is to provide some mathematical foundations, about graph theory and nonnegative matrices, to the considered estimation problem.

We shall strive for rigor in presentation and shall not discuss the applicability of the concepts to real world. This is postponed to later chapters. In this chapter we begin by surveying some basic notions from graph theory [1], [2] and recall an important result about connectivity of digraphs. Next, we explore the theory of nonnegative matrices with emphasis on the connections between nonnegative matrices and directed graphs. An excellent reference on nonnegative matrices is [3].

## 1.2 Matrix Theory

Here, we present basic notions and results about matrix theory, following the treatments in [5] and [6]. Let $\mathbb{R}^{n\times m}$ and $\mathbb{C}^{n\times m}$ denote the set of $n \times m$ real and complex matrices respectivily. Given a real matrix $A$ and a complex matrix $U$, we denote with $A^T$ and $U^*$ the transpose of $A$ and the conjugate transpose matrix of $U$, respectively. $I_n$ denotes the $n \times n$ identity matrix. For a square matrix $A$, we write $A > 0$, resp. $A \geq 0$, if $A$ is symmetric positive definite, resp. symmetric positive semidefinite. For a real matrix $A$, we denote with $rank(A)$ the rank of $A$, respectively. Given a vector $v$, we denote with $diag(v)$ the diagonal square matrix whose diagonal elements are equal to the component $v$.

### 1.2.1 Matrix sets

A matrix $A \in \mathbb{R}^{n\times n}$ with entries $a_{ij}, (i, j \in 1, ..., n)$, is

(i) **Orthogonal** if $AA^T = I_n$, and is special orthogonal if it is orthogonal with $det(A) = +1$. The set of orthogonal matrices is a group.

(ii) **Nonnegative** (resp., positive) if all its entries are nonnegative (resp., positive).

(iii) **Row-stochastic** (or stochastic for brevity) if it is nonnegative and $\sum_{j=1}^{n} a_{ij} = 1$, for all $i \in 1, ..., n$; in other words, $A$ is rowstochastic if

$$A\mathbf{1}_n = \mathbf{1}_n. \tag{1}$$

where $1_n$ denotes the following vector

$$1_n = [1, 1, ..., 1]^T \in \mathbb{R}^n. \quad (2)$$

(iv) **Column-stochastic** if it is nonnegative and $\sum_{j=1}^{n} a_{ij} = 1$, for all $j \in 1, ..., n$.

(v) **Doubly stochastic** if $A$ is both row-stochastic and column-stochastic.

The scalars $\mu_1, ..., \mu_k$ are convex combination coefficients if $\mu_i \geq 0$, for $i \in 1, ..., k$, and $\sum_{i=1}^{k} \mu_i = 1$. (Each row of a row-stochastic matrix contains convex combination coefficients.) A convex combination of vectors is a linear combination of the vectors with convex combination coefficients. A subset $U$ of a vector space $V$ is convex if the convex combination of any two elements of $U$ takes value in $U$. The set of stochastic matrices and the set of doubly stochastic matrices are convex.

(vi) **Normal** if $A^T A = AA^T$ .

(vii) $A$ is a **permutation matrix** if $A$ has precisely one entry equal to one in each row, one entry equal to one in each column, and all other entries equal to zero. The set of permutation matrices is a group.

(viii) An $n \times n$ matrix $A = (a_{ij})$ is said to be **dominant diagonal** if

$$|a_{ii}| > \sum_{j=1, j \neq i}^{n} |a_{ij}|, \forall i \in \mathbb{N} \quad (3)$$

If, in addition, $a_{ii} > 0$ ($a_{ii} < 0$ for all $i \in \mathbb{N}$, then $A$ is **dominant positive** (negative) **diagonal**

(ix) An $n \times n$ matrix $A = (a_{ij})$ is said to be **quasidominant diagonal** if there exist positive numbers $d_j$, $j \in \mathbb{N}$, such that either

$$d_i|a_{ii}| > \sum_{j=1, j \neq i}^{n} d_j|a_{ij}|, \forall i \in \mathbb{N} \quad (4)$$

or

$$d_j|a_{ii}| > \sum_{i=1, i \neq j}^{n} d_i|a_{ij}|, \forall j \in \mathbb{N} \quad (5)$$

Again, if all $a_{ii}$ are positive (negative) and either (4) or (5) is true, then $A$ is quasidominant positive (negative) diagonal. If in (4) $d_i = 1$ for all $i \in \mathbb{N}$ then it reduces to (3). It will be shown later that (4) and (5) are equivalent, whereas an analogous statement for (3) is not true in general.

(x) An $n \times n$ matrix $A = (a_{ij})$ is a **Metzler** matrix if

$$
\begin{aligned}
a_{ij} &< 0 \quad \forall i = j, &(6)\\
a_{ij} &\geq 0 \quad \forall i \neq j, \quad i, j = 1, 2, ...n. &(7)
\end{aligned}
$$

.

**Theorem 1** *A Metzler matrix A is Hurwitz if and only if it is quasidominant negative diagonal [11].*

**Example 1** *Let us consider the following matrices:*

$$
A_1 = \begin{bmatrix} -2 & 1 & 0 \\ 2 & 0 & 2 \\ 3 & 2 & -1 \end{bmatrix}; A_2 = \begin{bmatrix} -2 & 1 & 0 \\ 2 & -1 & 2 \\ 3 & 0 & -1 \end{bmatrix}; A_3 = \begin{bmatrix} -2 & 1 & 0 \\ 2 & -4 & 2 \\ 3 & 0 & -4 \end{bmatrix}; \tag{8}
$$

*First matrix $A_1$ is not Metzler because $A_1(2,2)$ is not $< 0$.*

*Matrix $A_2$ is Metzler because all diagonal terms are $< 0$ and the others are $\geq 0$, but it is neither diagonal nor quasi dominant diagonal. So it is not Hurwitz and in fact,*

$$
eig(A_2) = [0.8455, -2.4227 + 1.1077i, -2.4227 - 1.1077i] \tag{9}
$$

*Matrix $A_3$ is Metzler because all diagonal terms are $< 0$ and the others are $\geq 0$, it is not diagonal dominant but it is quasi dominant diagonal. So it is Hurwitz and in fact,*

$$
eig(A_3) = [-0.793, -4.6035 + 1.2275i, -4.6035 - 1.2275i] \tag{10}
$$

### 1.2.2 Eigenvalues, singular values, and induced norms

Let us introduce the notion of eigenvalue and of simple eigenvalue, that is, an eigenvalue with algebraic and geometric multiplicity equal to 1. The set of eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$ is called its spectrum and is denoted by $spec(A) \in \mathbb{C}^n$. The singular values of the matrix $A \in \mathbb{R}^{n \times n}$ are the positive square roots of the eigenvalues of $A^T A$. We begin with a well-known property of the spectrum of a matrix.

**Theorem 2** *(Gersgorin disks). Let $A$ be an $n \times n$ matrix. Then*

$$
spec(A) \subset \bigcup_{i=1,...,n} \{z \in \mathbb{C} | \|z - a_{ii}\|_{\mathbb{C}} \leq \sum_{j=1, j \neq i}^{n} |a_{ij}|\}. \tag{11}
$$

For a square matrix $A = (a_{ij})$, around every entry $a_{ii}$ on the diagonal of the matrix, draw a closed disc of radius $\sum_{j \neq i} |a_{ij}|$. Such discs are called Gershgorin discs. Every eigenvalue of $A$ lies in a Gershgorin disc.

**Example 2** *Let us consider the following matrix:*

$$A_4 = \begin{bmatrix} 0 & 5 & 0 & 0 \\ 2 & 0 & 0 & 6 \\ 0 & 3 & 0 & 0 \\ 1.5 & 0 & 0.5 & 0 \end{bmatrix}; \tag{12}$$

*For the matrix $A_4$ in (2), the Gershgorin discs are drawn in Fig.1 and the eigenvalues lie in the union of these discs, i.e., the largest disc.*

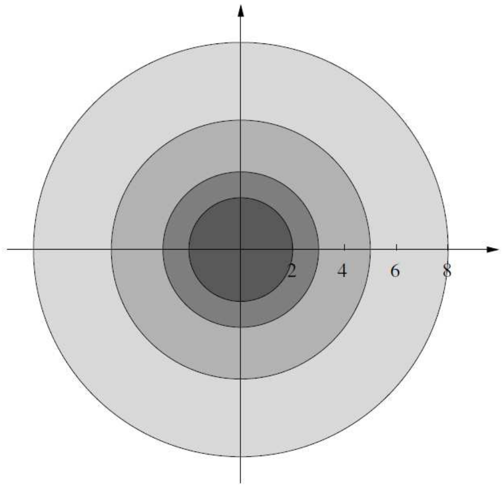



Figure 1: Gershgorin discs

Next, we review a few facts about normal matrices, their eigenvectors and their singular values.

**Lemma 1** *(Normal matrices [15] ). For a matrix $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent:*

*(i) $A$ is normal;*

*(ii) $A$ has a complete orthonormal set of eigenvectors; and*

*(iii) $A$ is unitarily similar to a diagonal matrix, that is, there exists a unitary matrix $U$ such that $U^*AU$ is diagonal.*

**Lemma 2** *(Singular values of a normal matrix [15]). If a normal matrix has eigenvalues $\lambda_1, ..., \lambda_n$, then its singular values are $|\lambda_1|, ..., |\lambda_n|$. Real symmetric matrices are normal, are* *diagonalizable by orthogonal matrices, and have real eigenvalues.*

We conclude by defining the notion of induced norm of a matrix. For $p \in [1, 2, ..., \infty]$, the p-induced norm of $A \in \mathbb{R}^{n \times n}$ is

$$x \in C^n; x = [x_1, x_2, ..., x_n]; \|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}; \|A\|_p = max\{\|Ax\|_p; \|x\|_p = 1\}. \tag{13}$$

One can see that

$$\begin{aligned} \|A\|_1 &= max_{j \in \{1,...,n\}} \sum_{i=1}^{n} |a_{ij}|, &(14)\\ \|A\|_\infty &= max_{i \in \{1,...,n\}} \sum_{j=1}^{n} |a_{ij}|, &(15)\\ \|A\|_2 &= max\{\sigma | \sigma \text{ } is \text{ } a \text{ } singular \text{ } value \text{ } of \text{ } A\}. &(16)\end{aligned}$$

### 1.2.3 Spectral radius and convergent matrices

The spectral radius of a matrix $A \in \mathbb{R}^{n \times n}$ is

$$\rho(A) = max\{\|\lambda\|_{\mathbb{C}}; \lambda \in spec(A)\}. \tag{17}$$

In other words, $\rho(A)$ is the radius of the smallest disk centered at the origin that contains the spectrum of $A$.

**Lemma 3** *(Induced norms and spectral radius [15]). For any square matrix A and in any norm* $p \in [1, 2, ..., \infty]; \rho(A) \leq \|A\|_p$ .

We will often deal with matrices with an eigenvalue equal to 1 and all other eigenvalues strictly inside the unit disk. Accordingly, we generalize the notion of spectral radius as follows. For a square matrix $A$ with $\rho(A) = 1$, we define the essential spectral radius

$$\rho_{ess}(A) = max\{\|\lambda\|_{\mathbb{C}}; \lambda \in spec(A) \setminus \{1\}\}. \tag{18}$$

Next, we shall consider matrices with useful convergence properties.

**Definition 1** *(Convergent and semi-convergent matrices). A matrix* $A \in \mathbb{R}^{n\times n}$ *is*

*(i) semi-convergent if* $lim_{\ell\to+\infty}A^{\ell}$ *exists; and*

*(ii) convergent if it is semi-convergent and* $lim_{\ell\to+\infty}A^{\ell} = 0$ *.*

These two notions are characterized as follows.

**Lemma 4** *The square matrix* $A$ *is convergent if and only if* $\rho(A) < 1$*. Furthermore,* $A$ *is semiconvergent if and only if the following three properties hold:*

*(i)* $\rho(A) \leq 1$*;*

*(ii)* $\rho_{ess}(A) < 1$*, that is, 1 is an eigenvalue and* $\lambda_1 = 1$ *is the only eigenvalue on the unit circle; and*

*(iii) the eigenvalue* $\lambda_1 = 1$ *is semisimple, that is, it has equal algebraic and geometric multiplicity (possibly larger than one). In other words,* $A$ *is semi-convergent if and only if there exists a nonsingular matrix* $T$ *such that*

$$A = T\begin{bmatrix} I_k & 0 \\ 0 & B \end{bmatrix}T^{-1} \tag{19}$$

*where* $B \in \mathbb{R}^{(n-k)\times(n-k)}$ *is convergent, that is,* $\rho(B) < 1$*. With this notation, we have* $\rho_{ess}(A) = \rho(B)$ *and the algebraic and geometric multiplicity of the eigenvalue* $\lambda_1 = 1$ *is* $k$ *[15].*

**Example 3** *Let us consider the following matrix* $A$*:*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & -2 \end{bmatrix} \tag{20}$$

*A is semi-convergent, in fact its eigenvalues are:*

$$\lambda_1 = 1; \lambda_2 = 0.866i; \lambda_3 = -0.866i; \tag{21}$$

*If we consider the following non-singular matrix* $T$*:*

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{22}$$

*we can write matrix A like in (19) where:*

$$I_k = 1; B = \begin{bmatrix} 0.5 & -1 \\ 1 & -0.5 \end{bmatrix} \tag{23}$$

### 1.2.4 Perron-Frobenius theory

Positive and nonnegative matrices enjoy useful spectral properties. In what follows, Theorem amounts to the original Perrons Theorem for positive matrices and the successive theorems , are the extensions due to Frobenius for certain nonnegative matrices. We refer to [5] for a detailed treatment.

**Theorem 3** *(Perron-Frobenius for positive matrices [15]). If the square matrix $A$ is positive, then*

*(i) $\rho(A) > 0$;*

*(ii) $\rho(A)$ is a simple eigenvalue, of $A$ and $\rho(A)$ is strictly larger than the magnitude of any other eigenvalue;*

*(iii) the eigenvalue $\rho(A)$ has an eigenvector with positive components.*

Requiring the matrix to be strictly positive is a key assumption that limits the applicability of this theorem. It turns out that it is possible to obtain the same results of the theorem under weaker assumptions.

**Definition 2** *(Irreducible matrix [15]). A nonnegative matrix $A \in \mathbb{R}^{n\times n}$ is irreducible if, for any nontrivial partition $J \cup K$ of the index set 1, . . . , n, with $J \bigcap K = \{.\}$, there exist $j \in J$ and $k \in K$ such that $a_{jk} \neq 0$.*

**Remark 1** *(Properties of irreducible matrices [15]). An equivalent definition of irreducibility is given as follows. A matrix $A \in \mathbb{R}^{n\times n}$ is irreducible if it is not reducible, and is reducible if either:*

*(i) n = 1 and $A = a_{11} \in \Re$, with $a_{11} = 0$ ; or*

*(ii) there exists a permutation matrix $P \in \mathbb{R}^{n\times n}$ and an integer number $r \in \{1, ..., n-1\}$ such that $P^T AP$ is block upper triangular with diagonal blocks of dimensions $r\times r$ and $(n-r)\times(n-r)$.*

It is an immediate consequence that the property of irreducibility depends upon only the patterns of zeros and nonzero elements of the matrix. We can now weaken the assumption in Theorem 3 and obtain a comparable, but weaker, result for irreducible matrices.

**Theorem 4** *(PerronFrobenius for irreducible matrices [15]). If the nonnegative square matrix $A$ is irreducible, then*

*(i) $\rho(A) > 0$;*

*(ii) $\rho(A)$ is an eigenvalue, and it is simple; and*

*(iii) $\rho(A)$ has an eigenvector with positive components.*

In general, the spectral radius of a nonnegative irreducible matrix does not need to be the only eigenvalue of maximum magnitude. For example, the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has eigenvalues $\{1, -1\}$. In other words, irreducible matrices do indeed have weaker spectral properties than positive matrices. Therefore, it remains unclear which nonnegative matrices have the same properties as those stated for positive matrices in Theorem 3.

**Definition 3** *(Primitive matrix). A nonnegative square matrix $A$ is primitive if there exists $k \in \mathbb{N}$ such that $A^k$ is positive.*

It is easy to see that if a nonnegative square matrix is primitive, then it is irreducible. In later sections we will provide a graph-theoretical characterization of primitive matrices; for now, we are finally in a position to sharpen the results of Theorem 4.

**Theorem 5** *(PerronFrobenius for primitive matrices [15]). If the nonnegative square matrix $A$ is primitive, then it is also irreducible, which means that*

*(i) $\rho(A) > 0$;*

*(ii) $\rho(A)$ is an eigenvalue, it is simple, and $\rho(A)$ is strictly larger than the magnitude of any other eigenvalue; and*

*(iii) $\rho(A)$ has an eigenvector with positive components.*

We conclude this section by noting the following convergence property that is an immediate corollary to Lemma 4 and to Theorem 5.

**Corollary 1** . *If the nonnegative square matrix $A$ is primitive, then the matrix $\rho(A)^{-1}A$ is semi-convergent [15].*

### 1.3 Digraphs, Neighbors and Degrees

A directed graph, (in short, **digraph**) of order $n$ is a pair $G = (V, E)$, where $V$ is a set with n elements called vertices (or nodes) and E is a set of ordered pair of vertices called edges. In other words, $E \subseteq V \times V$ . We call $V$ and $E$ the vertex set and edge set, respectively. When convenient, we let $V(G)$ and $E(G)$ denote the vertices and edges of $G$, respectively. For $u$, $v \in V$, the ordered pair $(u, v)$ denotes an edge from $u$ to $v$.

An undirected graph in short, **graph**, consists of a vertex set $V$ and of a set $E$ of unordered pairs of vertices. For $u, v \in V$ and $u \neq v$, the set $\{u, v\}$ denotes an unordered edge. A digraph is undirected if $(v, u) \in E$ anytime $(u, v) \in E$. It is possible and convenient to identify an undirected digraph with the corresponding graph; vice versa, the directed version of a graph $(V, E)$ is the digraph $(V', E')$ with the property that $(u, v) \in E'$ if and only if $u, v \in E$. In what follows, our convention is to allow self-loops in both graphs and digraphs.

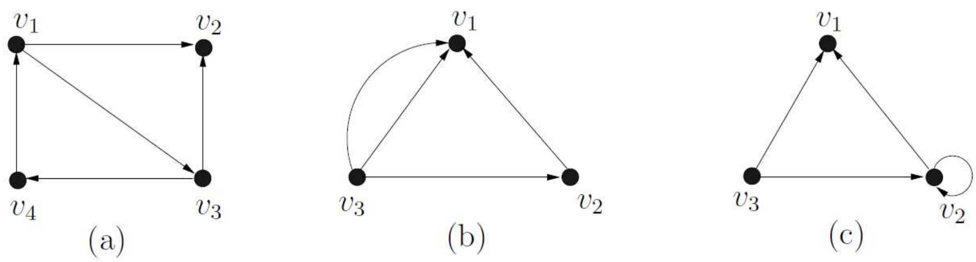



Figure 2: Digraphs.

A digraph $(V', E')$ is a **subgraph** of a digraph $(V, E)$ if $V' \subset V$ and $E' \subset E$; additionally, a digraph $(V', E')$ is a **spanning subgraph** if it is a subgraph and $V' = V$ . The subgraph of $(V, E)$ induced by $V' \subset V$ is the digraph $(V', E')$, where $E'$ contains all edges in $E$ between two vertices in $V'$. For two digraphs $G = (V, E)$ and $G' = (V', E')$, the intersection and union of $G$ and $G'$ are defined by

$$G \cap G' = (V \cap V', E \cap E'), \tag{24}$$

$$G \cup G' = (V \cup V', E \cup E'). \tag{25}$$

Analogous definitions may be given for graphs. In a digraph $G$ with an edge $(u, v) \in E$, $u$ is called an **in-neighbor** of $v$, and $v$ is called an **out-neighbor** of $u$. We let $N_G^{in}(v)$ (resp., $N_G^{out}(v)$) denote the set of in-neighbors, (resp. the set of out-neighbors) of $v$ in the digraph $G$. We will drop the subscript $G$ when the graph $G$ is clear from the context. The **in-degree** and **out-degree** of $v$ are the cardinality of $N^{in}(v)$ and $N^{out}(v)$, respectively. A digraph is **topologically balanced** if each vertex has the same in-and out-degrees (even if distinct vertices have distinct degrees). Likewise, in an undirected graph $G$, the vertices $u$ and $v$ are neighbors if $\{u, v\}$ is an undirected edge. We let $N_G(v)$ denote the set of neighbors of $v$ in the undirected graph $G$. As in the directed case, we will drop the subscript $G$ when the graph $G$ is clear from the context.

The degree of $v$ is the cardinality of $N(v)$.

**Remark 2** *(Additional notions [15]). For a digraph $G = (V, E)$, the* ***reverse digraph*** *$rev(G)$ has vertex set $V$ and edge set $rev(E)$ composed of all edges in $E$ with reversed direction. A digraph $G = (V, E)$ is* ***complete*** *if $E = V \times V$ . A* ***clique*** *(V',E') of a digraph (V,E) is a subgraph of (V,E) which is complete, that is, such that $E' = V' \times V'$. Note that a clique is fully determined by its set of vertices, and hence there is no loss of precision in denoting it by V '.* ***A maximal clique*** *V' of an edge of a digraph is a clique of the digraph with the following two properties: it contains the edge, and any other subgraph of the digraph that strictly contains $(V', V' \times V')$ is not a clique.*

#### 1.3.1 Connectivity notions

Let us now review some basic **connectivity** notions for digraphs and graphs. We begin with the setting of undirected graphs because of its simplicity. A path in a graph is an ordered sequence of vertices such that any pair of consecutive vertices in the sequence is an edge of the graph. A graph is connected if there exists a path between any two vertices. If a graph is not connected, then it is composed of multiple connected components, that is, multiple connected subgraphs. A **path** is **simple** if no vertices appear more than once in it, except possibly for initial and final vertex. A **cycle** is a simple path that starts and ends at the same vertex. A graph is **acyclic** if it contains no cycles. A connected acyclic graph is a tree. A **forest** is a graph that can be written as the disjoint union of trees. Trees have interesting properties: for example, $G = (V, E)$ is a tree if and only if $G$ is connected and $|E| = |V| - 1$. Alternatively, $G = (V, E)$ is a tree if and only if $G$ is acyclic and $|E| = |V| - 1$. Figure (3) illustrates these notions.

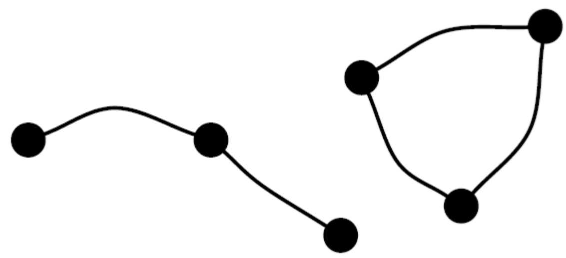
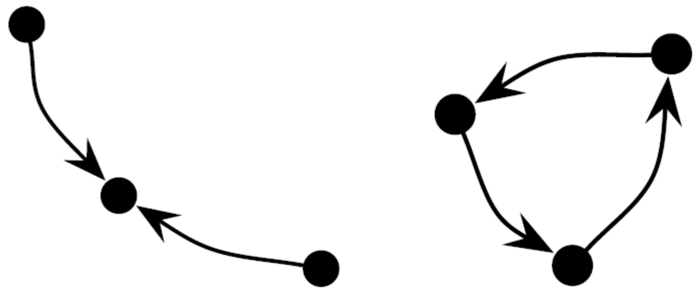


Figure 3: An illustration of connectivity notions on graphs. The graph has two connected components. The leftmost connected component is a tree, while the rightmost connected component is a cycle.

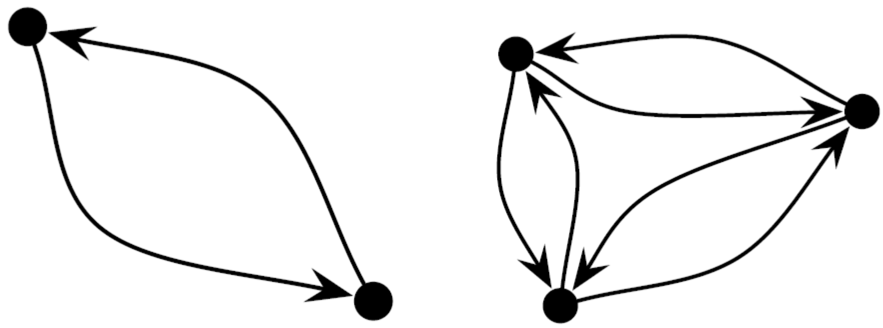
Next, we generalize these notions to the case of digraphs. A directed path in a digraph is an ordered sequence of vertices such that any ordered pair of vertices appearing consecutively in the sequence is an edge of the digraph. A cycle in a digraph is a directed path that starts and ends at the same vertex and that contains no repeated vertex as except for the initial and the final vertex. A digraph is acyclic if it contains no cycles. In an acyclic graph, every vertex of in-degree 0 is named a source, and every vertex of out-degree 0 is named a sink. Every acyclic digraph has at least one source and at least one sink. Figure (4) illustrates these notions.



(a) (b)

Figure 4: Illustrations of connectivity notions on a digraph: (a) shows an acyclic digraph with one sink and two sources; (b) shows a directed path which is also a cycle.

The set of cycles of a directed graph is finite. A directed graph is **aperiodic** if there exists no $k > 1$ that divides the length of every cycle of the graph. In other words, a digraph is aperiodic if the greatest common divisor of the lengths of its cycles is one. A digraph is **periodic** if it is not aperiodic. Figure (5) shows examples of a periodic and an aperiodic digraph.



(a) (b)

Figure 5: (a) A periodic digraph. (b) An aperiodic digraph with cycles of length 2 and 3.

A vertex of a digraph is **globally reachable** if it can be reached from any other vertex by traversing a directed path. A digraph is **strongly connected** if every vertex is globally reachable. A **directed tree** (sometimes called a rooted tree) is an acyclic digraph with the following property: there exists a vertex, called the root, such that any other vertex of the digraph can be reached by one and only one directed path starting at the root. In a directed tree, every in-neighbor of a vertex is called a **parent** and every out-neighbor is called a child. Two vertices with the same parent are called **siblings**. A successor of a vertex $u$ is any other node that can be reached with a directed path starting at $u$. A predecessor of a vertex $v$ is any other node such that a directed path exists starting at it and reaching $v$. A directed spanning tree, or simply a spanning tree, of a digraph is a spanning subgraph that is a directed tree. Clearly, a digraph contains a spanning tree if and only if the reverse digraph contains a globally reachable vertex. A (directed) **chain** is a directed tree with exactly one source and one sink. A (directed) **ring digraph** is the cycle obtained by adding to the edge set of a chain a new edge from its sink to its source. Figure (6) illustrates some of these notions.

**Lemma 5** *(Connectivity in topologically balanced digraphs [15]). Let $G$ be a digraph. The following statements hold:*

*(i) if $G$ is strongly connected, then it contains a globally reachable vertex and a spanning tree; and*

*(ii) if $G$ is topologically balanced and contains either a globally reachable vertex or a spanning tree, then $G$ is strongly connected and is Eulerian (that is a graph with a cycle that visits all the graph edges exactly once). Given a digraph $G = (V,E)$, an in-neighbor of a nonempty set of nodes $U$ is a node $v \in V \backslash U$ for which there exists an edge $(v, u) \in E$ for some $u \in U$.*

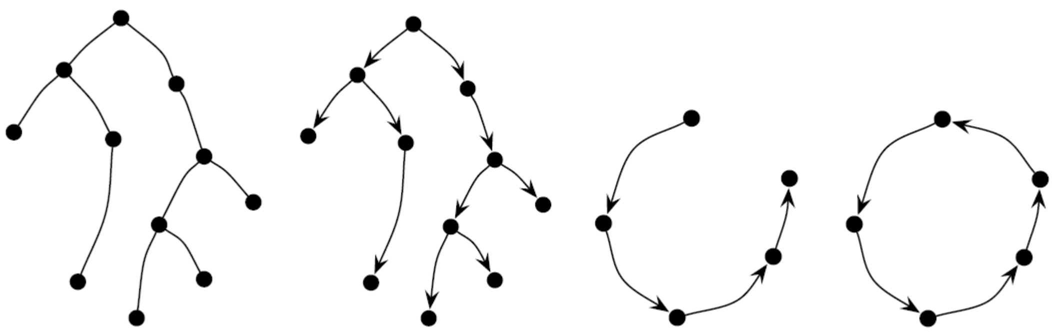


Figure 6: From left to right, tree, directed tree, chain, and ring digraphs.

**Lemma 6** *(Disjoint subsets and spanning trees [15]). Given a digraph $G$ with at least two nodes, the following two properties are equivalent: (i) $G$ has a spanning tree; and (ii) for any pair of nonempty disjoint subsets $U_1, U_2 \subset V$, either $U_1$ has an in-neighbor or $U_2$ has an in-neighbor.*

The result is illustrated in Figure (7). We can also state the result in terms of global reachability: G has a globally reachable node if and only if, for any pair of nonempty disjoint subsets $U_1, U_2 \subset V$ , either $U_1$ has an out-neighbor or $U_2$ has an out-neighbor. The definition of the out-neighbor of a set can be trivially made analogously.

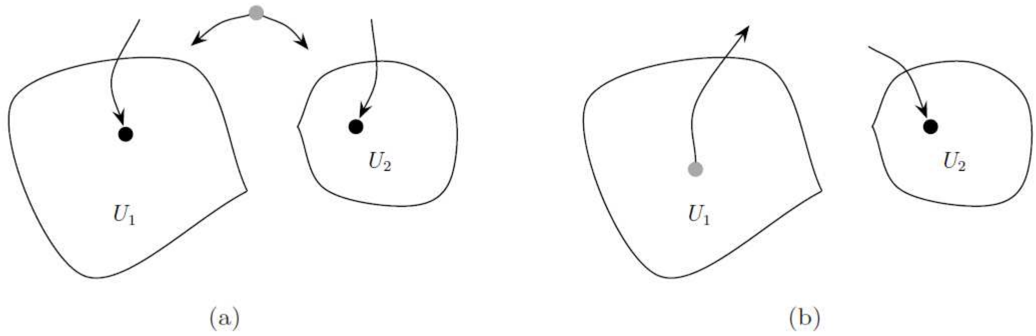



Figure 7: An illustration of Last Lemma. The root of the spanning tree is plotted in gray. In (a), the root is outside the sets $U_1$ and $U_2$. Because these sets are non-empty, there exists a directed path from the root to a vertex in each one of these sets. Therefore, both $U_1$ and $U_2$ have in-neighbors. In (b), the root is contained in $U_1$. Because $U_2$ is non-empty, there exists a directed path from the root to a vertex in $U_2$, and, therefore, $U_2$ has in-neighbors. The case when the root belongs to $U_2$ is treated analogously.

### 1.3.2 Weighted digraphs

A weighted digraph is a triplet $G = (V, E, A)$, where the pair $(V, E)$ is a digraph with nodes $V = v_1, ..., v_n$, and where the nonnegative matrix $A \in R^{n\times n}_{>0}$ is a weighted **adjacency matrix** with the following property: for $i, j \in 1, ..., n$, the entry $a_{ij} > 0$ if $(v_i, v_j)$ is an edge of $G$, and $aij = 0$ otherwise. In other words, the scalars $a_{ij}$ , for all $(v_i, v_j) \in E$, are a set of weights for the edges of $G$. Note that the edge set is uniquely determined by the weighted adjacency matrix and it can therefore be omitted. When convenient, we denote the adjacency matrix of a weighted digraph $G$ by $A(G)$. Figure (8) shows an example of a weighted digraph.

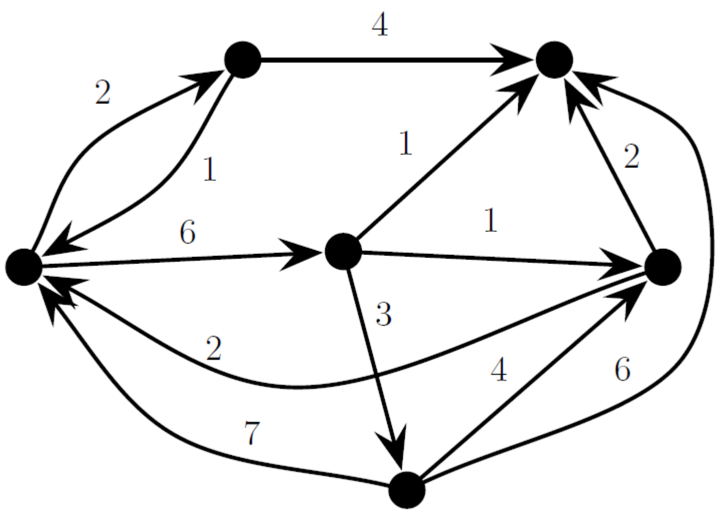



Figure 8: A weighted digraph with natural weights.

A digraph $G = (V, E)$ can be naturally thought of as a weighted digraph by defining the weighted adjacency matrix $A \in \{0,1\}^{n \times n}$ as

$$a_{ij} = \begin{cases} 1 & if (v_i, v_j) \in E \\ 0 & otherwise \end{cases} \tag{26}$$

where $V = \{v_1, ..., v_n\}$. The adjacency matrix of a graph is the adjacency matrix of the directed version of the graph. Reciprocally, given a weighted digraph $G = (V, E, A)$, we refer to the digraph $(V, E)$ as the unweighted version of $G$ and to its associated adjacency matrix as the unweighted adjacency matrix. A weighted digraph is undirected if $a_{ij} = a_{ji}$ for all $i, j \in \{1, ..., n\}$. Clearly, $G$ is undirected if and only if $A(G)$ is symmetric.

Numerous concepts introduced for digraphs remain equally valid for the case of weighted digraphs, including the connectivity notions and the definitions of in- and out-neighbors. Finally, we generalize the notions of in- and out-degree to weighted digraphs. In a weighted digraph $G = (V, E, A)$ with $V = \{v_1, ..., v_n\}$, the weighted out-degree and the weighted in-degree of vertex $v_i$ are defined by, respectively,

$$d_{out}(v_i) = \sum_{j=1}^{n} a_{ij}, \tag{27}$$

and

$$d_{in}(v_i) = \sum_{j=1}^{n} a_{ji}, \tag{28}$$

The weighted digraph $G$ is weight-balanced if $d_{out}(vi) = d_{in}(vi)$ for all $v_i \in V$. The weighted out-degree matrix $D_{out}(G)$ and the weighted in-degree matrix $D_{in}(G)$ are the diagonal matrices defined by

$$D_{out}(G) = diag(A\mathbf{1}_n), \tag{29}$$

and

$$D_{in}(G) = diag(A^T\mathbf{1}_n), \tag{30}$$

that is, $(D_{out}(G))_{ii} = d_{out}(v_i)$ and $(D_{in}(G))_{ii} = d_{in}(v_i)$, respectively.

**Example 4** *The graph of fig.2(a), in which a weight equal to 1 is considered for each edge, has the following adjacency, out-degree and in-degree matrices:*

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}; \quad (31) \\ D_{out} = diag(A\boldsymbol{1}_n) &= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad (32) \\ D_{in} = diag(A^T\boldsymbol{1}_n) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad (33) \end{aligned}$$

### 1.3.3 Distances on digraphs and weighted digraphs

We first present a few definitions for unweighted digraphs. Given a digraph $G$, the (topological) length of a directed path is the number of the edges composing it. Given two vertices $u$ and $v$ in the digraph $G$, the distance from $u$ to $v$, denoted $dist_G(u,v)$, is the smallest length of any directed path from $u$ to $v$, or $+\infty$ if there is no directed path from $u$ to $v$. That is,

$$dist_G(u,v) = min(\{length(p)|p\ is\ a\ directed\ path\ from\ u\ to\ v\} \cup \{+\infty\}). \quad (34)$$

Given a vertex $v$ of a digraph $G$, the **radius** of $v$ in $G$ is the maximum of all the distances from $v$ to any other vertex in $G$. That is,

$$radius(v,G) = maxdist_G(v,u)|u \in V(G). \quad (35)$$

If $T$ is a directed tree and $v$ is its root, then the depth of $T$ is radius $(v,T)$. Finally, the diameter of the digraph $G$ is

$$diam(G) = max\{dist_G(u,v) \ \mid \ u,v \in V(G)\}. \quad (36)$$

These definitions lead to the following simple results:

(i) radius$(v,G) \leq diam(G)$ for all vertices $v$ of $G$;

(ii) $G$ contains a spanning tree rooted at $v$ if and only if radius $(v, G) < +\infty$; and

(iii) G is strongly connected if and only if $diam(G) < +\infty$.

The definitions of path length, distance between vertices, radius of a vertex, and diameter of a digraph can be easily extended to undirected graphs, too. Next, we consider weighted digraphs. Given two vertices $u$ and $v$ in the weighted digraph $G$, the weighted distance from $u$ to $v$, denoted $wdist_G(u, v)$, is the smallest weight of any directed path from $u$ to $v$, or $+\infty$ if there is no directed path from $u$ to $v$. That is,

$$wdist_G(u, v) = min(\{weight(p) \ \mid \ p \, is \, a \, directed \, path \, from \, u \, to \, v\} \cup \{+\infty\}). \tag{37}$$

Here, the weight of a subgraph of a weighted digraph is the sum of the weights of all the edges of the subgraph. Note that when a digraph is thought of as a weighted digraph (with the unweighted adjacency matrix (1.3.2)), the notions of weight and weighted distance correspond to the usual notions of length and distance, respectively. We leave it the reader to provide the definitions of weighted radius, weighted depth, and weighted diameter.

### 1.3.4 Algebraic graph theory

Algebraic graph theory investigates the properties of matrices defined by digraphs. In this section, we expose two topics. First, we study the equivalence between properties of graphs and of their associated adjacency matrices. We also specify how to associate a digraph to a nonnegative matrix. Second, we introduce and characterize the **Laplacian matrix** of a weighted digraph. We begin by studying adjacency matrices. Note that the adjacency matrix of a weighted digraph is nonnegative and, in general, not stochastic. The following lemma expands on this point.

**Lemma 7** *(Weight-balanced digraphs and doubly stochastic adjacency matrices [15]). Let $G$ be a weighted digraph of order $n$ with weighted adjacency matrix $A$ and weighted out-degree matrix $D_{out}$. Define the matrix*

$$F = \{ \begin{array}{cc} D_{out}^{-1}A, & if \; each \; out-degree \; is \; strictly \; positive, \\ (I_n + D_{out})^{-1}(I_n + A), & otherwise \end{array} \tag{38}$$

*Then*

*(i) $F$ is row-stochastic; and*

*(ii) $F$ is doubly stochastic if $G$ is weight-balanced and the weighted degree is constant for all*

*vertices.*

**Proof 1** *Consider first the case when each vertex has an outgoing edge so that $D_{out}$ is invertible. We first note that $diag(v)^{-1}v = \boldsymbol{1}_n$, for each $v \in (\Re \setminus \{0\})^n$. Therefore*

$$(D_{out}^{-1}A)\boldsymbol{1}_n = diag(A\boldsymbol{1}_n)^{-1}(A\boldsymbol{1}_n) = \boldsymbol{1}_n \tag{39}$$

*which proves (i). Furthermore, if $D_{out} = D_{in} = dI_n$ for some $d \in \Re_{>0}$, then*

$$(D_{out}^{-1}A)^T\boldsymbol{1}_n = \frac{1}{d}(A^T\boldsymbol{1}_n) = D_{in}^{-1}(A^T\boldsymbol{1}_n) = diag(A^T\boldsymbol{1}_n)^{-1}(A^T\boldsymbol{1}_n) = \boldsymbol{1}_n \tag{40}$$

*which proves (ii). Finally, if $(V, E, A)$ does not have outgoing edges at each vertex, then apply the statement to the weighted digraph $(V, E \cup \{(i,i)|i \in \{1, ..., n\}\}, A + I_n)$.*

The next result characterizes the relationship between the adjacency matrix and directed paths in the digraph.

**Lemma 8** *(Directed paths and powers of the adjacency matrix [15]).*

*Let $G$ be a weighted digraph of order $n$ with weighted adjacency matrix $A$, with unweighted adjacency matrix $A_{0,1} \in \{0,1\}^{n\times n}$, and possibly with selfloops. For all $i, j, k \in 1, ..., n$*

*(i) the $(i, j)$ entry of $A_{0,1}^k$ equals the number of directed paths of length $k$ (including paths with self-loops) from node $i$ to node $j$; and*

*(ii) the $(i, j)$ entry of $A^k$ is positive if and only if there exists a directed path of length $k$ (including paths with self-loops) from node $i$ to node $j$.*

**Proof 2** . *The second statement is a direct consequence of the first. The first statement is proved by induction. The statement is clearly true for $k = 1$. Next, we assume the statement is true for some $k \geq 1$ and we prove it for $k + 1$. By assumption, the entry $(A^k)_{ij}$ equals the number of directed paths from $i$ to $j$ of length $k$. Note that each path from $i$ to $j$ of length $k + 1$ identifies (1) a unique node $\ell$ such that $(i, \ell)$ is an edge of $G$ and (2) a unique path from $\ell$ to $j$ of length $k$. We write $A^{k+1} = AA^k$ in components as*

$$A_{ij}^{k+1} = \sum_{\ell=1}^{n} A_{i\ell}(A^k)_{\ell j}. \tag{41}$$

*Therefore, it is true that the entry $A_{ij}^{k+1}$ equals the number of directed paths from $i$ to $j$ of length $k + 1$. This concludes the induction argument. Lemma 8 is proven.*

The following proposition characterizes in detail the relationship between various connectivity properties of the digraph and algebraic properties of the adjacency matrix.

**Proposition 1** *(Connectivity properties of the digraph and positive powers of the adjacency matrix [15]).*

*Let $G$ be a weighted digraph of order $n$ with weighted adjacency matrix $A$. The following statements are equivalent:*

*(i) $G$ is strongly connected;*

*(ii) $A$ is irreducible; and*

*(iii) $\sum_{k=0}^{n} A^k$ is positive.*

*For any $j \in 1, ..., n$, the following two statements are equivalent:*

*(iv) the $j$th node of $G$ is globally reachable; and*

*(v) the $j$th column of $\sum_{k=0}^{n-1} A^k$ has positive entries.*

Stronger statements can be given for digraphs with self-loops.

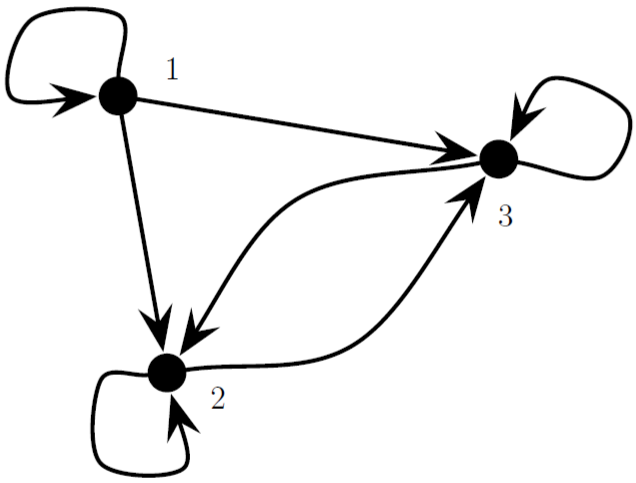



Figure 9: An illustration of Proposition 1. Even though vertices 2 and 3 are globally reachable, the digraph is not strongly connected because vertex 1 has no in-neighbor other than itself. Therefore, the associated adjacency matrix $A = (a_{ij})$ with $(a_{1j}) = \mathbf{1}_3, (a_{2j}) = (a_{3j}) = (0, 1, 1)$, is reducible.

**Proposition 2** *(Connectivity properties of the digraph and positive powers of the adjacency matrix [15]). Let $G$ be a weighted digraph of order $n$ with weighted adjacency matrix $A$ and with self-loops at each node. The following statements are equivalent:*

*(iv) $G$ is strongly connected; and*

*(v) $A^{n-1}$ is positive. For any $j \in 1, ..., n$, the following two statements are equivalent:*

*(iv) the $j$th node of $G$ is globally reachable; and*

*(v) the $j$th column of $A^{n-1}$ has positive entries.*

Next, we characterize the relationship between irreducible aperiodic digraphs and primitive matrices.

**Proposition 3** *(Strongly connected and aperiodic digraph and primitive adjacency matrix [15]). Let $G$ be a weighted digraph of order $n$ with weighted adjacency matrix $A$. The following two statements are equivalent:*

*(i) $G$ is strongly connected and aperiodic; and*

*(ii) $A$ is primitive, that is, there exists $k \in \mathbb{N}$ such that $A^k$ is positive.*

This concludes our study of adjacency matrices associated to weighted digraphs. Next, we emphasize how all results obtained so far have analogs that hold when the original object is a nonnegative matrix, instead of a weighted digraph.

**Remark 3** *(From a nonnegative matrix to its associated digraphs [15]). Given a nonnegative $n \times n$ matrix $A$, its associated weighted digraph is the weighted digraph with nodes $\{1, ..., n\}$, and weighted adjacency matrix $A$. The unweighted version of this weighted digraph is called the associated digraph. The following statements are analogs of the previous lemmas:*

*(i) if $A$ is stochastic, then its associated digraph has weighted out-degree matrix equal to $I_n$;*

*(ii) if $A$ is doubly stochastic, then its associated weighted digraph is weight-balanced and, additionally, both in-degree and out-degree matrices are equal to $I_n$; and*

*(iii) $A$ is irreducible if and only if its associated weighted digraph is strongly connected.*

So far, we have analyzed in detail the properties of adjacency matrices. We conclude this section by studying a second relevant matrix associated to a digraph, called the Laplacian matrix. The Laplacian matrix of the weighted digraph $G$ is

$$L(G) = Dout(G) - A(G). \tag{42}$$

Some immediate consequences of this definition are the following:

(i) $L(G)\mathbf{1}_n = \mathbf{0}_n$, that is, 0 is an eigenvalue of $L(G)$ with eigenvector $\mathbf{1}_n$;

(ii) $G$ is undirected if and only if $L(G)$ is symmetric; and

(iii) $L(G)$ equals the Laplacian matrix of the digraph obtained by adding to or removing from $G$ any self-loop with arbitrary weight.

Further properties are established as follows.

**Theorem 6** *(Properties of the Laplacian matrix [15]). Let $G$ be a weighted digraph of order $n$. The following statements hold:*

*(i) all eigenvalues of $L(G)$ have nonnegative real part (thus, if $G$ is undirected, then $L(G)$ is symmetric positive semidefinite);*

*(ii) if $G$ is strongly connected, then $rank(L(G)) = n - 1$, that is, 0 is a simple eigenvalue of $L(G)$;*

*(iii) $G$ contains a globally reachable vertex if and only if $rank(L(G)) = n - 1$;*

*(iv) the following three statements are equivalent:*

*(a) $G$ is weight-balanced;*

*(b) $\mathbf{1}_n^T L(G) = \mathbf{0}_n^T$; and*

*(c) $L(G) + L(G)^T$ is positive semi-definite.*

**Example 5** *Let us consider the graph in fig.(10). It is a weight-balanced graph, in fact any node the weighted in-degree is equal to the weighted out-degree, 3 for node A, 2 for node B, 1 for node C.*

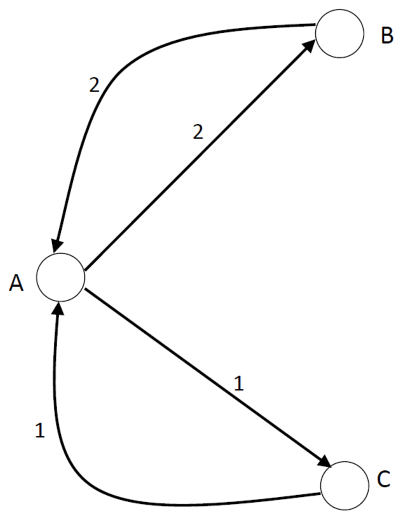



Figure 10: A weight-balanced graph

*The Adjacency matrix is the following:*

$$A = \begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}; \tag{43}$$

*The out-degree and in-degree matrix:*

$$D = D_{out} = D_{in} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \tag{44}$$

*The Laplacian matrix:*

$$L(G) = D - A = \begin{bmatrix} 3 & -2 & -1 \\ -2 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}; \tag{45}$$

*Let us verify defined properties:*

$$\boldsymbol{1}_n^T L(G) = \boldsymbol{1}_n^T \begin{bmatrix} 3 & -2 & -1 \\ -2 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \boldsymbol{0}_n^T; \tag{46}$$

*Property (b) is verified.*

$$eig(L(G) + L(G)^T) = eig(\begin{bmatrix} 6 & -4 & -2 \\ -4 & 4 & 0 \\ -2 & 0 & 2 \end{bmatrix}) = (0, 2.54, 9.46); \tag{47}$$

*Property (c) is verified too.*

## 1.4 Conclusions

In this chapter we have provided a mathematical foundation, based on the theory of graphs and nonnegative matrices.

We have not discussed the applicability of the concepts in the real world. This is postponed for the later chapters where we apply the results developed here to various aspects of system analysis. In this chapter we have given some basic notions from graph theory and develop important results about connectivity of digraphs. We have also explored the theory of nonnegative matrices with emphasis on the deeper connections between nonnegative matrices and directed graphs.

## References


[1] J. Bang-Jensen and G. Gutin, Digraphs: Theory, Algorithms and Applications. Springer-Verlag, 2002.

[2] L. R. Foulds, Graph Theory Applications. Springer-Verlag, 1992.

[3] A. Berman and R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences. Classics in Appl. Math. 9, SIAM, 1994.

[4] Z. Lin, Coupled Dynamic Systems: From Structure Towards Stability and Stabilizability, Ph.D. dissertation, University of Toronto, Canada, 2006.

[5] Horn, R. A. and Johnson, C. R., Matrix Analysis, Cambridge University Press, 1985.

[6] Meyer, C. D., Matrix Analysis and Applied Linear Algebra SIAM, 2001.

[7] M. Fiedler. Algebraic connectivity of graphs. Czechoslovak Mathematical Journal, 23(98):298-305, 1973.

[8] C. Godsil and G. Royle. Algebraic Graph Theory, volume 207 of Graduate Texts in Mathematics. Springer, 2001.

[9] R. Merris. Laplacian matrices of a graph: a survey. Linear Algebra and its Applications, 197:143-176, 1994.

[10] B. Mohar. The Laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O. Ollermann, and A. Schwenk, editors, Graph Theory, Combinatorics, and Applications, pages 871-898, New York, 1991. John Wiley.

[11] A. Muhammad and M. Egerstedt. Connectivity graphs as models of local interactions. Journal of Applied Math. and Computation, 2004.

[12] J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. Proc. American Mathematical Socity, 15:733-736, 1963.

[13] M. M. Zavlanos and G. J. Pappas. Controlling connectivity of dynamic graphs. 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference (CDC-ECC 05), pages 6388-6393, Dec. 2005. 100.

[14] D.D. Šiljak,'“Large Scale Dynamic Systems: Stability and Structure.”. Amsterdam, The Netherlands: North Holland, 1978.

[15] Distributed Control of Robotic Networks by F. Bullo, J. Cortes and S. Martnez. Manuscript under contract. Version: March 22, 2009

## 2 Observability and Strong Observability in LTI Systems

### 2.1 Introduction

We'll limit our analysis to linear and time invariant (LTI) systems. In particular, since observability depends to the pair of matrices $(A, C)$, in this section we'll limit to consider autonomous systems, that is systems in which there is no external inputs.

**Definition 4** *The autonomous LTI system*

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{cases} \tag{48}$$

*is said to be observable if and only if, for all initial states $x_0 = x(0)$, the value $x_0$ can be recovered on the basis of the observation of the output evolution $y(t)$ over a finite interval $t \in [0, t_f], t_f < \infty$.*

Let introduce a simple physical example that allows us to illustrate this concept.

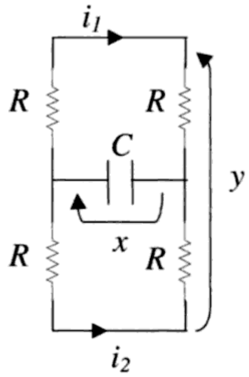



Figure 11: An unobservable network.

**Example 6** *Let us consider the network of figure (11), where the state variable is the voltage of the capacitor, that is $x(t) = v_C(t)$. Because of the symmetry of the network, it is easy to prove that, for all initial values $x(0)$ of the capacitor's voltage, the output voltage $y$ is zero. In fact, for all $t \geq 0$, we can write $y(t) = Ri_1(t) - Ri_2(t) = 0$, since $i_1(t) = i_2(t)$, for all initial value of the capacitor's voltage. The measure of the output $y(t)$ for a certain time interval does not allow us to discover the initial state of the system. This implies that the system is unobservable.*

### 2.2 Verification of Observability

For verification of the observability of a system, we'll provide two different criteria of analysis, both based on the calculation of appropriate matrix. The first criterion is based on the verification of the full rank of a matrix called observability grama, the second criterion is based on the calculation of the rank of observability matrix. The second criterion is more immediate but the first is more important because it provides a procedure to reconstruct the initial state of the system, by knowning the output for a finite time interval.

**Definition 5** *Let us consider system (48), where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^p$, we define the observability grama the $n \times n$ ti matrix*

$$O(t) = \int_0^t e^{A^T\tau} C^T C e^{A\tau} d\tau \tag{49}$$

**Theorem 7** *System (48) is observable if and only if the observability grama is nonsingular for all $t > 0$ [28].*

**Theorem 8** *Let us consider system (48), where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^p$, we define observability matrix the $(p \cdot n) \times n$ matrix*

$$\Lambda = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{50}$$

*The system (48) is observable if and only if*

$$rank(\Lambda) = n \tag{51}$$

*[28]*

.

## 2.3 Luenberger Observer

Let us consider the system

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{cases} \tag{52}$$

with $x \in \mathbb{R}^n$, $u \in \mathbb{R}^r$ and $y \in \mathbb{R}^p$. The LTI system:

$$\begin{cases} \dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + K_0(y(t) - \hat{y}(t)) \\ \hat{y}(t) &= C\hat{x}(t) \end{cases} \tag{53}$$

with $\hat{x} \in \mathbb{R}^n$, $\hat{y} \in \mathbb{R}^p$ where $K_0 \in \mathbb{R}^{n \times p}$ is any matrix for which matrix $A - K_0C$ is Hurwitz, is called a Luenberger Observer with respect to the system (52).

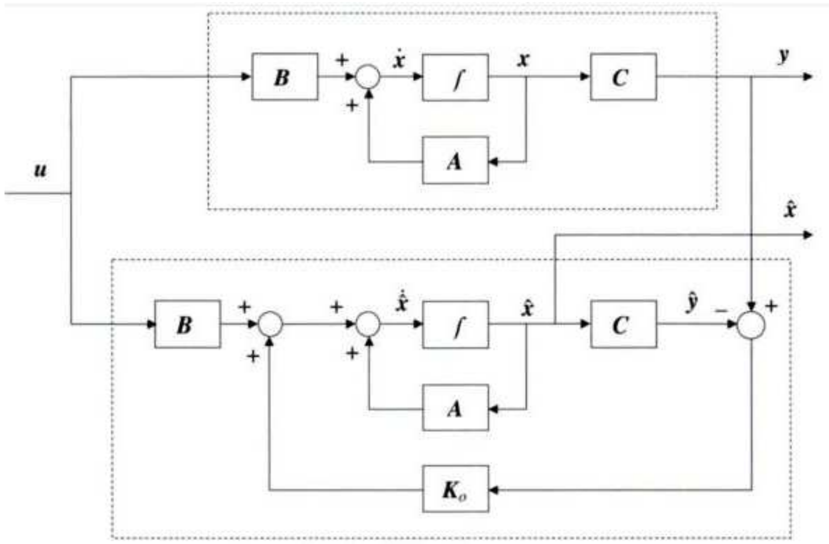



Figure 12: Luenberger Observer.

**Theorem 9** *Let us consider system (52). The Luenberger Observer is an asymptotic state observer of that system.*

**Proof 3** *Let us denote with*

$$e(t) = x(t) - \hat{x}(t) \tag{54}$$

*the estimation error which measures the difference between the state $x(t)$ and the estimated state $\hat{x}(t)$.*

*Subtracting (53) from (52) we get*

$$\begin{aligned} \dot{e}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) \\ &= Ax(t) + Bu(t) - \hat{A}\hat{x}(t) - \hat{B}\hat{u}(t) - K_0(y(t) - \hat{y}(t)) \\ &= Ae(t) + Bu(t) + K_0 C\hat{x}(t) - Bu(t) - K_0 y(t) \\ &= Ae(t) - K_0 C(x(t) - \hat{x}(t)) \\ &= (A - K_0 C)e(t) \end{aligned} \tag{55}$$

*which means that the error dinamics are governed by the autonomous system*

$$\dot{e}(t) = (A - K_0 C)e(t); \qquad e(0) = x(0) - \hat{x}(0) \tag{56}$$

*from which follows the validity of the statement, being $A - K_0 C$ Hurwitz which implies*

$$\lim_{t\to\infty} \|e(t)\| = \lim_{t\to\infty} \|x(t) - \hat{x}(t)\| = 0 \tag{57}$$

*for all input functions $u(t)$ and initial states $x(0)$ and $\hat{x}(0)$.*

The Luenberger Observer is a LTI system with the same order $n$ of the system of whose state we want estimate the (this is the reason because it is also called "of full order"); its input is provided by the system's input $u(t)$ and output $y(t)$, and its output transformation is the same as that of the observed system. The scheme of Luenberger Observer is given in fig(12). Of course not all systems with that structure are asymptotic estimators. Condition (57) or another statement about eigenvalues of $A - K_0C$ must be satisfied.

## 2.4 Steady-State Kalman Filter

Let us consider system (52) and let us suppose that the pair $(A, C)$ is observable. In the previous section we have seen that the state $x$ can be asymptotically estimated by the observer (53), where $K_0$ is designed such that

$$Re[\lambda(A - K_0C)] < 0 \tag{58}$$

$K_0$ can also be designed in an optimal way with the following method.

Choose a $p \times p$ symmetric and positive-definite matrix $V$ and an $n \times l$ matrix $\Gamma$ such that $V = V^T > 0$and $(A^T, \Gamma^T)$ is observable. Then

$$K_0 = P_eCV^{-1}, \tag{59}$$

where $P_e = P_e^T \geq 0$ is the solution of the so called Algebraic Riccati Equation (ARE):

$$P_eA^T + AP_e + \Gamma\Gamma^T - P_eC^TV^{-1}CP_e = 0. \tag{60}$$

Is there a special meaning of this design of $K_0$?

Consider the system

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) + \Gamma w \\ y(t) &= Cx(t) + v \end{cases} \tag{61}$$

where $w(t) \in \mathbb{R}^l$ and $v(t) \in \mathbb{R}^p$ are uncorrelated white zero-mean Gaussian stochastic processes

$$
\begin{aligned}
E\{w(t)\} &= 0; \qquad E\{w(t)w(t+\tau)^T\} = I_l\delta(t-\tau), && (62)\\
E\{v(t)\} &= 0; \qquad E\{v(t)v(t+\tau)^T\} = V\delta(t-\tau), && (63)\\
E\{v(t)w(t)^T\} &= 0, \qquad \forall t, \tau > 0. && (64)
\end{aligned}
$$

In this case, $\hat{x}(t)$ provided by the observer (53), where $K_0$ is obtained in (59), is the optimal estimate that minimize

$$
\lim_{t\to\infty} E\{e(t)e(t)^T\} = 0, \tag{65}
$$

where $e$ is the estimate error defined in the second of (56), and the minimum value is given by the trace of matrix $P_e$.

The observer (53) with $K_0$ defined in (59) is called Steady-State Kalman Filter.

### 2.5 Sliding Mode Observers (SMO) for Linear Systems

The concept of sliding mode (SM) control has been applied to the problem of state estimation for linear systems, possibly uncertain and some classes of nonlinear systems as well. The observer trajectories are constrained to evolve, after a finite time, along a suitable sliding manifold by means of an injection signal designed according to a SM control algorithm. The sliding manifold is usually given by the difference between the observer and the system outputs, therefore in such cases we refer to the control signal as *output injection signal*.

Let us consider the linear system described by

$$
\begin{cases} \dot{x} = Ax(t) + Bu(t) \\ y = Cx(t) \end{cases} \tag{66}
$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$ and $p \geq m$. Assume that the matrices $B$ and $C$ are of full rank and the pair $(A, C)$ is observable. It is convenient to introduce a coordinate transformation so that the output variables appear as the last $p$ components of the states. One possibility is to consider the non-singular transformation $x \to T_c x$ as

$$
T_c = \begin{bmatrix} {N_c}^T \\ C \end{bmatrix} \tag{67}
$$

where the columns of $N_c \in \mathbb{R}^{n\times(n-p)}$ span the null space of $C$. This transformation is nonsingular, and defining

$$T_c x = \begin{bmatrix} x_1 \\ y \end{bmatrix} \begin{array}{l} \updownarrow n-p \\ \updownarrow p \end{array} \tag{68}$$

the transformed system's dynamics are

$$\begin{cases} \dot{x}_1(t) = A_{11}x_1(t) + A_{12}y(t) + B_1u(t) \\ \dot{y}(t) = A_{21}x_1(t) + A_{22}y(t) + B_2u(t) \end{cases} \tag{69}$$

where

$$\tilde{A} = T_c A T_c^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}; \;\; \tilde{B} = T_c B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}; \;\; \tilde{C} = C T_c^{-1} = [0 \; I_p] \tag{70}$$

The observer proposed by Utkin [29]-[30] has the form

$$\begin{cases} \dot{\hat{x}}_1(t) = A_{11}\hat{x}_1(t) + A_{12}\hat{y}(t) + B_1u(t) + L\nu \\ \dot{\hat{y}}(t) = A_{21}\hat{x}_1(t) + A_{22}\hat{y}(t) + B_2u(t) - \nu \end{cases} \tag{71}$$

where $(\hat{x}_1, \hat{y})$ represent the state estimates for $x_1$ and $y$, $L \in \mathbb{R}^{(n-p)\times p}$ is a constant feedback gain matrix and the discontinuous vector $\nu$, of appropriate dimension, is defined componentwise by

$$\nu_i = M sgn(\hat{y}_i - y_i) \tag{72}$$

where $M \in \mathbb{R}_+$. Define $e_1 = \hat{x}_1 - x_1$ and $e_y = \hat{y} - y$. Then from equations (69) and (71) the following error dynamics system are

$$\begin{cases} \dot{e}_1(t) = A_{11}e_1(t) + A_{12}e_y(t) + L\nu \\ \dot{e}_y(t) = A_{21}e_1(t) + A_{22}e_y(t) - \nu \end{cases} \tag{73}$$

that in compact form can be rewritten as follows

$$\dot{e} = \tilde{A}e(t) + \Gamma\nu \quad where \quad \Gamma = \begin{bmatrix} L \\ -I_p \end{bmatrix} \tag{74}$$

Since the pair $(A, C)$ is observable, the pair $(A_{11}, A_{21})$ is also observable. As a consequence, $L$ can be chosen to make the spectrum of $A_{11} + LA_{21}$ lie in $\mathbb{C}_-$. Define a further change of coordinates, dependent on L, by

$$T = \begin{bmatrix} I_{n-p} & L \\ 0 & I_p \end{bmatrix} \tag{75}$$

it results

$$\tilde{e} = \begin{bmatrix} \tilde{e}_1 \\ \tilde{e}_y \end{bmatrix} = Te = \begin{bmatrix} e_1(t) + Le_y(t) \\ e_y(t) \end{bmatrix} \tag{76}$$

and system in compact form (74), with respect to the new coordinates, can be further manipulated as

$$\dot{\tilde{e}} = T\tilde{A}e(t) + T\Gamma\nu \tag{77}$$

From (76), system (77) can be rewritten as

$$\begin{cases} \dot{\tilde{e}}_1(t) = \tilde{A}_{11}\tilde{e}_1(t) + \tilde{A}_{12}e_y(t) \\ \dot{e}_y(t) = A_{21}\tilde{e}_1(t) + \tilde{A}_{22}e_y(t) - \nu \end{cases} \tag{78}$$

where $\tilde{A}_{11} = A_{11} + LA_{21}$, $\tilde{A}_{12} = A_{12} + LA_{22} - \tilde{A}_{11}L$ and $\tilde{A}_{22} = A_{22} - A_{21}L$.
It follows from the 2-nd of (78) that in the domain

$$\Omega = \{(\tilde{e}_1(t), e_y) : \|A_{21}\tilde{e}_1(t)\| + \frac{1}{2}\lambda_{max}(\tilde{A}_{22} + \tilde{A}_{22}^T)\|e_y\| < M - \eta\} \tag{79}$$

where $\eta < M$ is some small positive scalar, the reachability condition

$$e_y^T \dot{e}_y < -\eta\|e_y\| \tag{80}$$

is satisfied. Consequently, an ideal sliding motion will take place on the surface

$$S_o = \{(\tilde{e}_1, e_y) : e_y = 0\} \tag{81}$$

It follows that after some finite time $t_s$, for all subsequent time, $e_y = 0$ and $\dot{e}_y = 0$. Therefore from equation (78) result

$$\dot{\tilde{e}}_1(t) = \tilde{A}_{11}\tilde{e}_1(t) \tag{82}$$

which, by choice of $L$, represents a stable system and so $\tilde{e}_1 \to 0$, i.e., $e_1 \to 0$ and consequently $\hat{x}_1 \to x_1$ asymptotically. Equation (82) represents the reduced order sliding mode error dynamics.

**Example 7** *Consider now the problem of designing a sliding mode observer for system in (66) specialized with*

$$A = \begin{bmatrix} 0 & 1 \\ -2 & 0 \end{bmatrix}; \ B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \ C = \begin{bmatrix} 1 & 1 \end{bmatrix} \tag{83}$$

*which represent a simple harmonic oscillator that is observable since* $(rank([C; CA]) = 2)$*. For simplicity assume* $u = 0$*. Define a nonsingular matrix*

$$T_c = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{84}$$

*and the change of coordinates according to (69)-(75)*

$$\tilde{C} = CT_c^{-1} = [0\ 1]; \ \ \tilde{A} = T_cAT_c^{-1} = \begin{bmatrix} -1 & 1 \\ -3 & 1 \end{bmatrix}; \ \tilde{B} = T_cB = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{85}$$

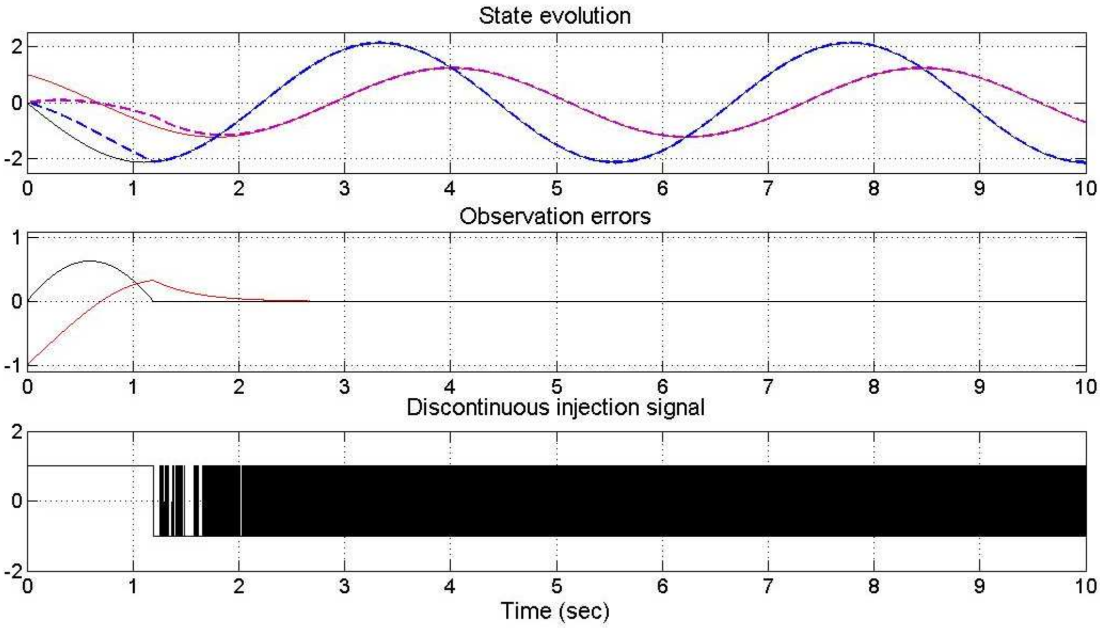



Figure 13: Utkin Observer

*The system (83) is now in the form (69). An appropriate choice of gain in the observer given in (71) is* $L = 0.5$ *which results in an error system governed by* $\tilde{A}_{11}$*. The simulation results which follows (13) were obtained setting the gain of the discontinuous output injection term* $M = 1$ *and the following initial conditions:*
$[x_1(0), y(0)] = [1 \;\; 0]$, $[\hat{x}_1(0), \hat{y}(0)] = [0 \;\; 0]$.

### 2.6 Strong observability and UIO design for linear systems with unknown inputs

Consider the linear time invariant dynamics

$$\begin{array}{rcl} \dot{x} & = & Ax + Gu + F\xi \\ y & = & Cx \end{array} \tag{86}$$

where $x(t) \in \mathbb{R}^n$ and $y(t) \in \mathbb{R}^p$ are the state and output variables, $u(t) \in \mathbb{R}^h$ is a ***known input*** to the system, $\xi(t) \in \mathbb{R}^m$ is an ***unknown input*** term, and $A, G, F, C$ are known constant matrices of appropriate dimension.

Let us make the following assumptions:

**A1**. The matrix triplet $(A, F, C)$ is strongly observable

**A2**. $rank\,(CF) = rank\; F = m$.

The notion of strong observability has been introduced more than thirty years ago [10, 27] in the framework of the unknown-input observers theory. Recently it has been exploited to design robust observers based on the high-order sliding mode approach [25]. It has been shown in [10] that the following property holds

**Theorem 10** *The triplet* $(A, F, C)$ *is strongly observable if and only if it has no invariant zeros [10].*

If conditions A1 and A2 are both satisfied then it can be systematically found a state coordinates transformation together with an output coordinates change which decouple the unknown input $\xi$ from a certain subsystem in the new coordinates. Such a transformation is outlined below.

For the generic matrix $J \in \mathbb{R}^{n_r \times n_c}$ with $rank(J) = r$, we define $J^{\perp} \in \mathbb{R}^{n_r - r \times n_r}$ as a matrix such that $J^{\perp} J = 0$ and $rank(J)^{\perp} = n_r - r$. Matrix $J^{\perp}$ always exists and, furthermore, it is not unique[1]. Let $\Gamma^{+} = [\Gamma^T \Gamma]^{-1} \Gamma^T$ denote the left pseudo-inverse of $\Gamma$ such that $\Gamma^{+}\Gamma = I_{n_c}$, with $I_{n_c}$ being the identity matrix of order $n_c$.

Consider the following transformation matrices $T$ and $U$:

$$T = \begin{bmatrix} F^{\perp} \\ (CF)^{+} C \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}, \quad U = \begin{bmatrix} (CF)^{\perp} \\ (CF)^{+} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}. \tag{87}$$

and the transformed state and output vectors

$$\bar{x} = Tx = \begin{bmatrix} T_1 x \\ T_2 x \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \quad \bar{x}_1 \in R^{n-m} \quad \bar{x}_2 \in R^{m} \tag{88}$$

$$\bar{y} = Uy = \begin{bmatrix} U_1 y \\ U_2 y \end{bmatrix} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix}, \quad \bar{y}_1 \in R^{p-m} \quad \bar{y}_2 \in R^{m} \tag{89}$$

The subcomponents of the transformed vectors take the form

$$\bar{x}_1 = F^{\perp} x, \qquad \bar{x}_2 = (CF)^{+} Cx \tag{90}$$

$$\bar{y}_1 = (CF)^{\perp} y, \qquad \bar{y}_2 = (CF)^{+} y \tag{91}$$

After simple algebraic manipulations the **transformed dynamics** in the new coordinates take the form:

$$\begin{array}{rcl} \dot{\bar{x}}_1 & = & \bar{A}_{11}\bar{x}_1 + \bar{A}_{12}\bar{x}_2 + F^{\perp} G u \\ \dot{\bar{x}}_2 & = & \bar{A}_{21}\bar{x}_1 + \bar{A}_{22}\bar{x}_2 + (CF)^{+} CGu + \xi \\ \bar{y}_1 & = & \bar{C}_1 \bar{x}_1 \\ \bar{y}_2 & = & \bar{x}_2 \end{array} \tag{92}$$

[1] A Matlab instruction for computing $M_b = M^{\perp}$ for a generic matrix $M$ is `Mb = null(M')'`

with the matrices $\bar{A}_{11}$, ..., $\bar{A}_{22}$, $\bar{C}_1$ such that

$$\begin{bmatrix} \bar{A}_{11} & \bar{A}_{11} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} = TAT^{-1}, \quad \bar{C}_1 = (CF)^{\perp} CT_1. \tag{93}$$

It turns out that the triple $(A, C, F)$ is strongly observable if, and only if, **the pair $(\bar{A}_{11}, \bar{C}_1)$ is observable** [10, 27]. In light of the Assumption A1, this property, that can be also understood in terms of a simplified algebraic test to check the strong observability of a matrix triple, and its satisfaction opens the route to design stable observers for the state of the transformed dynamics (92).

The peculiarity of the transformed system (92) is that $\bar{x}_2$ is available for measurements since it constitutes a part of the transformed output vector $\bar{y}$. Hence, **state observation for system (92) can be accomplished by estimating $\bar{x}_1$ only**, whose dynamics is not affected by the unknown input vector.

The observability of the $(\bar{A}_{11}, \bar{C}_1)$ pair permits the implementation of the following Luenberger observer for the $\bar{x}_1$ subsystem of (92):

$$\dot{\hat{\bar{x}}}_1 = \bar{A}_{11}\hat{\bar{x}}_1 + \bar{A}_{12}\bar{y}_2 + F^{\perp}Gu + L(\bar{y}_1 - \bar{C}_1\hat{\bar{x}}_1) \tag{94}$$

which gives rise to the error dynamics

$$\dot{e}_1 = (\bar{A}_{11} - L\bar{C}_1)e_1, \qquad e_1 = \hat{\bar{x}}_1 - \bar{x}_1 \tag{95}$$

whose eigenvalues can be arbitrarily located by a proper selection of the matrix $L$. Therefore, with properly chosen $L$ we have that

$$\hat{\bar{x}}_1 \to \bar{x}_1 \quad as \quad t \to \infty \tag{96}$$

which implies that the overall system state can be reconstructed by the following relationships

$$\hat{x} = T^{-1} \begin{bmatrix} \hat{\bar{x}}_1 \\ \bar{y}_2 \end{bmatrix} \tag{97}$$

Note that the convergence of $\hat{\bar{x}}_1$ to $\bar{x}_1$ is exponential and can be made as fast as desired. Remarkably, the above estimation is correct in spite of the presence of unmeasurable, possibly very large, external inputs.

## 2.7 Unknown input reconstruction

An estimator can be designed which gives an exponentially converging estimate of the unknown input vector $\xi$. Consider the following estimator dynamics

$$\dot{\hat{\bar{x}}}_2 = \bar{A}_{21}\hat{\bar{x}}_1 + \bar{A}_{22}\bar{y}_2 + (CF)^{+} CGu + v(t) \tag{98}$$

with the estimator injection input $v(t)$ yet to be specified.

Let it can be found a constant $\Xi_d$ such that

$$|\dot{\xi}(t)| \leq \Xi_d \tag{99}$$

Define the estimator "sliding variable"

$$\sigma(t) = \hat{\bar{x}}_2 - \bar{y}_2 = \hat{\bar{x}}_2 - \bar{x}_2 \tag{100}$$

By (98) and (92), the dynamics of the sliding variable $\sigma$ takes the form

$$\dot{\sigma} = f(t) - v(t), \quad f(t) = \bar{A}_{21}\bar{e}_1(t) + \xi(t). \tag{101}$$

Considering (95), the time derivative of the uncertain term $f(t)$ can be evaluated as

$$\dot{f}(t) = \bar{A}_{21}(\bar{A}_{11} - L\bar{C}_1)e_1(t) + \dot{\xi}(t) \tag{102}$$

where $e_1(t)$ is exponentially vanishing. Then, considering (99), by taking any $\overline{\Psi} > \Xi_d$, the next condition

$$|\dot{f}(t)| \leq \overline{\Psi}, \qquad t > T_f, \qquad T_f < \infty \tag{103}$$

will be established starting from a finite time instant $t = T_f$ on.

As shown in [31], if the estimator injection input $v(t)$ is designed according to the next "*Super-Twisting*" algorithm

$$v\,(t) \;=\; \lambda\,|\sigma(t)|^{1/2}\,sign\sigma(t) + v_1(t) \tag{104}$$

$$\dot{v}_1\,(t) \;=\; \alpha sign\;\sigma(t), \tag{105}$$

with the tuning parameters $\alpha$ and $\lambda$ chosen according to the next inequalities

$$\alpha > \overline{\Psi}, \qquad \lambda > \frac{1-\theta}{1+\theta}\sqrt{\frac{\alpha - \overline{\Psi}}{\alpha + \overline{\Psi}}}, \qquad \theta \in (0,1) \tag{106}$$

then both $\sigma$ and its time derivative $\dot{\sigma}$ tend to zero in finite time. Therefore, condition

$$v(t) = \xi(t) + \bar{A}_{21}(\bar{A}_{11} - L\bar{C}_1)e_1(t), \quad t \geq T^*. \tag{107}$$

holds starting from some finite time instant $T^*$ . Since $e_1(t)$ is vanishing, it follows that

$$|v(t) - \xi(t)| \rightarrow 0 \quad as \quad t \rightarrow \infty \tag{108}$$

and, furthermore, the convergence process takes place exponentially. Therefore, under the condition (99), the estimator (98), (100), (104)-(106) allows one to reconstruct the unknown input vector $\xi$ acting on the original system (86).

**Remark 1** It shall be pointed out that the above procedure of state estimation and unknown input reconstruction implies that the next conditions involving the system state, output and unknown input dimensions hold

$$n > m, \qquad p > m \tag{109}$$

**Remark 2** If the vector $\bar{x}_1$ is available for measurements, then a modified estimator can be implemented as follows

$$\dot{\hat{\bar{x}}}_2 = \bar{A}_{21}\bar{x}_1 + \bar{A}_{22}\bar{y}_2 + (CF)^+ CGu + v(t) \tag{110}$$

which, along with the equations (104)-(106),(100),(103) allows for the **finite-time** exact reconstruction of the unknown input vector $\xi$, i.e., given some finite time $T^* > 0$

$$v(t) = \xi(t), \qquad \forall t \geq T^*. \tag{111}$$

The proof can be easily derived by letting $e(t) = 0$ in the previous treatment, particulary relation (107)

This is also another way for reconstruct of the unknown input. We can generate the estimator injection input $v(t)$ in the following manner:

$$v = k_p\sigma + k_i \int_0^t \sigma(\tau)d\tau \tag{112}$$

The closed loop system (110)-(112) can be represented as in the Figure 22. It follows from the scheme in Fig. 22 that the closed loop transfer function between the "input" $\xi + \bar{A}_{21}(x_1 - \hat{x_1})$ and the "output" $v$ is the following

$$P_i(s) = \frac{k_p s + k_i}{s^2 + k_p s + k_i} \tag{113}$$

Then it can be selected the free design parameters $k_p$ and $k_i$ in order to guarantee that such a transfer function is close to the unitary value in a prescribed frequency range $[0, \omega_b]$. Thus, with properly selected gains $k_p$ and $k_i$ the following identity approximately holds

$$v \approx \xi \tag{114}$$

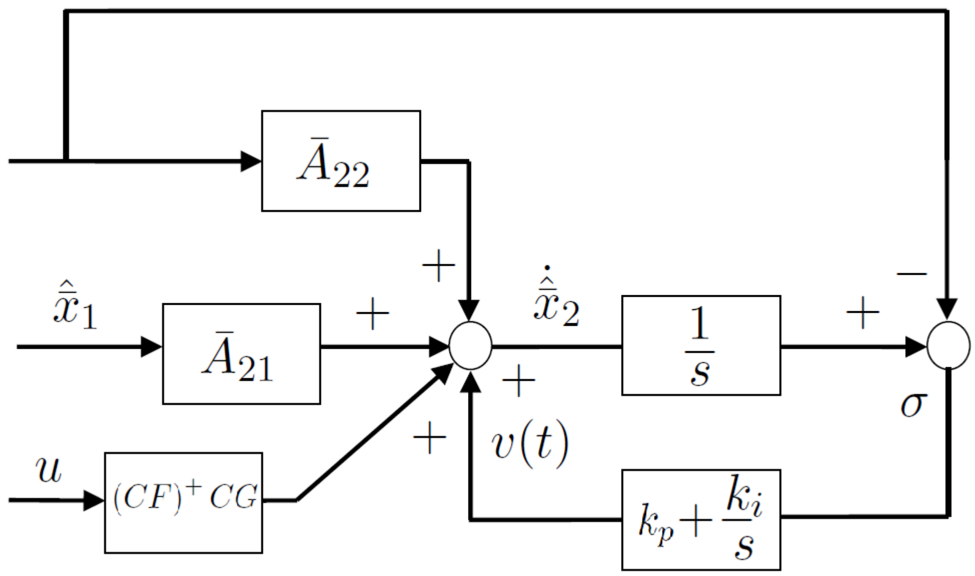



Figure 14: The equivalent block scheme for the closed loop system (110)-(112).

## 2.8 Example

Let the linear system under consideration be autonomous and represented by a third order model with $n = 3$, $m = 1$, $p = 2$, and the next system matrices

$$A = \begin{bmatrix} -3 & 2 & 1 \\ 1 & -4 & 1 \\ -1 & 2 & -3 \end{bmatrix}, \quad F = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{115}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{116}$$

The scalar unknown input is selected as

$$\xi(t) = (1 + sin(t)) \tag{117}$$

Let us calculate the transformation matrices (87):

$$T = \begin{bmatrix} F^{\perp} \\ (CF)^{+}C \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \tag{118}$$

$$U = \begin{bmatrix} (CF)^{\perp} \\ (CF)^{+} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{119}$$

$$\bar{A}_{11} = \begin{bmatrix} -4 & 1 \\ 2 & -3 \end{bmatrix}, \quad \bar{A}_{12} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \bar{A}_{21} = \begin{bmatrix} 2 & 1 \end{bmatrix}, \quad \bar{A}_{22} = \begin{bmatrix} -3 \end{bmatrix}, \tag{120}$$

$$\bar{C}_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}; \quad \bar{C}_2 = \begin{bmatrix} 1 \end{bmatrix}. \tag{121}$$

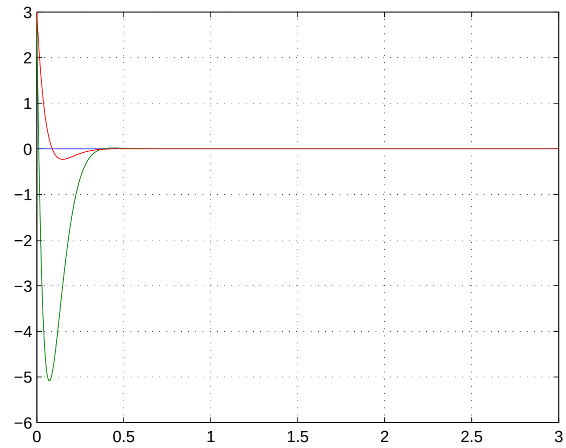



Figure 15: The components of the observation error vector

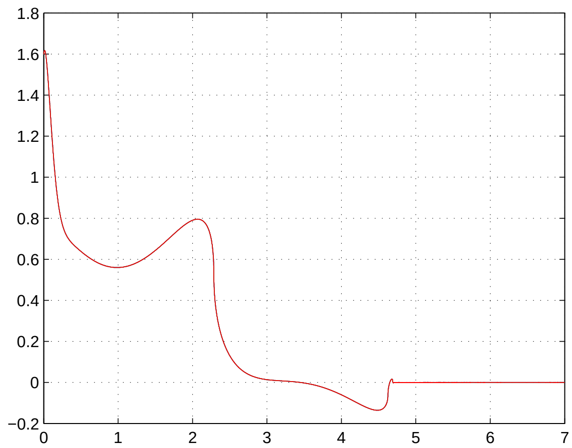



Figure 16: The difference between the unknown input and its reconstruction with SM Observer

System (115) is strongly observable, in fact:

$$rank(\bar{A}_{11}, \bar{C}_1) = n - m = 2. \tag{122}$$

So, we can use the observer (94) for $x_1$. Considering a Luenberger Observer, in which initial state condition are $[3, 3, 3]^T$, the gain $L = [98, 23]^T$ is chose in order to have the eigenvalues of $(\bar{A}_{11} - L\bar{C}_1)$ in $[-10, -20]$ we obtain the following behavior of the error $x - \hat{x}$ (see fig.(15)).

Now let us address the reconstruction of the sinusoidal unknown input $u(t)$. According to the suggested procedure with the SM "super twisting" observer (98) has been implemented with $\lambda = 1.5$ and $\alpha = 1$. In the present example we obtain figure (16).

In the next figure (17), still in order to reconstruct the sinusoidal unknown input, it is shown

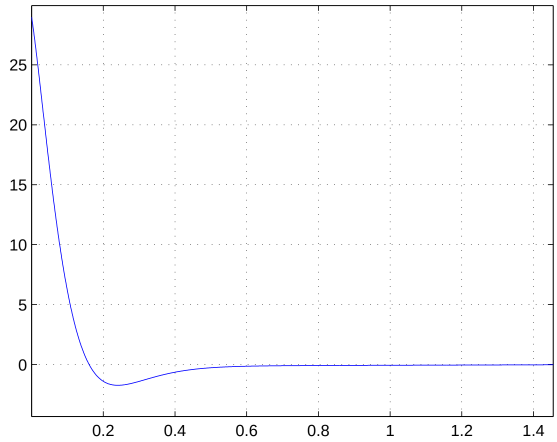



Figure 17: The difference between the unknown input and its reconstruction with PI Observer

the behavior of the "PI observer" (112), in which $k_p = 1$ and $k_i = 2$.

# References

[1] D. G. Luenberger, Observing the state of a linear system, IEEE Trans. mil, Electron., vol. MIL-8, pp. 74-80, Apr. 1964.

[2] D. G. Luenberger, Observers for multivariable systems, IEEE Trans, automat. Contr., vol. AC-11, pp. 190-197, Apr. 1966.

[3] M. Aoki and J. R. Huddle, Estimation of state vector of a dubmat. Contr., vol. AC-11, pp. 190-197, Apr. 1966. linear stochastic system with a constrained estimator, IEEE Trans. Automat. Contr., vol. AC-12, pp. 432- 433.

[4] W. A. Wolovich, On state estimation of observable systems, in 1968 Joint Automatic Control Conf., pp. 210-220.

[5] J. J. Bongiorno and D. C. Youla, On observers in multivariable control systems, Int. J. Contr., vol. 8, no. 3, pp. 221-243, 1968. 1968.

[6] F. Dellon and P. E. Sarachik, Optimal control of unstable linear plants with inaccwible state,, IEEE Trans. Automat. Contr., vol. AC-13, pp. 491-495, Oct. 1968.

[7] R. H. Ash and I. Lee, State estimation in linear systems; a unified theory of minimum order observers", 3rd Hawaii Sys. Conf., Jan. 1970.

[8] G. W. Johnson, A deterministic theory of estimation and control, IEEE Trans. Automat. Contr., vol. AC-14, pp. 380-384, Aug. 1969.

[9] E. Tse and M. Athans, Optimal minimal-order observer estimators for discrete linear the-varying systems, IEEE Trans. Automat. Contr., vol. AC-1.5, pp. 416-426, 1970.

[10] S. D. G. Cumming, Design of observers of reduced dynamics, Electron. Lett., vol. 5, pp. 213-214, 1969.

[11] R. T. K. Chen, On the construction of state observers in multivariable control systems, presented at the Nat. Electron Conf., Dec. 8-10, 1969.

[12] Y. O. Yiiksel and J. J. Bongiorno, Observers for linear multivariable systems Kith applications, pp. 603-613.

[13] B. E. Bona, Designing observers for time-varying state systems, in Rec. 4th Asi1oma.r Conf. Circuits a.nd Systems, 1970.

[15] M.M. Newmann, A continuous-time reduced-order filter for estimating the state vector of a linear stochastic system, Int. J. Contr., vol. 11, no. 2, pp. 229-239, 1970.

[16] V. V. S. Sarma and B. L.Deekshatulu, Optimal control when some of the state variables are not measurable, Int. J. Contr., vol. 7, no. 3, pp. 251-256, 1968.

[17] B. Porter and M. A. Woodhead, Performance of optimal control systems when some of the state variables are not measurable, Int. J. Contr., vol. 8, no. 2, pp. 191-195, 1968.

[18] M. M. Newman, Optimal and sub-optimal control using an observer when some of the state variables are not measurable Int. J. Contr., vol. 9, pp. 281-290, 1969.

[19] I. G. Sarnla and C. Jayaraj On the use of observers in finitetime optimal regulator problems Int. J. Contr., vol. 11, no. 3, pag.221-243, 1968.

[20] W.M. Wonham, Dynamic observers-geometric theory IEEE Trans. Automat. Cont., vol. AC-15, pp.238-259. Apr. 1970.

[21] A.E. Bryson,Jr., and D. G. Luenberger, The synthesis of regulator logic using state-variable concepts" Proc. IEEE, vol. 58, pp. 1803-1811, Nov. 1970.

[22] B. Gopinath, On the control of linear multiple input-output. systems" Bell Syst. Tech. J., Mar. 1971.

[23] K. Frederick, and G. F. Franklin, Design of piecewiselinear swit.ching functions for relay control systems, IEEE Trans. Automat. Contr., vol. AC-12, pp. 380-387, Aug. 1967.

[24] Molinari, B.P.(1976). "A strong contollability and observability in linear multivariable control". *IEEE Transaction on Automatic Control*, 21(5), 761-764.

[25] F.J. Bejarano, L. Fridman and A. Poznyak (2007). "Exact state estimation for linear systems with unknown inputs based on hierarchical super-twisting algorithm". *Int. J. Robust Nonlinear Control*, 17(18),1734-1753.

[26] Fridman, L., Levant, A. and Davila, J.(2007). "Observation of Linear Systems with Unknown Inputs via High-Order Sliding-Mode", *International Journal of Systems Science*, 38(10), 773-791.

[27] Hautus, M. L. J.(1983). "Strong detectability and observers" Linear Algebra and its Applications, 50, 353-368.

[28] A. Giua, C. Seatzu, "Analisi dei Sistemi Dinamici", Springer-Verlag, 2005.

[29] V. Utkin, "Principle of identification using sliding regimes". *Soviet Physics Doklady* 26, pp. 271-272, 1981.

[30] S. Drakunov, V. Utkin, "Sliding Mode Observers. Tutorial". *Proc. of the 34th CDC conference*, New Orleans (LA), December 1995.

[31] A. Levant, "Sliding order and sliding accuracy in sliding mode control", *Int. Journal of Control*, vol. 58, pp. 1247-1263, 1993.

# 3 Consensus Algorithms in Networked Systems

## 3.1 Introduction

This Chapter provides, using the concept of the previous chapters, a theoretical framework for the analysis of a consensus algorithms for multi-agent networked systems with an emphasis on the role of directed information flow, robustness to changes in network topology due to link/node failures, time-delays, and performance guarantees. An overview of basic concepts in information consensus in networks, methods of convergence and performance analysis for the algorithms are provided. Our analysis framework is based on tools from matrix theory, algebraic graph theory, and control theory. We discuss the connections between consensus problems in networked dynamic systems and same applications including synchronization of coupled oscillators, flocking, formation control, fast consensus in small-world networks, Markov processes and gossip-based algorithms, load balancing in networks, rendezvous in space, distributed sensor fusion in sensor networks, and belief propagation. We establish direct connections between spectral and structural properties of complex networks and the speed of information diffusion of consensus algorithms. A brief introduction is provided on networked systems with nonlocal information flow that are considerably faster than distributed systems with latticetype nearest neighbor interactions. Simulation results are presented that demonstrate the role of small-world effects on the speed of consensus algorithms and cooperative control of multi-vehicle formations.

## 3.2 Consensus in Networks

The interaction topology of a network of agents is represented using a directed graph $G = (V, E)$ with the set of nodes $V = 1, 2, ..., n$ and edges $E \subseteq V \times V$. The neighbors of agent $i$ are denoted by $N_i = \{j \in V : (i, j) \in E\}$. According to [70], a simple consensus algorithm to reach an agreement regarding the state of $n$ integrator agents with dynamics $\dot{x}_i(t) = u_i$ can be expressed as an $n$-th order linear system on a graph:

$$\dot{x}_i(t) = \sum_{j \in N_i} (x_j(t) - x_i(t)) + b_i(t), \qquad x_i(0) = z_i \in \mathbb{R}, b_i(t) = 0 \tag{123}$$

The collective dynamics of the group of agents following protocol (123) can be written as

$$\dot{x} = -Lx \tag{124}$$

where $L = [l_{ij}]$ is the *graph Laplacian* of the network, defined in (42), and in the case of unweighted graph its elements are defined as follows:

$$l_{ij} = \begin{cases} -1 & i \neq j, \\ |N_i| & i = j. \end{cases} \tag{125}$$

Here, $|N_i|$ denotes the number of neighbors of node i (or out-degree of node i). Fig (18) shows two equivalent forms of the consensus algorithm in equations (123) and (124) for agents with a scalar state. The role of the input bias $b$ in Fig(18-(b)) is defined later. According to the definition of graph Laplacian in (125), all row-sums of $L$ are zero because of $\sum_j l_{ij} = 0$. Therefore, $L$ has always a zero eigenvalue $\lambda_1 = 0$. This zero eigenvalues corresponds to the eigenvector $\mathbf{1} = (1, ..., 1)^T$ because $\mathbf{1}$ belongs to the null-space of $L$ ($L\mathbf{1} = 0$). In other words, an equilibrium of system (124) is a state in the form $x^* = (\alpha, ..., \alpha)^T = \alpha\mathbf{1}$ where all nodes agree. Based on analytical tools from algebraic graph theory [2], we later show that $x^*$ is a unique equilibrium of (124) (up to a constant multiplicative factor) for connected graphs.

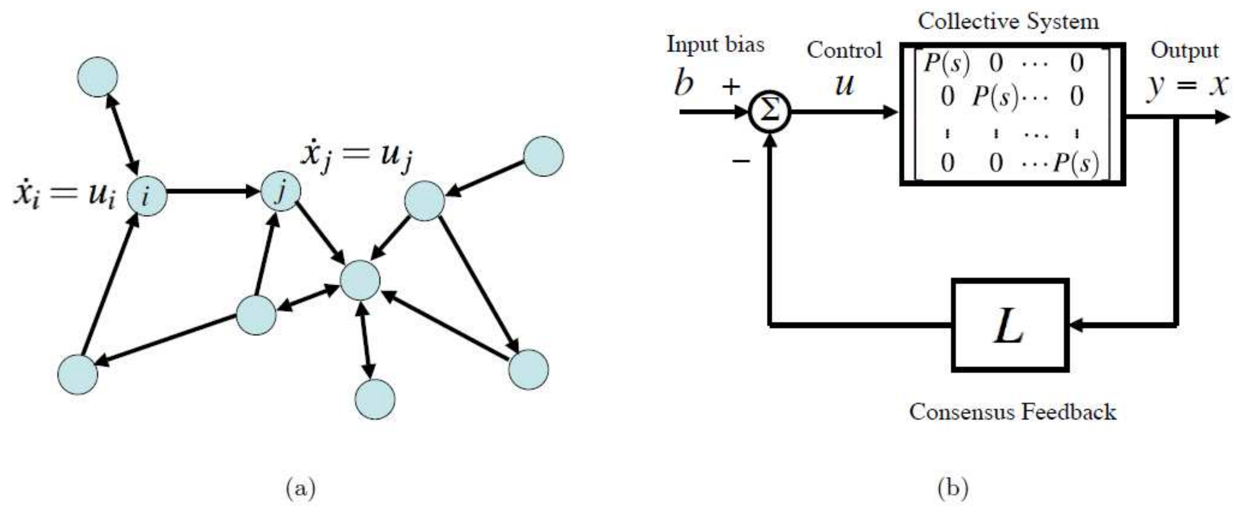



Figure 18: Two equivalent forms of consensus algorithms: (a) a network of integrator agents in which agent $i$ receives the state $x_j$ of its neighbor, agent $j$, if there is a link $(i, j)$ connecting the two nodes; and (b) the block diagram for a network of interconnected dynamic systems all with identical transfer functions $P(s) = 1/s$. The collective networked system has a diagonal transfer function and is a MIMO (multi-input multi-output) linear system.

One can show that for a connected network, the equilibrium $x^* = (\alpha, ..., \alpha)^T$ is globally exponentially stable. Moreover, the consensus value is $\alpha = 1/n \sum_i z_i$ that is equal to the average of the initial values.

This implies that irrespective of the initial value of the state of each agent, all agents reach an asymptotic consensus regarding the value of the function $f(z) = 1/n \sum_i z_i$. While the calculation of $f(z)$ is simple for small networks, its implications for very large networks is more interesting. For example, if a network has $n = 10^6$ nodes and each node can only talk to $log_{10}(n) = 6$ neighbors, finding the average value of the initial conditions of the nodes is more complicated. The role of protocol (123) is to provide a systematic consensus mechanism in such a large network to compute the average. There are a variety of functions that can be computed in a similar fashion using synchronous or asynchronous distributed algorithms.

### 3.3 Information Consensus in Networked Systems

Consider a network of decision-making agents with dynamics $\dot{x}_i = u_i$ interested in reaching a consensus via local communication with their neighbors on a graph $G = (V, E)$. By "reaching a consensus", we mean asymptotically converging to a one-dimensional agreement space characterized by the following equation

$$x_1 = x_2 = ... = x_n \tag{126}$$

This agreement space can be expressed as $x = \alpha\mathbf{1}$ where $\mathbf{1} = (1, ..., 1)^T$ and $\alpha \in \mathbb{R}$ is the collective decision of the group of agents. Let $A = [a_{ij}]$ be the adjacency matrix of graph $G$. The set of neighbors of agent $i$ is $N_i$ and defined by

$$N_i = \{j \in V : a_{ij} \neq 0\}; \qquad V = \{1, ..., n\}. \tag{127}$$

Agent $i$ communicates with agent $j$ if $j$ is a neighbor of $i$ (or $a_{ij} \neq 0$). The set of all nodes and their neighbors defines the edge set of the graph as $E = \{(i, j) \in V \times V : a_{ij} \neq 0\}$. A **dynamic graph** $G(t) = (V, E(t))$ is a graph in which the set of edges $E(t)$ and the adjacency matrix $A(t)$ are time-varying. Clearly, the set of neighbors $N_i(t)$ of every agent in a dynamic graph is a time-varying set as well. Dynamic graphs are useful for describing the network topology of mobile sensor networks and flocks [66]. It is shown in [70] that the linear system

$$\dot{x}_i(t) = \sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t)) \tag{128}$$

is a distributed consensus algorithm, i.e. it guarantees convergence to a collective decision via local inter-agent interactions. Assuming that the graph is undirected ($a_{ij} = a_{ji}$ for all $i$, $j$), it follows that the sum of the state of all nodes is an invariant quantity, or $\sum_i \dot{x}_i = 0$. In particular, applying this condition twice at times $t = 0$ and $t = \infty$ gives the following result

$$\alpha = \frac{1}{n} \sum_i x_i(0). \tag{129}$$

In other words, if a consensus is asymptotically reached, then necessarily the collective decision is equal to the average of the initial state of all nodes. A consensus algorithm with this specific invariance property is called an **average-consensus** algorithm [8] and has broad applications in distributed computing on networks (e.g. sensor fusion in sensor networks). The dynamics of system (128) can be expressed in a compact form as

$$\dot{x_i} = -Lx \tag{130}$$

where $L$ is known as the **graph Laplacian** of $G$. The graph Laplacian is defined as

$$L = D - A \tag{131}$$

where $D = diag(d_1, ..., d_n)$ is the degree matrix of $G$ with elements $d_i = \sum_{j \neq i} a_{ij}$ and zero off-diagonal elements. By definition, $L$ has a right eigenvector $\mathbf{1}$ associated with the zero eigenvalue because of the identity $L\mathbf{1} = 0$. For the case of undirected graphs, graph Laplacian satisfies the following sum-of-squares (SOS) property

$$x^T Lx = \frac{1}{2} \sum_{(i,j) \in E} a_{ij}(x_j - x_i)^2. \tag{132}$$

By defining a quadratic disagreement function as

$$\varphi(x) = \frac{1}{2} x^T Lx \tag{133}$$

it becomes apparent that algorithm (128) is the same as

$$\dot{x} = -\nabla \varphi(x) \tag{134}$$

or the gradient-descent algorithm. This algorithm globally asymptotically converges to the agreement space provided that two conditions hold:

1) $L$ is a positive semidefinite matrix and

2) the only equilibrium of (128) is $\alpha\mathbf{1}$ for some $\alpha$.

Both of these conditions hold for a connected graph and follow from the SOS property of graph Laplacian in (132). Therefore, an average-consensus is asymptotically reached for all initial states. This fact is summarized in the following Lemma:

**Lemma 9** . *Let $G$ be a connected undirected graph. Then, the algorithm (128) asymptotically solves the average-consensus problem for all initial states.*

### 3.3.1 Algebraic Connectivity and Spectral Properties of Graphs

Spectral properties of Laplacian matrix are instrumental in analysis of convergence of the class of linear consensus algorithms in (128). According to Gershgorin theorem [5], all eigenvalues of $L$ in the complex plane are located in a closed disk centered at $\Delta + 0j$ with a radius of $\Delta = max_i d_i$, i.e. the maximum degree of a graph. For undirected graphs, $L$ is a symmetric matrix with real eigenvalues and therefore the set of eigenvalues of $L$ can be ordered sequentially in an ascending order as

$$0 = \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n \leq 2\Delta \tag{135}$$

The zero eigenvalue is known as the trivial eigenvalue of $L$. For a connected graph $G$, $\lambda_2 > 0$ (i.e. the zero eigenvalue is isolated). The second smallest eigenvalue of Laplacian $\lambda_2$ is called **algebraic connectivity** of a graph [1]. Algebraic connectivity of the network topology is a measure of performance/speed of consensus algorithms [6].

**Example 8** . *Fig. (19) shows two examples of networks with different topologies. Both graphs are undirected and have 0-1 weights. Every node of the graph in Fig.(19) (a) is connected to its 4 nearest neighbors on a ring. The other graph is a proximity graph of points that are distributed uniformly at random in a square. Every node is connected to all of its spatial neighbors within a closed ball of radius $r > 0$. Here are the important degree information and Laplacian eigenvalues of these graphs:*

$$a)\lambda_1 = 0, \lambda_2 = 0.48, \lambda_n = 6.24, \Delta = 4 \tag{136}$$

$$b)\lambda_1 = 0, \lambda_2 = 0.25, \lambda_n = 9.37, \Delta = 10 \tag{137}$$

In both cases, $\lambda_i < 2\Delta$ for all $i$.

### 3.3.2 Convergence Analysis for Directed Networks

The convergence analysis of the consensus algorithm (128) is equivalent to proving that the agreement space characterized by $x = \alpha_1$, $\alpha \in \mathbb{R}$ is an asymptotically stable equilibrium of system (128). The stability properties of system (128) is completely determined by the location of

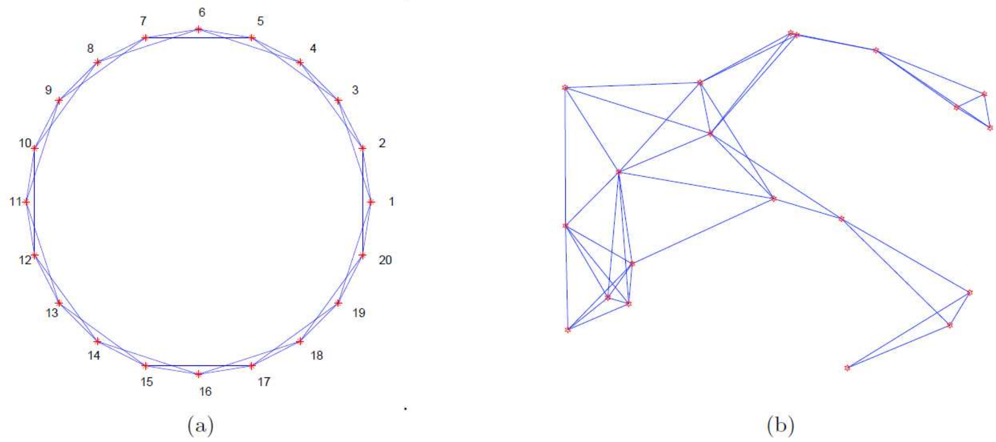



Figure 19: Examples of networks with n = 20 nodes: a) a regular network with 80 links and b) a random network with 65 links.

the Laplacian eigenvalues of the network. The eigenvalues of the adjacency matrix are irrelevant to the stability analysis of system (128), unless the network is $k$-regular (all of its nodes have the same degree k). The following lemma combines a well-known rank property of graph Laplacians with Gershgorin theorem to provide spectral characterization of Laplacian of a fixed directed network $G$. Before stating the lemma, we need to define the notion of strong connectivity of graphs. A graph is strongly connected (SC) if there is a directed path connecting any two arbitrary nodes of the graph.

**Lemma 10** . *(spectral localization) Let $G$ be a strongly connected digraph on $n$ nodes. Then $rank(L) = n - 1$ and all nontrivial eigenvalues of $L$ have positive real parts. Furthermore, suppose $G$ has $c \geq 1$ strongly connected components, then $rank(L) = n - c$.*

**Proof 4** . *The proof of the rank property for digraphs is given in [6]. The proof for undirected graphs is available in the algebraic graph theory literature [2]. The positivity of the real parts of the eigenvalues follow from the fact that all eigenvalues are located in a Gershgorin disk in the closed right-hand plane that touches the imaginary axis at zero. The second part follows from the first part after relabeling the nodes of the digraph so that its Laplacian becomes a block diagonal matrix.*

**Remark 4** . *Lemma (10) holds under a weaker condition of existence of a directed spanning tree for $G$. $G$ has a directed spanning tree if there exists a node $r$ (a root) such that all other nodes can be linked to $r$ via a directed path. This type of condition on existence of directed spanning trees have appeared in [3], [4], [7]. The root node is commonly known as a leader [3]. The essential results regarding convergence and decision value of Laplacian-based consensus algorithms for directed networks with a fixed topology are summarized in the following theorem.*

*Before stating this theorem, we need to define an important class of digraphs that shall appear frequently throughout this section.*

**Definition 6** *(balanced digraphs [6]) A digraph $G$ is called balanced if $\sum_{j\neq i} a_{ij} = \sum_{j\neq i} a_{ji}$ for all $i \in V$. In a balanced digraph, the total weight of edges entering a node and leaving the same node are equal for all nodes. The most important property of balanced digraphs is that $w = \mathbf{1}$ is also a left eigenvector of their Laplacian (or$\mathbf{1}^T L = 0$).*

**Theorem 11** . *Consider a network of $n$ agents with topology $G$ applying the consensus algorithm*

$$\dot{x}_i(t) = \sum_{j\in N_i} a_{ij}(x_j(t) - x_i(t)), \tag{138}$$

$$x(0) = z \tag{139}$$

*Suppose $G$ is a strongly connected digraph. Let $L$ be the Laplacian of $G$ with a left eigenvector $\gamma = (\gamma_1, ..., \gamma_n)$ satisfying $\gamma^T L = 0$. Then*

*i) A consensus is asymptotically reached for all initial states;*

*ii) The algorithm solves the $f$-consensus problem with the linear function $f(z) = (\gamma^T x)/(\gamma^T \mathbf{1})$, i.e. the group decision is $\alpha = \sum_i w_i x_i$ with $\sum_i w_i = 1$;*

*iii) If the digraph is also balanced, an average-consensus is asymptotically reached and $\alpha = (\sum_i x_i(0))/n$.*

**Proof 5** . *The convergence of the consensus algorithm follows from Lemma 2. To show part ii), note that the collective dynamics of the network is $\dot{x} = -Lx$. This means that $y = \gamma^T x$ is an invariant quantity due to $\dot{y} = -\gamma^T Lx = 0, \forall x$. Thus, $lim_{t\to\infty} y(t) = y(0)$, or $\gamma^T(\alpha\mathbf{1}) = \gamma^T x(0)$ that implies the group decision is $\alpha = (\gamma^T z)/\sum_i \gamma_i$. Setting $w_i = \gamma_i/\sum_i \gamma_i$, we get $\alpha = w^T z$. Part iii) follows as a special case of the statement in part ii) because for a balanced digraph $\gamma = \mathbf{1}$ and $w_i = 1/n, \forall i$.*

# References


[1] M. Fiedler. Algebraic connectivity of graphs. Czechoslovak Mathematical Journal, 23(98):298-305, 1973.

[2] C. Godsil and G. Royle. Algebraic Graph Theory, volume 207 of Graduate Texts in Mathematics. Springer, 2001.

[3] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans. on Automatic Control, 48(6):988-1001, June 2003.

[4] L. Moreau. Stability of multiagent systems with time-dependent communication links. IEEE Trans. on Automatic Control, 50(2):169-182, 2005.

[5] R. Olfati-Saber. Flocking for Multi-Agent Dynamic Systems: Algorithms and Theory. IEEE Trans. on Automatic Control, 51(3), Mar. 2006.

[6] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans. on Automatic Control, 49(9):1520-1533, Sep. 2004.

[7] W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. IEEE Trans. on Automatic Control, 50(5):655-661, 2005.

[8] R. O. Saber and R. M. Murray. Consensus protocols for networks of dynamic agents. Proceeding of the 2003 American Control Conference, 2:951-956, 2003.

# 4 Decentralized Estimator Based on Consensus Algorithm

## 4.1 Preliminaries

The previous chapter has been devoted to explore the concept of "consensus", and to show some general cases in which it is possible to achieve it. In this Chapter we show a recent solution of the problem of decentralized state estimation in which it is used a combination of Kalman filters, already introduced in chapter 2 and consensus algorithms.

## 4.2 Introduction

A great deal of attention has been paid among researchers working in field of complex dynamic systems to the problem of decentralized estimation. This notion should be considered in its generic sense, since under this term one can consider different structures that are totally or partially decentralized and/or hierarchical. The key requirement is that the large scale system has to be modelled as an interconnection of subsystems, and that each subsystem have a local decision maker (intelligent agent) associated with it. Depending on the available resources, the agent might have access to different information, such as the characteristics of the sensors, properties and models of the system and its environment and communication channels between the agents. One class of decentralized estimators is obtained starting from the parallelization of the globally optimal Kalman filter; typically, such estimators possess a fusion center which generates the global estimate (e.g., see [8]). Attempts to provide an insight into basic principles and structures of decentralized estimation can be found in [13], [11]. One of general design tools has been found to be the inclusion principle, starting from the expansion/contraction paradigm: the observed large scale system is expanded, decomposed into subsystems and contracted back to the original system space after designing local estimators for the extracted subsystems [4]. Successful applications of this approach have been reported in [8]. However, none of the existing methodologies is able to provide a systematic and general way of designing communication links between the agents without recurring to a strong fusion center. In the 1980s, important results were obtained in the area of distributed asynchronous iterations in parallel computation and distributed optimization (e.g. [1]). On the other hand, a very intensive research has been carried out recently in the field of multi-agent systems, including coordinated control of multiple vehicle systems such as unmanned air vehicles (UAV) formation flying, coordinated rendezvous, and multiple robot coordination, as well as in the field of sensor networks with broad applications in surveillance and environment monitoring, collaborative processing of information, and gathering scientific data from spatially distributed sources (e.g., see [8], [15], [11]). These references have a common methodology: they all use the **consensus** strategy, which is found to be beneficial for solving diverse problems in distributed computation, distributed signal processing and multi-agent networks. The state estimation problem itself is deeply embedded in this line of thought either implicitly, through the very definition of consensus algorithms (e.g., see [18]), or explicitly, where the dynamic consensus strategy between multiple agents is used for obtaining (on the basis of averaging) estimates of the quantities used subsequently for generating optimal parameter or state estimates (e.g., see [21], [20]). However, none of these approaches is aimed at establishing any type of collaboration between local estimators in the overlapping decentralized estimation problem. In this chapter a recent state estimation algorithm for linear complex systems is recalled [9] based on:

1) overlapping system decomposition;

2) implementation of local state estimators by intelligent agents according to their sensing and computing resources;

3) application of a consensus strategy providing the global state estimates to all the agents in the network. The main definition of the problem is given in Section 4.3. In Section 4.4 the recalled [9] estimation algorithm is described. The algorithm works in continuous time, and structurally resembles the distributed computation algorithm recalled in [1]. Formally speaking, it is composed of overlapping decentralized Kalman filters combined together in a multi-agent network on the basis of a consensus strategy ([18]). In Section 4.5 stability of the recalled [9] scheme is discussed. It is proved that it is possible to find, under general conditions concerning the local estimators with their a priori knowledge and the network topology, such a consensus scheme which ensures asymptotic stability of the whole estimator. Section 4.6 provides a strategy aimed at obtaining the gains of the consensus scheme on the basis of optimization: it is shown how to minimize the total mean-square estimation error with respect to the unknown consensus gains and to obtain in such a way a general and efficient estimation scheme. Section 4.7 contains several examples illustrating the main properties of the recalled [9] estimation algorithm.

## 4.3 Overlapping Decentralized Estimation

Let a continuous-time complex stochastic system be represented by

$$S : \left\{ \begin{array}{rcl} \dot{x} & = & Ax + \Gamma e \\ y & = & Cx + v \end{array} \right. \tag{140}$$

where $x = (x_1, ..., x_n)^T$ is the state vector, $y = (y_1, ..., y_p)^T$ is the output vector, $e = (e_1, ..., e_q)^T$

and $v = (v_1, ..., v_p)^T$ are stochastic noise signals, and $A, \Gamma$ and $C$ are constant matrices of appropriate dimension.

It is assumed that $e(t)$ and $v(t)$ are mutually independent stochastic gaussian processes, with zero-mean and covariance matrices as follows:

$$E\{e(t)e(\tau)^T\} = Q\delta(t-\tau), \qquad E\{v(t)v(\tau)^T\} = R\delta(t-\tau). \tag{141}$$

We shall consider the problem of decentralized estimation of the state $x$ of $\mathbf{S}$.

In the decentralized scheme for state vector estimation, we assume to have $N$ autonomous agents which have the goal to generate their estimates of the state vector $x(t)$ on the basis of **locally available measurements** (each agent can measure only a subset of the elements of the output vector $y$). More precisely we assume that the $i$-th agent can measure, with some noise, the $p_i$-dimensional vector $y^{(i)}$ containing those output variables with indices specified by the so-called agent's **output index set** $I_y^i$, which is defined as follows:

$$I_i^y = \{l_1^i, ..., l_{p_i}^i\}, \qquad l_1^i, ..., l_{p_i}^i \in \{1, ..., p\}, \quad l_1^i < ... < l_{p_i}^i, \quad p_i \leq p. \tag{142}$$

We also define **state index set** $I_i^x$ of $i$-th agent as follows.

$$I_i^x = \{l_1^i, ..., l_{n_i}^i\}, \qquad l_1^i, ..., l_{n_i}^i \in \{1, ..., n\}, \quad l_1^i < ... < l_{n_i}^i, \quad n_i \leq n \tag{143}$$

The set $I_i^x$ denotes the indices of the original state variables $x_i$'s that explicitly affect the local measurement output equation.

The **local system model** available to the $i-th$ agent is defined as follows

$$S_i : \left\{ \begin{array}{lcl} \dot{x}^{(i)} & = & A^{(i)}x^{(i)} + \Gamma^i e \\ y^{(i)} & = & C^{(i)}x^{(i)} + v^{(i)} \end{array} \right. , \quad i = 1, 2, ..., N \tag{144}$$

where $x^{(i)}$ is a vector of dimension $n_i \leq n$, $y^{(i)}$ was previously defined as the local output vector with dimension $p_i \leq p$, and $A^i$, $C^i$, $\Gamma^i$ are constant matrices of appropriate dimension. Clearly, the matrices $C^i$, that define the local output measurement equations, contain a subset of the entries of the "complete" output matrix $C$. More precisely, matric $C^i$ contains the elements of $C$ having the row and column indexes specified by the set $I_i^y \times I_i^x$. Similarly, matrix $A^i$ contains those elements of the matrix $A$ with the row and column indexes taken in the set $I_i^x \times I_i^x$. Matrix $\Gamma^i$ contains a subset of the rows of the complete matrix $\Gamma$, with the selected row indexes specified in the state index set $I_i^x$. Vector $v^{(i)}$ represents a "reduced-order" measurement noise vector having zero mean and covariance matrix

$$E\left[v^{(i)}(t)\left(v^{(i)}(\tau)\right)^T\right] = R^{(i)}\delta(t-\tau) \tag{145}$$

Matrix $R^{(i)}$ contains the elements of the original matrix $R$ having the row and column indexes taken in the set $I_i^y \times I_i^y$. Note that the noise vectors $v^{(i)}$ correspond to different, uncorrelated, realizations of the statistical process.

System (144) defines $N$ **overlapping local subsystems** of the overall plant (140). Note that the decomposition of (140) into the overlapping subsystems (144) does not rely on any inherent decomposition of the original plant matrices $A$ in (140). Hence, the agents need not a complete knowledge of the overall system dynamics, but only of a part of it.

Starting from the model $S_i$ and the accessible measurements $y^{(i)}$, each agent is able to generate autonomously its own local estimate $\hat{x}^{(i)}$ of the vector $x^{(i)}$ using an estimator which can be designed on the basis of (144). Having in mind the nature of the whole system $S$, the following local steady-state Kalman filters will be assumed to be implemented by each agent [1]:

$$\overline{E_i} : \dot{\hat{x}}^{(i)} = A^{(i)}\hat{x}^{(i)} + L^{(i)}(y^{(i)} - C^{(i)}\hat{x}^{(i)}) \tag{146}$$

where $L^{(i)}$ is the steady state Kalman gain given by

$$L^{(i)} = P^{(i)}C^{(i)T}R^{(i)-1} \tag{147}$$

where $P^{(i)}$ is a solution of the algebraic Riccati equation

$$A^{(i)}P^{(i)} + P^{(i)}A^{(i)^T} - P^{(i)}C^{(i)^T}R^{(i)^{-1}}C^{(i)}P^{(i)} + Q\Gamma^{(i)^T} = 0. \tag{148}$$

We shall assume in the sequel that the pairs $(A^{(i)}, \Gamma^{(i)})$ are stabilizable and the pairs $(A^{(i)}, C^{(i)})$ detectable, so that $A^{(i)} - L^{(i)}C^{(i)}$, the error matrices of the estimators (146), are asymptotically stable and $P^{(i)} > 0, i = 1, ..., N$ [1].

## 4.4 Consensus based state observation

As stated in the Introduction, our task is to formulate an estimation strategy which would enable all the agents in the network to get reliable estimates of the whole state vector $x$ on the basis of:

(1) local estimates $\hat{x}^{(i)}$ available at each node, and

(2) decentralized communication strategy uniform for all the nodes. We recall in this chapter an algorithm based on the introduction of a consensus scheme (see e.g. [1], [15], [11], [18]). Namely,

the estimate of $x$ generated by the $i$-th agent is given by

$$E_i: \quad \dot{\xi}^i = A^i \xi^i + \sum_{j=1, j\neq i}^{N} K_{ij}(\tilde{\xi}^{i,j} - \xi^i) + L_i \left( y^{(i)} - C^i \xi^i \right), \quad i = 1, ..., N. \tag{149}$$

Where:

· $A^i \in R^{n\times n}$ and whose elements having the row and column indexes in the set $I_x^i \times I_x^i$ are equal to the corresponding entries of $A^{(i)}$, while the remaining elements are zeros,

· $C^i \in R^{p_i \times n}$ and whose elements having the column index in the set $I_x^i$ are equal to the corresponding entries of $C^{(i)}$, while the remaining elements are zeros,

Furthermore:

· $K_{ij} \in R^{n\times n}$ are constant consensus gain matrices,

· $\tilde{\xi}^{i,j}$ is the noisy estimate $\xi^j$ communicated by the $j$-th agent ($j \neq i$), i.e., $\tilde{\xi}^{i,j} = \xi^j + w_{ij}$.

· $L_i \in R^{n\times p_i}$ and whose elements having the row index in the set $I_x^i$ are equal to the corresponding entries of the steady-gain matrix of the local Kalman filter for the system (144), while the remaining elements are zeros,

It is possible to observe that the algorithm is based on a combination of: a) decentralized overlapping estimators represented by (146) and b) a consensus scheme with matrix gains $K_{ij}$, [18]. The scheme reduces to the local estimators when the "consensus part" is eliminated $K_{ij} = 0$. When the local estimators are eliminated, the "consensus part" alone asymptotically provides $\xi^{(i)} = \xi$ under a proper choice of the matrices $K_{ij}$ when the communication noise $w_{ij}$ can be neglected, where $\xi$ is a weighted sum of the a priori estimates $\xi^{(i)}(t_0)$ and $t_0$ is the initial time instant [18]. Notice that the estimator $E_i$ reminds structurally of the discrete-time distributed optimization algorithm recalled in [1], it performs "computation" by evaluating the updating part described by (146), and enforces the "agreement" between the agents by evaluating the remaining terms in (149). Formally, the "estimation part" of (149) is obtained simply by rewriting $\overline{E_i}$ and placing the estimates $\hat{x}^{(i)}$ generated by $\overline{E_i}$ at the correct positions in $\xi^{(i)}$. The "consensus part" requires, however, additional specifications. The second term in the right hand side of (149) is obtained by generalizing the consensus scheme presented, for example, in [18]. Having in mind that any analysis of the properties of this scheme is faced with serious problems in the general case when $K_{ij}$ are full matrices (which could obscure presentation of the main properties of the recalled [9] estimation scheme), we shall assume in the sequel that $K_{ij}$ are diagonal matrices $K_{ij} = diag\{k_1^{ij}, ..., k_n^{ij}\}$, where $k_\nu^{ij}$, $\nu = 1, ..., n$, $i, j = 1, ..., N$; moreover, all theoretical and experimental analyses we have done show that the resulting structure gives enough freedom for obtaining reliable estimation schemes. We shall assume further that $k_\nu^{ij} = g_\nu^{ij} h_\nu^{ij}$ where $g_\nu^{ij}$ directly reflects structural properties of $S$ and $S_j$ and the uncertainty in the local estimate $\hat{x}^{(j)}$, and $h_\nu^{ij}$ reflects properties of communication links. For example, $g_\nu^{ij}$ can be chosen to be nonzero for $\nu \in I_x^j$ and inversely proportional to the variance of the local estimate of the $\nu - th$ component of the state vector, and $h_\nu^{ij} = h_{ij}$ to be the scalar communication gain between node $j$ and the node $i$. Therefore the whole multi-agent network can be represented as a a collection of n directed graphs (digraphs) with N nodes corresponding to the particular agents, and with edges representing transmission of the particular components of the vectors $\xi^i$ between the nodes, characterized by the prespecified gains $k_\nu^{ij}$. Let $G_\nu$ and $L^{G_\nu}$ represent, respectively, the digraph and the Laplacian matrix connected to the $\nu - th$ component $x_\nu$ of $x$ (or $\xi_\nu^i$ ), $\nu = 1, ..., n$.

### 4.5 Stability

One of the main issues to be addressed in relation with the recalled [9] estimation algorithm is its asymptotic stability. If $\Xi = ((\xi^1)^T, ..., (\xi^N)^T)^T$ is the vector composed of all the estimates in the agents network, the following model, with the system output measurements as the known input and the communication noises as the disturbance input, describes its global behavior:

$$\mathbf{E}: \quad \dot{\Xi} = A_\Xi \Xi + L_\Xi Y + K_\Xi \Sigma \tag{150}$$

where:

$$A_\Xi = \tilde{A} + \tilde{K}, \quad \tilde{A} = diag\{A^1 - L_1 C^1, ..., A^N - L_N C^N\}, \tag{151}$$

$$\tilde{K} = \begin{bmatrix} -\sum_{j,j\neq i} K_{1j} & K_{12} & \dots & K_{1N} \\ K_{21} & -\sum_{j,j\neq i} K_{2j} & & \dots \\ \dots & & & \\ K_{N1} & \dots & & -\sum_{j,j\neq i} K_{Nj} \end{bmatrix}, \tag{152}$$

so that

$$A_{\Xi} = \begin{bmatrix} A_E^1 & K_{12} & \dots & K_{1N} \\ K_{21} & A_E^2 & & \dots \\ \dots & & & \\ K_{N1} & \dots & & A_E^N \end{bmatrix}, \tag{153}$$

and

$$A_E^i = A^i - L_i C^i - \sum_{j,j \neq i} K_{ij}, \tag{154}$$

$$L_{\Xi} = diag\{L_1, ..., L_N\}, \tag{155}$$

$$\tilde{K}^i = \left[ K_{i1} \vdots K_{i2} \vdots \dots \vdots K_{iN} \right], \quad K_{ii} = 0, \tag{156}$$

$$K_{\Xi} = diag\{\tilde{K}^1, ..., \tilde{K}^N\}, \quad Y = ((y^1)^T, ..., (y^N)^T)^T, \tag{157}$$

$$\Sigma = (w_{11}^T, ..., w_{1N}^T, w_{21}^T, ..., w_{2N}^T, ..., w_{N1}^T, ..., w_{NN}^T)^T, \quad w_{ii} = 0. \tag{158}$$

Notice that $\tilde{K}$, the part of $A_{\Xi}$ due to consensus, is cogredient (i.e., related by permutation transformations) to $diag\{L^{G_1}, ..., L^{G_\nu}, ..., L^{G_n}\}$.

The starting assumption for all our further stability considerations is:

(A.1) the local estimators $\overline{E_i}$ are asymptotically stable, i.e., the matrices $A^{(i)} - L^{(i)}C^{(i)}$ are Hurwitz ($i = 1, ..., N$).

We shall also adopt the following basic assumption related to the network topology:

(A.2) For each $G_\nu, \nu = 1, ..., n$, there is at least one center node $\mu$ satisfying $\nu \in I_\mu^x$, i.e. each component of the state vector is updated in at least one node by the local estimator, and all the remaining nodes are accessible from this node [5]. The following analysis concerns two different case: in case A (Disjoint subsystems), the main system is separate in different and not interconnected subsystems; in case B (Overlapping subsystems), subsystems are interconnected different parts of the main system.

**A. Disjoint Subsystems**:

We shall first adopt the following simplifying assumptions, allowing a direct insight into the main structures and the related stability properties:

(A.3') the state vectors of the subsystems $S_i$ are disjoint, that is

$$I_i^x \bigcap I_j^x = 0, \quad \forall i, j = 1, ..., N, i \neq j. \tag{159}$$

(A.4') $\bigcup_{i=1}^{N} I_i^x = \{1, 2, ..., n\}$.

According to (A.3') and (A.4'), every state is estimated by one and only one local estimator, and all the states in $x$ are estimated ($\sum_{i=1}^{N} n_i = n$).

(A.5') The center nodes for all the graphs $G_\nu, \nu = 1, ..., n$, are at the same time source nodes [5].

Assumption (A.5') ensures direct distribution of the local estimates to all the remaining nodes (with no arcs entering the related node).

**Theorem 12** *[9]. Let the assumptions (A.1), (A.2), (A.3'), (A.4') and (A.5') be satisfied, and let $h_\nu^{ij} = h_{ij} \geq 0$ for $\nu = 1, ..., n$. Then the estimator* $\boldsymbol{E}$ *is asymptotically stable for any selected $h_{ij} \geq 0$ and $g_{ij} \geq 0, \quad i, j = 1, ..., N, \quad \nu = 1, ..., n$.*

The proof of Theorem 12, together with the proof of Theorem 13, can be derived from that of Theorem 14 in next chapter.

**B. Overlapping Subsystems**:

In the general case, the sets $I_i^x$ overlap, implying a more complicated structure of the matrix $A_\Xi$. In this situation, it is not possible to achieve stability of **E** by an arbitrary choice of consensus gains, like in Theorem 12. We shall concentrate here on demonstrating that for any set of stable matrices $A^{(i)} - L^{(i)}C^{(i)}$ the estimator is stabilizable, i.e. there exist such gains $K_{ij}$ which ensure asymptotic stability of **E**. In the next section we shall indicate a practical direction for choosing these gains based on estimator optimization. In order to demonstrate stabilizability of **E** by a proper choice of the consensus gains, we shall adopt a slightly more complex structure of matrices $K_{ij}$. Namely, we shall assume that $h_\nu^{ij} = h'_{ij} \geq 0$ for $\nu \in I_i^x$, and $h_\nu^{ij} = h_{ij}" \geq 0$ for $\nu \in \tilde{I}_i^x = 1, ..., n \setminus I_i^x, \quad \nu = 1, ..., n$. We shall also introduce $G_1^{ij} = diag\{g_{\nu_1^i}^{ij}, ..., g_{\nu_{n_i}^i}^{ij}\}$ and $G_2^{ij} = diag\{g_{\tilde{\nu}_1^i}^{ij}, ..., g_{\tilde{\nu}_{n-n_i}^i}^{ij}\}$ where

$$\nu_1^i < ... < \nu_{n_i}^i \in I_i^x \tag{160}$$

and

$$\tilde{\nu}_1^i < ... < \tilde{\nu}_{n-n_i}^i \in \tilde{I}_i^x \tag{161}$$

we introduce also

$$K_{ij}^{1,0} = h'_{ij} G_1^{ij} \quad and \quad K_{ij}^{2,0} = h_{ij}" G_2^{ij}. \tag{162}$$

The assumptions (A.3') and (A.4') become now:

(A.3") the state vectors of the subsystems $S_i$ are overlapping, that is

$$\bigcup_{i,j=1,...N, i \neq j} (I_i^x \bigcap I_j^x) \neq \emptyset; \tag{163}$$

(A.4") $\bigcup_{i=1}^{N} I_i^x = 1, 2, ..., n$.

Assumptions (A.3") and (A.4") imply that all the components of the state vector x in S are estimated, and there is at least one component estimated by more than one local estimator.

**Theorem 13** . *Let the assumptions (A.1),(A.2),(A.3") and (A.4") hold. Then, for any given* $h"_{ij} \geq 0$ *and* $g_\nu^{ij} \geq 0$*, it is possible to find such* $h'_{ij} \geq 0$ *that the estimator E is asymptotically stable,* $i, j = 1, ..., N, \nu = 1, ..., n$.

Proof: See ([9]).

## 4.6 Optimization

In the previous section we recalled some results which demonstrate that a stabilizing consensus scheme exists for the considered overlapping decentralized estimation problem. However, the practical problem of how to tune the consensus gains remains open. In this section we shall recall [9] a methodology for the consensus-based estimator design based on minimization of the steady-state estimation error variance. Recalling the estimator model (150), we can replace $Y$ for

$$Y = C_\Xi X + V; \tag{164}$$

where

$$X = (x^T, ..., x^T)^T, \quad C_\Xi = diag\{C_1, ...C_N\}, \tag{165}$$

and $V = ((v^{(1)})^T, ..., (v^{(N)})^T)^T$ is a white noise term with zero mean and covariance $R_V$, which can easily derived from $R$. As a result we have:

$$\dot{\Xi} = A_{\Xi}\Xi + L_{\Xi}C_{\Xi}X + L_{\Xi}V + K_{\Xi}\Sigma \tag{166}$$

Furthermore, we obtain, considering the dynamic's error of $\Xi$ in respect to $X$

$$\dot{\Xi} - \dot{X} = A_{\Xi}(\Xi - X) + \Delta AX + L_{\Xi}V + K_{\Xi}\Sigma - \Gamma_D E \tag{167}$$

where

$$\Delta A = diag\{(A_1 - A), ..., (A_N - A)\}, \quad \Gamma_D = diag\{\Gamma, ..., \Gamma\}, \quad E^T = (e^T, ..., e^T)^T, \tag{168}$$

having in mind that $\tilde{K}X = 0$ ($\tilde{K}$ is cogredient to the Laplacian matrix). If we define the new state vector of the whole system consisting of the system $S$ itself and the decentralized estimator **E** as $Z = (x^T, (\Xi - X)^T)^T$ , we obtain the state model

$$SE : \dot{Z} = A_Z Z + B\Psi = A_Z Z + B\Psi, \tag{169}$$

where

$$A_Z = \begin{bmatrix} A & 0 \\ A_{\Delta} & A_{\Xi} \end{bmatrix}, \quad A_{\Delta} = col\{(A_1 - A), ..., (A_N - A)\} \tag{170}$$

(col{.} denotes the block-column matrix composed of the listed elements),

$$B = diag\{-\Gamma_{\Delta}, -L_{\Xi}, K_{\Xi}\}, \quad \Gamma_{\Delta} = col\{\Gamma, ..., \Gamma\}, \quad \Psi = (e^T, V^T, \Sigma^T)^T. \tag{171}$$

Obviously, **SE** represent a stochastic system with the white noise $\Psi$ as a stochastic input. If we assume that $A_Z$ is Hurwitz, the steady-state covariance $P$ of $Z$ is defined by the positive semi-definite solution of the following Lyapunov equation

$$A_Z P + PA_Z^T + BR_{\Psi}B^T = 0 \tag{172}$$

where $R_{\Psi} = cov\Psi$, which can be easily constructed starting from the definition of the constituent terms of $\Psi$ under the assumption that these terms are independent. If we define vector $\theta$ containing all the unknown parameters of the consensus scheme in **E**, we can formulate the following optimization problem:

$$\min_{\theta} J = \min_{\theta}\{TrP\} \tag{173}$$

Solution to the problem (173), $\theta^* = Arg\min_{\theta} J$, will define the consensus parameters in the recalled [9] "optimized" estimation scheme. It is difficult to draw any conclusion about the nature of the above optimization problem in the general case. As it will be shown in the experimental part of the chapter, numerical optimization procedures can efficiently be applied to finding $\theta^*$. Having in mind the adopted structure of the consensus matrix gains $K_{ij}$ , we shall assume that the above optimization procedure encompasses, as elements of the parameter vector $\theta$, only scalar weighting parameters $h_{ij}$ (or $h'_{ij}$ and $h_{ij}$" ). In general, nonzero parameters $g^{ij}_{\nu}, \nu = 1, ..., N$, which correspond to the indices $\nu \in I^x_j$ , ensure proper propagation of the local state estimates through the network; these parameters will be assumed to be specified a priori. It has been found to be beneficial to adopt that $g^{ij}_{\nu}$ is equal to the $\nu - th$ diagonal element of the inverse of the local covariance matrix $P^{(j)}$ in (148).

**Remark 5** . *One of important and interesting issues is related to the influence of the consensus strategy on measurement noise suppression resulting from implicit ensemble averaging, in the case when the number of nodes increases. The basic problem of consensus averaging and asymptotic agreement has been studied in [21]. In the case of the recalled [9] algorithm, it is possible to show, assuming simplified structures of the communication gains, that the elimination of measurement noise influence can be achieved when the number of nodes tends to infinity provided the network possesses a sufficient connectivity degree [9]. The achieved performance depends on the absolute value of the real part of the eigenvalues of the Laplacians (characterizing the underlying scalar digraphs) which are closest to the imaginary axis (in analogy with the Fiedlers algebraic complexity in the case of undirected graphs). It is interesting to notice that the strong connectivity alone does not ensure complete asymptotic denoising [9].*

## 4.7 Numerical Illustrations

Some characteristic properties of the recalled [9] estimation scheme will be illustrated by simple examples. We shall assume that the state of **S** is represented by a fourth order model with

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{174}$$

where

| | $h_{12}$ | $h_{21}$ | $J$ |
|---|---|---|---|
| $R_1$=1 | 1521.9 | 855.5 | 1.9819 |
| $R_1$=10 | 898.4 | 49.56 | 2.0102 |
| $R_1$=100 | 170.2 | 1.927 | 2.0109 |
| $R_1$=1000 | 110.2 | 0.026 | 2.0110 |

Figure 20: Optimization Results for Different Measurement Noise Levels

$$A_{11} = \begin{bmatrix} -1 & 0 \\ -1 & 2 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} 0 & 0.1 \\ 0.1 & 0 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 0 & 1 \\ -3 & -5 \end{bmatrix}. \tag{175}$$

with $\Gamma = Q = I$.

We shall consider several cases differing by the information available to the agents and their a priori knowledge about the system.

*Case A.* Agent 1 observes the states using

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}. \tag{176}$$

with noise variance $R_1$ and Agent 2 using

$$C_2 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}. \tag{177}$$

with noise variance $R_2$; both agents possess the knowledge of the entire state model, i.e. the whole matrix $A$ is used within the local estimators. The communication noise is characterized by variances $W_{12} = 0.01$ and $W_{21} = 0.01$. It is assumed that the consensus gains are defined by $K_{12} = h_{12}G_{12}$ and $K_{21} = h_{21}G_{21}$, where $G_{12} = diag\{P^{(2)-1}\}$ and $G_{21} = diag\{P^{(1)-1}\}$, and $h_{12} \geq 0$ and $h_{21} \geq 0$ represent unknown parameters to be determined by optimization; $P^{(1)}$ and $P^{(2)}$ are the estimation error covariances of the local Kalman filters. Table (20) shows the results obtained by the recalled [9] optimization procedure, for $R_2 = 1$ and different values of $R_1$.

The results show high efficiency of the recalled [9] consensus scheme: the criterion values are only slightly deteriorated by the increase of $R_1$. It is also evident that both gains $h_{12}$ and $h_{21}$ are higher for lower measurement noise levels; however, $h_{21}$ decreases much more rapidly and for high values of $R_1$ is close to zero, having in mind that the quality of the first local estimator $\overline{E_1}$ becomes low.

*Case B*. We consider three agents, the first two being the same as above (with $R_1 = R_2 = 1$), while the third observes the system using

$$C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{178}$$

and

$$R_3 = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}, \tag{179}$$

Optimization provides now six parameters, (two per agent); the obtained results are: $h_{12} = 0.155, h_{13} = 0.355, h_{21} = 0.460, h_{23} = 0.300, h_{31} = 0$ and $h_{32} = 0$, taking, as above, diagonal matrices $G_{ij}$ equal to the main diagonals of the corresponding local estimation error covariance inverses. The scheme behaves again as predicted: agent 3, with the globally optimal Kalman estimator, does not need any help, so that the weights of the edges leading to it are equal to zero. On the other hand, Agents 1 and 2 take the more accurate estimates obtained from Agent 3 with higher gains.

*Case C*. In this case we consider Agents 1 and 2 observing the same outputs as above, but we assume that the agent resources are such that the local estimators are based on the second order local subsystem models with the state matrices

$$A_1 = \begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & A_{22} \end{bmatrix}, \tag{180}$$

(see the notation in (149) and the corresponding two-dimensional stochastic inputs. The optimization provides now $h_{12} = 0.6311$ and $h_{21} = 0.8088$, with $J = 2.0271$. The consensus scheme succeeds to efficiently compensate not only for the missing measurements, but also for the modelling imprecision. Figure (21) depicts the form of the criterion function: the function is convex in the considered two-dimensional case.

*Case D*. In this case we consider two agents in two situations: in the first, the subsystem models are disjoint as in the case C, while in the second the subsystem models are of third order, and are, obviously, overlapping. We assume now that $A$ in $S$ is composed of

$$A_{11} = \begin{bmatrix} 1 & 1 \\ -1 & 0.2 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} 0.1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} -0.1 & 1 \\ -0.3 & -5 \end{bmatrix}, \tag{181}$$

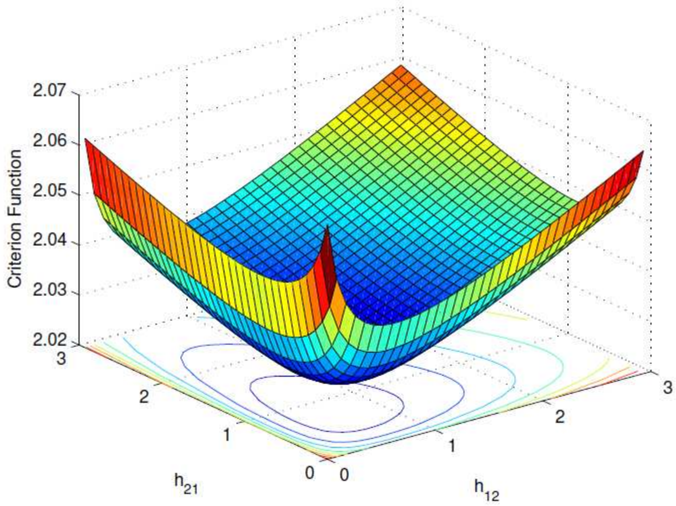



Figure 21: Criterion function

and that in situation I Agent 1 utilizes $A_{11}$, and Agent 2 utilizes $A_{22}$, as in Case C. In situation II, we assume overlapping subsystems, so that Agent 1 utilizes

$$A_1 = \begin{bmatrix} 1 & 1 & 0 & \vdots & 0 \\ -1 & 0.2 & -1 & \vdots & 0 \\ 0.1 & 0 & -0.1 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \vdots & 0 \end{bmatrix}, \tag{182}$$

and Agent 2

$$A_2 = \begin{bmatrix} 0 & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & 0.2 & -1 & 0 \\ 0 & \vdots & 0 & -0.1 & 1 \\ 0 & \vdots & 1 & -0.3 & -5 \end{bmatrix}, \tag{183}$$

With the same noise levels as above, we obtained for situation I $h_{12} = 0.001$ and $h_{21} = 0.1791$, with $J = 35.43$, and for situation II $h_{12} = 2.5421$ and $h_{21} = 8.1781$, with $J = 7.8621$. This example shows possible advantages of overlapping decompositions with respect to the disjoint ones.

### 4.8 Conclusions

In this chapter a recent algorithm for solving the problem of overlapping decentralized state estimation of linear complex stochastic systems is recalled [9] based on combining a set of local Kalman filters with a consensus scheme. Considering local decision makers as multiple agents in a network with the topology induced by local communications, it has been proved that the recalled [9] algorithm can always be asymptotically stabilized by a proper choice of the consensus gains under general conditions related to the overlapping subsystems (resulting from the a priori knowledge of each agent) and the network topology. This chapter also recalls an optimization algorithm for obtaining the network gains by minimizing the total steady-state mean-square estimation error. A set of simple examples illustrate some specific properties of the recalled [9] algorithm, showing that it can represent a reliable tool for practical implementations.

## References


[1] D. P. Bertsekas, J. .N Tsitsiklis " Parallel and Distributed Computation: Numerical Methods". Englewood Cliffs, NJ:Prentice Hall, 1989.

[2] P. Ferguson and J. How, "Decentralized estimation algorithms for formation flying spacecraft". Proc. AIAA Guid. Nav. Control Conf, 2003.

[3] H.R. Hashimpour, S. Roy and A.J. Laub "Decentralized structures for parallel Kalman filtering", IEEE Trans. Aut. Control, vol. 33, pp 88-93, 1988

[4] M. Ikeda, D.D. Šiljak, D. E. White, " Decentralized Control with Overlapping Information Sets". Journal of optimization theory and applications: Vol.34, No.2. June 1981.

[5] Z. Lin, Coupled Dynamic Systems: From Structure Towards Stability and Stabilizability, Ph.D. dissertation, University of Toronto, Toronto, ON, Canada, 2006.

[6] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters". Proc. 2005 IEEE Conf. Decision and Control CDC'05.

[7] W. Ren, R.W. Beard and D.B. Kingston "Multi-agent Kalman consensus with relative uncertainty". Proc. 2005 American Control Conference ACC'05.

[8] D. Simon. Optimal State Estimation: Kalman, H , and Nonlinear Approaches. Wiley-Interscience, New Jersey, 2006.

[9] S.S. Stankovic', M.S. Stankovic'and D.M. Stipanovic' "Consensus based overlapping decentralized estimator". IEEE Transactions on Automatic Control, 54, 410-415, 2009.

[10] S.S. Stankovic', M.S. Stankovic'and D.M. Stipanovic' "Consensus based overlapping decentralized estimation with missing observations and communication faults". Automatica, 45, 6, 1397-1406, 2009.

[11] D.D. Šiljak,'"Large Scale Dynamic Systems: Stability and Structure.". Amsterdam, The Netherlands: North Holland, 1978.

[12] C.W. Sanders, E.C. Tacker and T.D. Linton "Specific structures for large scale state estimation algorithms having information exchange". IEEE Transactions on Automatic Control, 23, pp. 255-260, 1978.

[13] L. Moreau "Stability of multiagent systems with time-dependent communication links". IEEE Transactions on Automatic Control, vol. 50, pp 169- 182, 2005.

[14] R. Olfati-Saber, J.A. Fax, R.M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems". IEEE Proc., vol. 95, 1, pp. 215–233, 2007.

[15] W. Ren and R.W. Beard "Consensus seeking in multi-agent systems using dynamically changing interaction topologies". IEEE Transactions on Automatic Control, vol. 50, pp 655-661, 2005.

[16] S. Simani, C. Fantuzzi and R.J. Patton, "Model-based fault diagnosis in dynamic systems using identification techniques", Advances in Industrial Control, Springer–Verlag, London, 2002.

[17] A. Speranzon, C. Fischione and K. H. Johansson, " Distributed and collaborative estimation over wireless sensor networks". Proc. 2006 IEEE Conf. Decision and Control CDC'06.

[18] F. Xiao, L. Wang " Consensus problems for high-dimensional multi-agent systems". IET Control Theory & Applications, vol. 1, No.3. May 2007.

# 5 Decentralized Estimation in Systems with Unknown Inputs.

## 5.1 Preliminaries

In the previous Chapter we have seen how it is possible to design a decentralized state estimation when a system is affect by stochastic disturbance. But what can we do if disturbance is also deterministic input? As it possible to reconstruct the deterministic disturbances, the answer to these questions is given in this Chapter, in which a new estimation algorithm is proposed. It is based on the technique of chapter 4, combined with the theory of "strong observability", and related "robust observation" methodologies, that we have seen in chapter 2.

## 5.2 Introduction

As we have seen in the previous chapter, decentralized state estimation addresses the problem of reconstructing the state of a large-scale system under the key requirement that the system can be modelled as an interconnection of subsystems, and that each subsystem have a decision maker (an intelligent "agent") associated with it. The agents might have access to different informations of **local** nature, such as local output variables, properties and models of the local subsystems, the characteristics of the measurements sensors and of the communication channels between the agents.

Earlier works in the area started from the parallel implementation of the globally optimal Kalman filter [8]. Attempts to provide an insight into basic principles and paradigms

In the broad area of multi-agent systems, including coordinated control of multi-vehicle systems, sensor networks, and many other important applications (e.g., see [8, 15, 18, 20]), a common methodology, the **consensus** strategy, is becoming more and more popular as a beneficial and extremely appropriate paradigm for distributed signal-processing or decision making [14, 21].

The problem of state observation for linear time-invariant systems with unknown inputs has been widely studied during the last two decades. As we saw in chapter 2, it was shown that, under the condition of observability (see, e.g., [9, 10]) and a special additional requirement about the absence of invariant zeroes between output and unknown inputs (both conditions together were shown to be necessary, the first, necessary and sufficient, the second, and denoted in [9] as "**strong observability**"), a "decoupling" state transformation can be made such that the observation error dynamics in the transformed state coordinates is not contaminated by the unknown inputs. Then, a reduced-order linear observer can be designed which is capable of reconstructing the overall state vector. Unknown input observers have been widely used in the framework of fault detection and isolation [19]. It was discovered that sliding mode observers allow to reconstruct accurately the unknown input together with the system state, which is an important requirement in FDI schemes [11, 6]. In [1, 2, 14], second-order sliding-mode observers were suggested that allows to reconstruct the unknown input in finite time.

However, none of the above approaches is aimed at establishing any type of collaboration between local estimators in the overlapping decentralized estimation problem. Only recently overlapping decentralized Kalman filters have been put together in a multi-agent network on the basis of a consensus strategy, as we saw in chapter 4.

In application of consensus to decentralized state observation problems, $N$ agents possessing different plant informations provide independent state estimates $\hat{x}^i$ which will eventually align to the same estimate as time grows to infinity, thanks to appropriate communication protocols. The stabilizability of the collective observation error dynamics is proven by arguments heavily relying on graph theory, able to properly capture the main features of large-scale inter-agent communication.

In this chapter we extend the class of plants dealt with in chapter 4 by encompassing deterministic **unknown inputs** acting on the linear system. Furthermore, we not only aim to reconstruct the system state but we shall provide as well the approximate reconstruction of the unknown inputs acting on the system. We address the state observation problems from the perspective of consensus-based decentralized Kalman estimation. By following a similar approach as in last chapter 4, we also develop an optimization-based procedure for computing the consensus gain parameters. Finally, we solve the unknown-input reconstruction problem by means of an appropriate proportional-integral (PI) observer. As compared with the method used in [1, 2, 14], the use of a PI observer allows to relax the amount of prior information required about the unknown input.

This chapter is structured as follows. The problem formulation, and the combined state-output transformation that play a fundamental role in the present approach, are discussed in the Section 5.3. The proposed consensus-based decentralized state estimation procedure and a method for computing optimal values for the consensus gains are explained in the Section 5.4. The problem of the reconstruction of the unknown inputs is addressed in the Section 5.5. Section 5.6 reports some simulation results, and in Section 5.7 some concluding remarks are given along with possible lines of research for future related activities.

### 5.3 Problem Formulation

Let us consider the following class of continuous-time linear stochastic systems with unknown inputs:

$$\begin{cases} \dot{x} &= Ax + \Gamma e + Bu \\ y &= Cx + v \end{cases} \tag{184}$$

where $x = (x_1, ..., x_n)^T$ is the state vector, $y = (y_1, ..., y_p)^T$ is the output vector, $u = (u_1, ..., u_m)^T$, $m < p$, is an **unknown** input vector, $e = (e_1, ..., e_q)^T$ and $v = (v_1, ..., v_p)^T$ are stochastic noise signals, and $A, \Gamma, C$ and $B$ are constant matrices of appropriate dimension.

It is assumed that $e(t)$ and $v(t)$ are mutually independent stochastic gaussian processes, with zero-mean and covariance matrices as follows:

$$E\{e(t)e(\tau)^T\} = Q\delta(t-\tau), \qquad E\{v(t)v(\tau)^T\} = R\delta(t-\tau). \tag{185}$$

In the decentralized scheme for state vector estimation, $N$ autonomous agents have the goal to generate their estimates of the state vector $x(t)$ on the basis of **locally available measurements** (each agent can measure only a subset of the elements of the output vector $y$). More precisely we assume that the $i$-th agent can measure, with some noise, the $p_i$-dimensional vector $y^{(i)}$, with $p_i > m$, containing those output variables with indices specified by the so-called agent's **output index set** $I_y^i$, which is defined as follows:

$$I_i^y = \{l_1^i, ..., l_{p_i}^i\}, \qquad l_1^i, ..., l_{p_i}^i \in \{1, ..., p\}, \quad l_1^i < ... < l_{p_i}^i, \quad p_i \le p. \tag{186}$$

We also define, like in chapter 4, **state index set** $I_i^x$ of $i$-th agent as follows.

$$I_i^x = \{l_1^i, ..., l_{n_i}^i\}, \qquad l_1^i, ..., l_{n_i}^i \in \{1, ..., n\}, \quad l_1^i < ... < l_{n_i}^i, \quad n_i \le n \tag{187}$$

Vector $I_i^x$ denotes the indices of the original state variables $x_i$'s that explicitly affect the local measurement output equation.

The **local system model** available to the $i-th$ agent is defined as follows

$$\begin{array}{rcl} \dot{x}^{(i)} &=& A^i x^{(i)} + \Gamma^i e + B^i u \\ y^{(i)} &=& C^i x^{(i)} + v^{(i)} \end{array}, \quad i = 1, 2, ..., N \tag{188}$$

where $x^{(i)}$ is a vector of dimension $n_i \le n$, $y^{(i)}$ was previously defined as the local output vector with dimension $p_i \le p$, and $A^i$, $B^i$, $C^i$, $\Gamma^i$ are constant matrices of appropriate dimension. Clearly, the matrices $C^i$, that define the local output measurement equations, contain a subset of the entries of the "complete" output matrix $C$. More precisely, matric $C^i$ contains the elements of $C$ having the row and column indexes specified by the set $I_i^y \times I_i^x$. Similarly, matrix $A^i$ contains those elements of the matrix $A$ with the row and column indexes taken in the set $I_i^x \times I_i^x$. Matrices $\Gamma^i$ and $B^i$ contain a subset of the rows of the complete matrices $\Gamma$ and $B$, with the selected row indexes specified in the state index set $I_i^x$. Vector $v^{(i)}$ represents a "reduced-order" measurement noise vector having zero mean and covariance matrix

$$E\left[v^{(i)}(t)\left(v^{(i)}(\tau)\right)^T\right] = R^{(i)}\delta(t-\tau) \tag{189}$$

Matrix $R^{(i)}$ contains the elements of the original matrix $R$ having the row and column indexes taken in the set $I_i^y \times I_i^y$. Note that the noise vectors $v^{(i)}$ correspond to different, uncorrelated, realizations of the statistical process.

System (188) defines $N$ **overlapping local subsystems** of the overall plant (184). Note that the decomposition of (184) into the overlapping subsystems (188) does not rely on any inherent decomposition of the original plant matrices $A$ and $B$ in (184). Hence, the agents need not a complete knowledge of the overall system dynamics, but only of a part of it.

Concerning the local system models dynamics ($i = 1, 2, ..., N$) let us make the following assumptions:

**(A.1)** $rank\left(C^iB^i\right) = rankB^i$.

**(A.2)** The matrix triplets $\left(A^i, C^i, B^i\right)$ are **strongly observable** [10]

The notion of strong observability, recalled in chapter 2, has been introduced more than thirty years ago [10, 9] in the framework of the unknown-input observers theory. Recently it has been exploited to design robust observers based on the high-order sliding mode approach [1]. It has been shown in [10] that the following statements 1 and 2 are equivalent

1. The triple $\left(A^i, C^i, B^i\right)$ is strongly observable.
2. The invariant zeros of the triple $\left(A^i, C^i, B^i\right)$ have negative real part.

It is known [3, 2] that if condition A1 is satisfied then it can be systematically found a state coordinates transformation together with an output coordinates change which decouple the unknown input $u$ from a certain subsystem in the new coordinates. Such a transformation is outlined below.

For the generic matrix $J \in \mathbb{R}^{n_r \times n_c}$ with $rank(J) = r$, we define $J^{\perp} \in \mathbb{R}^{n_r - r \times n_r}$ as a matrix such that $J^{\perp}J = 0$ and $rank(J)^{\perp} = n_r - r$. Matrix $J^{\perp}$ always exists and, furthermore, it is not unique for a given $J$.

Consider the following transformation matrices $T^i$ and $U^i$ $(i = 1, 2, ...N)$:

$$T^i = \begin{bmatrix} B^{i\perp} \\ \left(C^i B^i\right)^+ C^i \end{bmatrix}, \qquad U^i = \begin{bmatrix} \left(C^i B^i\right)^{\perp} \\ \left(C^i B^i\right)^+ \end{bmatrix} = \begin{bmatrix} U_1^i \\ U_2^i \end{bmatrix} \tag{190}$$

$$\left(C^i B^i\right)^+ = \left[\left(C^i B^i\right)^T \left(C^i B^i\right)\right]^{-1} \left(C^i B^i\right)^T \tag{191}$$

and the transformed state and output vectors

$$\bar{x}^{(i)} = T^i x^{(i)}, \qquad \bar{y}^{(i)} = U^i y^{(i)} \tag{192}$$

The transformed dynamics in the new state and output coordinates are

$$\begin{array}{rcl} \dot{\bar{x}}^{(i)} & = & T^i A^i (T^i)^{-1} \bar{x}^{(i)} + T^i \Gamma^i e + T^i B^i u \\ \bar{y}^{(i)} & = & U^i C^i (T^i)^{-1} \bar{x}^{(i)} + U^i v^{(i)} \end{array}, \quad i = 1, 2, ..., N \tag{193}$$

Consider the following partitions of vectors $\bar{x}^{(i)}$ and $\bar{y}^{(i)}$

$$\bar{x}^{(i)} = \begin{bmatrix} \bar{x}_1^{(i)} \\ \bar{x}_2^{(i)} \end{bmatrix}, \qquad \bar{x}_1^{(i)} \in R^{n_i - m} \qquad \bar{x}_2^{(i)} \in R^m \tag{194}$$

$$\bar{y}^{(i)} = \begin{bmatrix} \bar{y}_1^{(i)} \\ \bar{y}_2^{(i)} \end{bmatrix} \qquad \bar{y}_1^{(i)} \in R^{p_i - m} \qquad \bar{y}_2^{(i)} \in R^m \tag{195}$$

The partitioned transformed vectors are given by

$$\bar{x}_1^{(i)} = B^{i\perp} x^{(i)}, \qquad \bar{x}_2^{(i)} = \left(C^i B^i\right)^+ C^i x^{(i)} \tag{196}$$

$$\bar{y}_1^{(i)} = \left(C^i B^i\right)^{\perp} y^{(i)} \qquad \bar{y}_2^{(i)} = \left(C^i B^i\right)^+ y^{(i)} \tag{197}$$

After simple algebraic manipulations the **transformed local system models** in the new coordinates can be expanded in the form:

$$\begin{array}{rcl} \dot{\bar{x}}_1^{(i)} & = & \bar{A}_{11}^i \bar{x}_1^{(i)} + \bar{A}_{12}^i \bar{x}_2^{(i)} + B^{i\perp} \Gamma^i e \\ \dot{\bar{x}}_2^{(i)} & = & \bar{A}_{21}^i \bar{x}_1^{(i)} + \bar{A}_{22}^i \bar{x}_2^{(i)} + \left(C^i B^i\right)^+ C^i \Gamma^i e + u \\ \bar{y}_1^{(i)} & = & \bar{C}_1^i \bar{x}_1^{(i)} + \bar{v}_1^{(i)} \\ \bar{y}_2^{(i)} & = & \bar{x}_2^{(i)} + \bar{v}_2^{(i)} \end{array} \tag{198}$$

with implicit definition of matrices $\bar{A}_{11}^i$, ..., $\bar{A}_{22}^i$ and with

$$\begin{bmatrix} \bar{v}_1^{(i)} \\ \bar{v}_2^{(i)} \end{bmatrix} = \bar{v}^{(i)} = U^i v^{(i)} = \begin{bmatrix} U_1^i \\ U_2^i \end{bmatrix} v^{(i)} \tag{199}$$

Signals $\overline{v}_1^{(i)}$ and $\overline{v}_2^{(i)}$ are then stochastic terms with zero-mean and, according to (189), the covariances

$$\begin{array}{l} E(\overline{v}_1^{(i)}(t)\overline{v}_1^{(i)}(\tau)^T) = [U_1^i R^{(i)}(U_1^i)^T]\delta(t-\tau) \\ E(\overline{v}_2^{(i)}(t)\overline{v}_1^{(i)}(\tau)^T) = [U_2^i R^{(i)}(U_2^i)^T]\delta(t-\tau) \end{array} \tag{200}$$

It turns out that the triple $\left(A^i, C^i, B^i\right)$ is strongly observable if, and only if, **the pair $(\bar{A}_{11}^i, \bar{C}_1^i)$ is observable** [10, 9]. In light of the Assumption A2, this property, that can be also understood in terms of a simplified algebraic test to check the strong observability of a matrix triple, opens the way to design stable observers for the transformed dynamics (194)-(198). The strong observability assumption for the local system models implies the requirements that $p_i > m$ and $n_i > m$ $(i = 1, 2, ..., N)$, otherwise the special form (194)-(198) for the transformed dynamics is no longer obtained and the observer design procedure that is going to be illustrated becomes unfeasible. The less strict condition $p_i \leq m$ could be possibly allowed, for some $i$, under the additional requirement that the corresponding matrix $\bar{A}_{11}^i$ is Hurwitz.

### 5.4 Consensus based state observation

A peculiarity of the transformed local system models (198) is that **vectors $\bar{x}_2^{(i)}$ are directly available to the $i$-th agent** through the (noisy and stochastical) measurement of the transformed output vector components $\bar{y}_2^i$.

Hence, once the expected value of vector $\bar{x}_1^{(i)}$ has been made available through appropriate observation techniques, the expected value of the original local subsystem state $x^{(i)}$ can be reconstructed by the following relationships

$$E(x^{(i)}) = (T^i)^{-1} \begin{bmatrix} \bar{x}_1^{(i)} \\ \bar{x}_2^{(i)} \end{bmatrix} \tag{201}$$

Note that vector $\bar{x}_1^{(i)}$ is of reduced dimension $(n_i - m)$ as compared to the original state vector $x$, and, furthermore, an important additional peculiarity of $\bar{x}_1^{(i)}$ is that the resulting dynamics are not affected by the unknown input vector.

It can be extracted from (198) the subset of equations involving vector $\bar{x}_1^{(i)}$, that can be simply manipulated as follows

$$\begin{array}{rcl} \dot{\bar{x}}_1^{(i)} & = & \bar{A}_{11}^i \bar{x}_1^{(i)} + \bar{A}_{12}^i \bar{y}_2^{(i)} + D^i \mu^{(i)} \\ \bar{y}_1^{(i)} & = & \bar{C}_1^i \bar{x}_1^{(i)} + \overline{v}_1^{(i)} \end{array} \tag{202}$$

where

$$D^i = [B^{i\perp}\Gamma^i \quad - \bar{A}_{12}^i], \qquad \mu^{(i)} = \begin{bmatrix} e \\ \overline{v}_2^{(i)} \end{bmatrix} \tag{203}$$

By taking into account the statistical properties of signals $e$ and $\overline{v}_2^{(i)}$, it can be shown that $\mu^{(i)}$ is a zero-mean signal with the covariance matrix $V^{(i)}\delta(t-\tau)$, with

$$V^{(i)} = \begin{bmatrix} Q & 0 \\ 0 & U_2^i R^{(i)} (U_2^i) \end{bmatrix} \tag{204}$$

By considering (190)–(199), equation (202) can be further rewritten as follows

$$\begin{array}{rcl} \dot{\bar{x}}_1^{(i)} & = & \bar{A}_{11}^i \bar{x}_1^{(i)} + \bar{A}_{12}^i (C^i B^i)^+ y^{(i)} + D^i \mu^{(i)} \\ \bar{y}_1^{(i)} & = & (C^i B^i)^\perp y^{(i)} = \bar{C}_1^i \bar{x}_1^{(i)} + \overline{v}_1^{(i)} \end{array} \tag{205}$$

We propose a decentralized state estimation algorithm based on the *consensus scheme* specifying communications between the agents.

Let us consider, in analogy with (190)-(197), the transformed state vectors of the **original** system

$$\bar{x} = Tx, \qquad T = \begin{bmatrix} B^\perp \\ (CB)^+ C \end{bmatrix} \tag{206}$$

and, in particular, the subcomponent $\bar{x}_1$ of $\bar{x}$

$$\bar{x}_1 = B^\perp x, \qquad \bar{x}_1 \in R^{n-m} \tag{207}$$

In association with the transformed local system models (198), it is possible to define $N$ additional "transformed state index set vectors" $\bar{I}_x^i$ ($i = 1, 2, ..., N$) for the local transformed subsystems which contains $n_i - m$ ordered elements taken in the set $\{1, 2, ..., n-m\}$. The transformed state index set vectors specifies the link between any element of vector $\bar{x}_1^{(i)}$ and the corresponding element in the vector $\bar{x}_1$. In other words, vector $\bar{x}_1^{(i)}$ can be associated to a selected subset of the entries of the "overall" transformed vector $\bar{x}_1$.

We assume that the communication between the $i$-th and $j$-th agents is corrupted by an additive noise $w_{ij}$. Such a noise is supposed to be white and uncorrelated from $e$ and $v$, with zero mean, and with covariance

$$E\{w_{ij}(t) w_{ij}(\tau)^T\} = W_{ij}\delta(t-\tau) \tag{208}$$

Every agents build its own estimate $\xi^i$ of $\bar{x}_1$ according to the next collective consensus dynamics

$$\dot{\xi}^i = A_{11}^i \xi^i + A_{12}^i (C^i B^i)^+ y^{(i)} + \textstyle\sum_{j=1, j\neq i}^N K_{ij}(\tilde{\xi}^{i,j} - \xi^i) + L_i \left[ (C^i B^i)^\perp y^{(i)} - C_1^i \xi^i \right] \tag{209}$$

Clearly, according to (207), the dimension of vector $\xi^i$ is $n-m$. Matrices $A_{11}^i$, $A_{12}^i$ and $C_1^i$ are obtained by "augmenting" the matrices $\bar{A}_{11}^i$, $\bar{A}_{12}^i$ and $\bar{C}_1^i$ with appropriately located zero elements. Vector $\bar{I}_x^i$ specifies the locations where the zero elements have to be added, in accordance with the following description of the structure of matrices $A_{11}^i$, $A_{12}^i$, $C_1^i$:

$\cdot\ A_{11}^i \in R^{(n-m)\times(n-m)}$ and whose elements having the row and column indexes in the set $\bar{I}_x^i \times \bar{I}_x^i$ are equal to the corresponding entries of $\bar{A}_{11}^i$, while the remaining elements are zeros,

$\cdot\ A_{12}^i \in R^{(n-m)\times(m)}$ and whose elements having the row index in the set $\bar{I}_x^i$ are equal to the corresponding entries of $\bar{A}_{12}^i$, while the remaining elements are zeros,

$\cdot\ C_1^i \in R^{(p_i-m)\times(n-m)}$ and whose elements having the column index in the set $\bar{I}_x^i$ are equal to the corresponding entries of $\bar{C}_1^i$, while the remaining elements are zeros,

Furthermore:

$\cdot\ K_{ij} \in R^{(n-m)\times(n-m)}$ are constant consensus gain matrices,

$\cdot\ \tilde{\xi}^{i,j}$ is the noisy estimate $\xi^j$ communicated by the $j$-th agent ($j \neq i$), i.e., $\tilde{\xi}^{i,j} = \xi^j + w_{ij}$.

$\cdot\ L_i \in R^{(n-m)\times(p_i)}$ and whose elements having the row index in the set $\bar{I}_x^i$ are equal to the corresponding entries of the steady-gain matrix of the local Kalman filter for the system (205), while the remaining elements are zeros,

The decentralized observer dynamics (209) contains three parts:

- the nominal plant model $(A_{11}^i\xi^i + A_{12}^i(C^iB^i)^+y^{(i)})$,
- the consensus term $\sum_{j=1,j\neq i}^N K_{ij}(\tilde{\xi}^{i,j} - \xi^i)$, and
- the output error injection $L_i\left[(C^iB^i)^{\perp}y^{(i)} - C_1^i\xi^i\right]$.

It shall be proven that a stabilizing consensus scheme can be implemented which means that the expected values of all agents' state would eventually align to a common, correct, estimate, i.e. $E(\xi^i) = \xi^*$, $i = 1, 2, ..., N$. In other words, it must be demonstrated the stabilizability of the collective observation error dynamics by proper choice of the consensus gain matrices $K_{ij}$.

We select the consensus gain matrices to be **diagonal**:

$$K_{ij} = \begin{bmatrix} k_1^{ij} & 0 & 0 & 0 \\ 0 & k_2^{ij} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & k_{n-m}^{ij} \end{bmatrix}, \qquad i, j = 1, 2, ..., N. \quad i \neq j \tag{210}$$

The generic element $k_\nu^{ij}$ is further expressed as follows

$$k_\nu^{ij} = h_\nu^{ij} g_\nu^j, \qquad h_\nu^{ij} \geq 0, \qquad g_\nu^j \geq 0, \qquad \nu = 1, 2, ..., n-m \tag{211}$$

where $h_\nu^{ij}$ reflects properties of the involved agent communication channel, and $g_\nu^j$ reflects structural properties of the $j$-th local model (188) and the corresponding uncertainty in the local estimate $\xi^j$. If agents $i$ and $j$ do not communicate, then the corresponding elements $h_\nu^{ij}$ are set to zero.

For each component $\nu = 1, 2, ..., n-m$ we define the directed graph (digraph) $G_\nu$, with $N$ nodes (corresponding to the agents) and edges having the gains $k_\nu^{ij}$, specifying transmission of particular components of the vector $\xi^i$ between the nodes. Let $L^{G_\nu}$ represent the Laplacian of $G_\nu$, i.e., $L^{G_\nu} = [L_{ij}^{G_\nu}]$, with

$$L_{ij}^{G_\nu} = \begin{cases} k_\nu^{ij} & i \neq j \\ -\sum_{j,j\neq i} k_\nu^{ij} & i = j \end{cases}, \qquad i,j = 1,...N. \tag{212}$$

The additional basic assumptions are:

**(A.3)** the pair $\bar{A}_{11}^i, \bar{C}_1^i$ is observable, and the pair $\bar{A}_{11}^i, D^i V^{i^{1/2}}$ is stabilizable $\forall i = 1, ..., N$;

**(A.4)** For each $\mathsf{G}_\nu, \nu = 1, ...n-m$, there is at least one center node $\mu$ (from which every node is reachable), satisfying $\nu \in \bar{I}_x^\mu$.

**(A.5)** $\bigcup_{i=1}^N \bar{I}_x^i = \{1, ..., n-m\}$;

**(A.6)** $\bigcup_{i,j=1,...,N; i\neq j} (\bar{I}_x^i \bigcap \bar{I}_x^j) \neq \emptyset$;

Assumptions (A.1)-(A.3) involves structural restrictions about the local subsystem dynamics (in particular, Assumption A.3 guarantees that the local Kalman estimators are asymptotically stable [8]).

Assumption (A.4) specifies the constraint on the agents' communication topology and it is often referred to as "quasi-strong" connectivity [5]. Assumptions (A.5) and (A.6) imply that all the components of $\bar{x}_1$ are reconstructed by at least one agent, and that there is at least one component estimated by more than one local estimator.

The next theorem, whose rationale and proof are similar to that of Theorem 1 in [9], is in force for the transformed collective observation dynamics in question.

**Theorem 14** *Consider the decentralized estimation algorithm (209). Under the given assumptions (A.1)-(A.6) there exist consensus gain matrices $K_{ij}$ of the form (210)-(211) such that the expected values of all agents' state would eventually align to a common estimate which is the* *expected value of the transformed vector component $\bar{x}_1$ according to*

$$E(\xi^i) = E(\bar{x}_1), \qquad i = 1, 2, ..., N \tag{213}$$

**Proof of Theorem 14**. See the Appendix.

The above theorem is a "feasibility" result that does not offer any guidelines about how to tune the consensus gain matrices.

In order to fully exploit the properties of the the decentralized estimation it is mandatory to select them in some optimal way by taking into account the properties of the deterministic and stochastic parts of the involved subsystems and of the noisy agents communication links. Consider the vector $Z = \Xi - X_1$, $Z \in \mathbb{R}^{(n-m)N}$, where $\Xi$ and $X_1$ are defined as follows

$$\Xi = [(\xi^1)^T, ..., ((\xi^N)^T]^T, \qquad \Xi \in \mathbb{R}^{(n-m)N} \tag{214}$$

$$X_1 = [\bar{x}_1^T \ \bar{x}_1^T \ ... \ \bar{x}_1^T]^T \qquad X_1 \in \mathbb{R}^{(n-m)N} \tag{215}$$

and the associated dynamics, which takes the form

$$\dot{Z} = \Phi Z + B_Z \Theta \tag{216}$$

with the characteristic matrix $\Phi = \{\Phi_{ij}\}$ having the following entries

$$\Phi_{ij} = \begin{cases} K_{ij} & i \neq j \\ A_{11}^i - L_i C_1^i - \sum_{j, j \neq i} K_{ij} & i = j \end{cases} \tag{217}$$

the stochastic input $\Theta = (e^T, V^T, \Sigma^T)^T$, with $V = ((v^1)^T, ..., (v^N)^T)^T$ and $\Sigma$ defined as follows

$$\Sigma = (w_{11}^T, ..., w_{1N}^T, w_{21}^T, ..., w_{2N}^T, ..., w_{N1}^T, ..., w_{NN}^T)^T \tag{218}$$

and the matrix $B_Z$ as follows

$$B_Z = [-\overline{B}^{\perp}\Gamma \ \mid \ \Lambda \ \mid \ K_{\Xi}] \tag{219}$$

$$\overline{B}^{\perp} = col((B^i)^{\perp}), \quad \Lambda = diag(L^i (C^i B^i)^{\perp} + \bar{A}_{12}^i (C^i B^i)^{+}), \quad i = 1, 2, ..., N \tag{220}$$

$$K_{\Xi} = \begin{bmatrix} \tilde{K}_1 & 0 & 0 & 0 \\ 0 & \tilde{K}_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \tilde{K}_N \end{bmatrix}, \qquad \tilde{K}_i = [K_{i1} \quad K_{i2} \quad ..... \quad K_{iN}] \tag{221}$$

The matrices $\Phi$ and $B_Z$ depend on he entries of the consensus gain matrices $K_{ij}$ in the form (210)-(211), that need to be tuned. Let $g_\nu^j$ be selected as the $\nu - th$ diagonal element of the inverse of the estimation error covariance matrix of the steady Kalman filter associated to the

$j$-th local system model (188). This choice appears to be reasonable and motivated, as discussed in [9].

Now, define a vector $H$ which contains all the weights $h_\nu^{ij}$ to be determined. Consider the steady-state covariance matrix $P$ of vector $Z$ which is defined by the positive semi-definite solution of the Lyapunov equation

$$\Phi P + P\Phi^T + B_Z R_Z B_Z^T = 0 \tag{222}$$

with $R_Z$ being the covariance matrix of the stochastic input vector $\Theta$. It can be formulated the following optimization problem:

**Consensus gains optimization**:

Find vector $H$ such that the nonnegative optimality index $J = \mathtt{Tr}\ P$ is minimized.

The above choice for $H$, i.e. for the consensus gain parameters, gives the decentralized estimator optimal steady state performance. Unfortunately, the above optimization problem is non convex, hence its solution requires special iterative routines like for example the `fmins` Matlab function.

A **simplified** optimization strategy, devoted to reduce the number of free parameters, might be based on selecting

$$h_\nu^{ij} \equiv h^{ij} \qquad \forall \nu \tag{223}$$

In the simulation example it will be considered the above simplified tuning formula (223).

## 5.5 Unknown input reconstruction

In order to reconstruct the unknown input $u(t)$, we now refer to the $m$-dimensional dynamics of vectors $\bar{x}_2^{(i)}$, which can be extracted from (198)

$$\dot{\bar{x}}_2^{(i)} \;=\; \bar{A}_{21}^i \bar{x}_1^{(i)} + \bar{A}_{22}^i \bar{x}_2^{(i)} + u + \left(C^i B\right)^+ C^i \Gamma e \tag{224}$$

It must be stressed that vector $\bar{x}_2^{(i)}$ is available as a "noisy" part of the transformed output $\bar{y}^{(i)}$ (in fact, by(198), one has that $E(\bar{x}_2^{(i)}) = E(\bar{y}_2^{(i)})$). Furthermore, an asymptotically converging estimate of $\bar{x}_1^{(i)}$ is available from the consensus based estimator previously described. Any of the agents estimates $\xi$ can be used to that purpose, but it appears appropriate to choose the agent having the minimimun error covariance, i.e., if

$$P = \begin{bmatrix} P_1 & 0 & ... & 0 \\ 0 & \ddots & ... & 0 \\ 0 & 0 & P_i & 0 \\ 0 & 0 & ... & P_N \end{bmatrix} \tag{225}$$

we select the agent $i^*$ such that

$$i^* = \texttt{argmin}_i \; \|P_i\|_2 \tag{226}$$

The right hand side of dynamics (224) has a known deterministic part $\bar{A}_{21}^i \bar{x}_1^{(i)} + \bar{A}_{22}^i \bar{y}_2^{(i)}$, an **unknown deterministic component** $u(t)$ that is wanted to be reconstructed, and a stochastic part $\left(C^i B\right)^+ C^i \Gamma e$. Consider the following observer

$$\dot{\bar{z}}_2^{(i)} \quad = \quad \bar{A}_{21}^i \xi^i + \bar{A}_{22}^i \bar{y}_2^{(i)} + u^i, \qquad i = i^* \tag{227}$$

where $\xi^i$ is the estimation of $\bar{x}_1^{(i)}$ provided by the $i$-th agent, and $u^i$ is an observer injection input to be designed.

Let us study the dynamics of the error variable $\varepsilon^i = \bar{x}_2^{(i)} - \bar{z}_2^{(i)}$. It yields

$$\dot{\varepsilon}^i = \bar{A}_{21}^i (\bar{x}_1^{(i)} - \xi^i) + \left(C^i B\right)^+ C^i \Gamma e + u - u^i \tag{228}$$

which can be further elaborated as follows:

$$\dot{\varepsilon}^i = u - u^i + \varphi^i(t) \tag{229}$$

where

$$\varphi^i(t) = \left(C^i B\right)^+ C^i \Gamma e + \bar{A}_{21}^i \gamma^i(t) \tag{230}$$

Remember that $\gamma^i(t) \equiv \bar{x}_1^{(i)} - \xi^i$ tends asymptotically to a zero-mean stochastic residual. The idea is to design the observer injection term $u^i$ in such a way that $\varepsilon^i$ and its derivative $\dot{\varepsilon}^i$ are steered as close as possible to zero. We select $u^i$ in the form of a proportional-integral feedback, with gains $k_p$ and $k_i$ as follows:

$$u^i = k_p \varepsilon^i + k_i \int_0^t \varepsilon^i(\tau) d\tau \tag{231}$$

The closed loop system (229)-(231) can be represented as in the Figure 22. It follows from the scheme in Fig. 22 that the closed loop transfer function between the "input" $u + \varphi^i$ and the "output" $u^i$ is the following

$$P_i(s) = \frac{k_p s + k_i}{s^2 + k_p s + k_i} \tag{232}$$

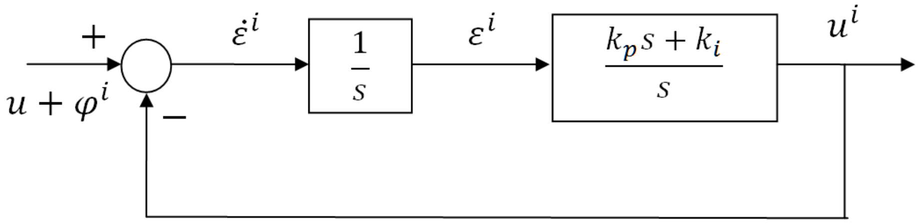



Figure 22: The equivalent block scheme for the closed loop system (229)-(231).

Then it can be selected the free design parameters $k_p$ and $k_i$ in order to guarantee that such a transfer function is close to the unitary value in a prescribed frequency range $[0, \omega_b]$. Thus, with properly selected gains $k_p$ and $k_i$ the following identity approximately holds

$$u^i \approx u + \varphi^i \tag{233}$$

If the magnitude of the stochastic term $\varphi^i$ is small enough with respect to the unknown input $u$, i.e. if $|\varphi^i| << |u|$ then the observer (227)-(232) guarantees that the following condition holds after the convergence transient

$$u^i \approx u \tag{234}$$

which means that the injection term of the observer (227) allows to reconstruct approximately the unknown input $u$.

The gains $k_p$ and $k_i$ should be designed in order to assign the transfer function $P_i(s)$ in (232) a bandwidth which includes the main spectral contents of the actual unknown input $u$. But, the higher the bandwidth of $P_i(s)$ the more statistical noise components will be injected in the unknown input estimate $u^i$. Hence, a careful tuning of those parameters requires some amount of a-priori information about the "spectral contents" of the unknown input. The amount of information required by this method is, however, milder than that required by the approaches in [1, 2, 14]. As an example, let the unknown input be an harmonic signal of the form $u(t) = Asin(\omega t)$. The effective application of our method requires to know an upperbound $\omega_b$ to the signal frequency ($\omega \in [0, \omega_b]$). Under this condition, the reconstruction of the unknown input is guaranteed for every of value of $A$. The methods in [1, 2, 14] requires to know a constant $M$ such that $|A\omega| \leq M$. Therefore, if the unknown input is large in magnitude ($A >> 1$), then it must be "sufficiently slowly varying" which can be not the case in some situation.

Note that **only one observer of the type (227) must be implemented**. The property that all agents converge to the same estimate makes it possible to select the "best" value $i^*$ of $i$ according to (226) and then implement the resulting observer.

## 5.6 Example

Let the linear system under consideration be represented by a fourth order model with $n = 4$, $m = 1$, $p = 3$, and the next system matrices

$$A = \begin{bmatrix} -1 & 0 & 1 & 0 \\ -1 & -2 & -1 & 0 \\ -1 & 0 & -1 & 0 \\ -0.1 & -1 & -3 & -5 \end{bmatrix}, \quad \Gamma = I_4, \quad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{235}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{236}$$

where $I_n$ represent the $n$-th order identity matrix. The scalar unknown input is selected as

$$u(t) = (10 + sin3t) \tag{237}$$

Assume that $N = 2$ **agents** get partial output measurements defined by the respective matrices

$$C^1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad C^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{238}$$

With reference to the first agent, its local system model has dimension $n_1 = 2$, the state index set $I_x^1 = \{1, 3\}$, the output index set $I_y^1 = \{1, 2\}$, and the next local subsystem matrices $A^1$ and $B^1$:

$$A^1 = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, \quad B^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{239}$$

Concerning the corresponding transformed local subsystem, the transformed state index set vector is $\bar{I}_x^1 = \{2\}$.

It can be checked that all the underlying assumptions $A.1$-$A.6$ are actually satisfied for the considered local system models.

The system and measurement covariance parameters are $Q = I_4$, $R^{(1)} = R^{(2)} = 0.01 I_2$, while the covariance parameters of the two inter-agent communication channels are $W_{12} = W_{21} = 0.01 I_3$. The suggested consensus-based observation scheme takes the following form.

$$\begin{aligned} \dot{\xi}^1 &= A_{11}^1 \xi^1 + A_{12}^1 (C^1 B^1)^+ y^{(1)} + K_{12}(\tilde{\xi}^{1,2} - \xi^1) + L_1 \left[ (C^1 B^1)^\perp y^{(1)} - C_1^1 \xi^1 \right] \\ \dot{\xi}^2 &= A_{11}^2 \xi^2 + A_{12}^2 (C^2 B^2)^+ y^{(2)} + K_{21}(\tilde{\xi}^{2,1} - \xi^2) + L_2 \left[ (C^2 B^2)^\perp y^{(2)} - C_1^2 \xi^2 \right] \end{aligned} \tag{240}$$

The tuning parameters $L^1$ and $L^2$ (the steady gains of the Kalman filter local) are evaluated by solving the appropriate Riccati equations. The consensus gains $K_{12}$ and $K_{21}$ have been evaluated by means of the **simplified** optimization procedure described in the section IV, that makes use of the simplified gain tuning relationship (223).

Let $P^1_{inv}$ and $P^2_{inv}$ be the inverse of the estimation error covariance matrix of the steady Kalman filter associated to the first and second local system model, respectively. They can be computed as follows

$$P^1_{inv} = 11.0499, \qquad P^2_{inv} = \begin{bmatrix} 3.7818 & 1.347 & 0.8959 \\ 1.347 & 4.2799 & 3.3486 \\ 0.8959 & 3.3486 & 15.6294 \end{bmatrix} \tag{241}$$

According to (223) and to the suggested choice for the $g^j_\nu$ elements (namely, the diagonal elements of matrices $P^1_{inv}$ and $P^2_{inv}$) the consensus gains $K_{12}$ and $K_{21}$ take the following form

$$K_{12} = h^{12} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 11.0499 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{242}$$

$$K_{21} = h^{21} \begin{bmatrix} 3.7818 & 0 & 0 \\ 0 & 4.2799 & 0 \\ 0 & 0 & 15.6294 \end{bmatrix} \tag{243}$$

The optimal values for the $h^{12}$ and $h^{21}$ coefficients were computed, by means of Matlab, as

$$h^{12} = 0.8116, \qquad h^{21} = 1.2381 \tag{244}$$

which give rise to the following expression for the consensus gain matrices

$$K_{12} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 8.2686 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad K_{21} = \begin{bmatrix} 4.6823 & 0 & 0 \\ 0 & 5.2991 & 0 \\ 0 & 0 & 19.3511 \end{bmatrix} \tag{245}$$

From the estimates $\xi^1$ and $\xi^2$ provided by the two agents, the state vector $x$ can be recovered according to equation (201). The transformation matrices $T^1$ and $T^2$ are:

$$T^1 = T^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{246}$$

which means that the transformed vector to be reconstructed is $\bar{x}_1 = [x_2, x_3, x_4]$. Let $\hat{\bar{x}}^i_1 = B^{i\perp}\xi^i$ denote the estimate of $\bar{x}_1$ provided by the $i$-th agent, according to the equation (196). Figure (23-24) depict the components of the observation errors $\bar{e}^i_1 = \hat{\bar{x}}^i_1 - \bar{x}_1$ with $i = 1, 2$.

In order to highlight the performance improvement due to the suggested consensus scheme, the estimates of the consensus based scheme are compared with the estimates obtained without communication between agents (i.e., by setting the consensus gain matrices $K_{12}$ and $K_{21}$ to

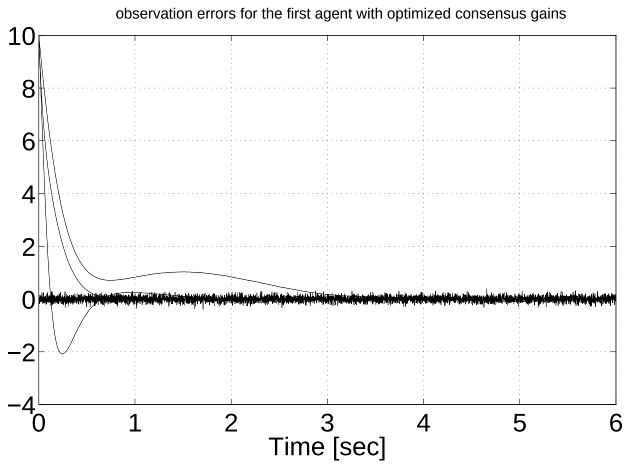



Figure 23: The components of the observation error vector $\bar{e}_1^1$

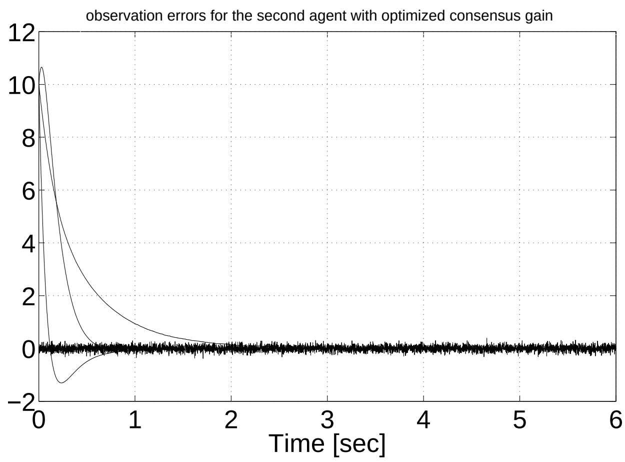



Figure 24: The components of the observation error vector $\bar{e}_1^2$

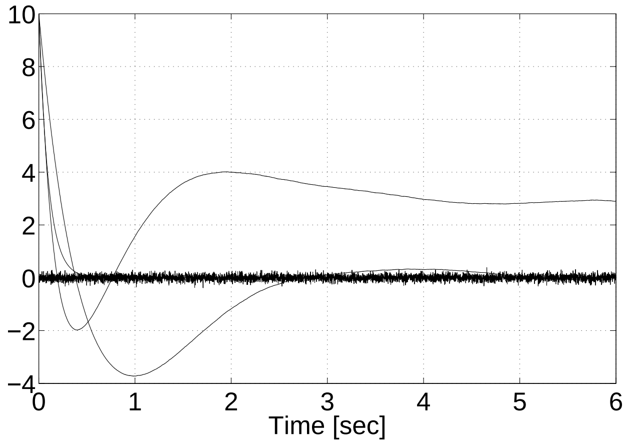



Figure 25: the observation error vector $\bar{e}_1^1$ without consensus.

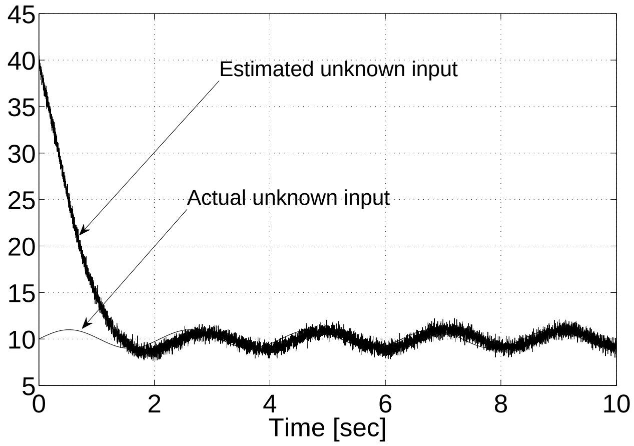



Figure 26: actual unknown input and its estimate $u^2$

zero.) The figure 25 shows the estimation error for the first agent **without** consensus. The benefits introduced by the communication between the agents are apparent.

Now let us address the reconstruction of the sinusoidal unknown input $u(t)$. According to the suggested procedure the observer (227) has been implemented selecting the agent $i$ which minimizes $\|n_i\|_2, i = 1, 2$, where $n_i$ is the vector which contains the diagonal elements of the covariance matrix $P_i$ associated to the $i$-th agent's estimation error. In the present example we obtain $\|n_1\|_2 = 0.5252, \|n_2\|_2 = 0.3397$, so we select the agent 2. The proportional and integral gains are selected as $k_p = p^2$, $k_i = 2p$, $p = 2$, in order to assign the transfer function $P_1(s)$ the pair of real negative poles $(-p, -p) \equiv (-2, -2)$. Figure 25 shows the actual and reconstructed unknown input, the latter which closely matches the actual profile after a short transient. The residual error is due to the propagation of the stochastic terms, which cannot be eliminated. To confirm that agent 2 is actually the best one, the estimation error featured by the two agents

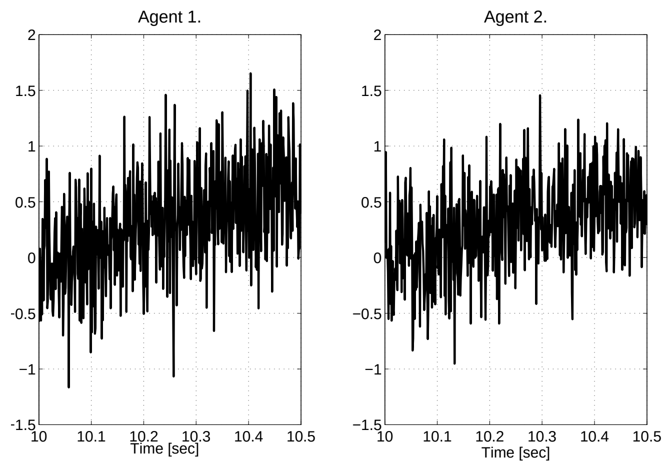



Figure 27: Unknown input estimation error for the agents 1 and 2

has been compared. Figure 27 shows a zoomed plot of signals $u^1 - u$ and $u^2 - u$. The computed variance of signal $u^1 - u$ is 0.396, while the computed variance of signal $u^2 - u$ is 0.351, which confirms the better performance of the second agent.

## 5.7 Conclusions

The problem of decentralized state estimation for a class of linear time-invariant systems affected by stochastic disturbances and deterministic **unknown input** has been addressed via a consensus based multi-agent Kalman estimator. A methodology for optimally tuning the consensus parameters is illustrated along with a scheme for reconstructing the unknown input vector, and all the procedures and methodologies have been verified by means of a thoroughly discussed simulation example.

The main improvement of the presented work over the recent related literature is the explicit introduction of deterministic, possibly large, unknown inputs in the plant dynamics. An interesting direction for further investigations could aim to relax the strong observability property for, at least, a subset of the local system models. In the present paper we have limited our attention to analyzing the consensus-based communication between **linear** Kalman observers. Consensus based communication between nonlinear observers, e.g. sliding mode observers, can be an interesting and promising enhancement of the present treatment, which is another challenging task for next research activities.

## APPENDIX

**Proof of Theorem 14**

Consider the vector $\bar{Z} = \bar{\Xi} - \bar{X_1}$, $\bar{Z} \in \mathbb{R}^{(n-m)N}$, where $\bar{\Xi}$ and $\bar{X_1}$ are defined as follows

$$\begin{aligned} \bar{\Xi} &= [E(\xi^1)^T, ..., E((\xi^N)^T]^T, \qquad \bar{\Xi} \in \mathbb{R}^{(n-m)N} & (247) \\ \bar{X_1} &= [E(\bar{x}_1)^T \;\; E(\bar{x}_1)^T \;\; ... \;\; E(\bar{x}_1)^T]^T \qquad \bar{X_1} \in \mathbb{R}^{(n-m)N} & (248) \end{aligned}$$

and the associated dynamics, which takes the form

$$\dot{\bar{Z}} = \Phi Z \tag{249}$$

with the characteristic matrix $\Phi = \{\Phi_{ij}\}$ having the following entries

$$\Phi_{ij} = \begin{cases} K_{ij} & i \neq j \\ A_{11}^i - L_i C_1^i - \sum_{j, j \neq i} K_{ij} & i = j \end{cases} \tag{250}$$

We shall also define:

$$h_\nu^{ij} = \begin{cases} \bar{h}^{ij} & \nu \in \bar{I_x^i} \\ \tilde{h}^{ij} & \nu \in \{1, ...n - m \backslash \bar{I_x^i}\} \end{cases}, \qquad \bar{h}^{ij}, \tilde{h}^{ij} > 0 \tag{251}$$

$$\begin{aligned} G_1^{ij} &= diag\{g_{\nu_1^i}^{ij}, ..., g_{\nu_{n_i-m}^i}^{ij}\}, \qquad \nu_1^i, ..., \nu_{n_i-m}^i \in \bar{I_x^i} & (252) \\ G_2^{ij} &= diag\{g_{\bar{\nu}_1^i}^{ij}, ..., g_{\bar{\nu}_{n-m-n_i}^i}^{ij}\}, \qquad \bar{\nu}_1^i, ..., \bar{\nu}_{n-m-n_i}^i \in \{1, ...n - m \backslash \bar{I_x^i}\} & (253) \end{aligned}$$

and let

$$K_{1,0}^{ij} = \bar{h}^{ij} G_1^{ij}, \qquad K_{2,0}^{ij} = \tilde{h}^{ij} G_2^{ij} \tag{254}$$

Stability of the consensus scheme corresponds to the stability of the characteristic matrix $\Phi$. Let us consider a matrix

$$\Phi^{'} = \begin{bmatrix} \Phi^{11} & \Phi^{12} \\ \Phi^{21} & \Phi^{22} \end{bmatrix} \tag{255}$$

where the generic entries of the submatrices in (255) is

$$\Phi_{ij}^{11} = \begin{cases} K_{ij}^{1,1} & i \neq j \\ \bar{A}_{11}^i - L^i \bar{C}_1^i - \Sigma_{j, j \neq i} K_{ij}^{1,0} & i = j \end{cases} \tag{256}$$

$$\Phi_{ij}^{12} = \begin{cases} K_{ij}^{1,2} & i \neq j \\ 0 & i = j \end{cases}, \qquad \Phi_{ij}^{21} = \begin{cases} K_{ij}^{2,1} & i \neq j \\ 0 & i = j \end{cases} \tag{257}$$

$$\Phi_{ij}^{22} = \begin{cases} K_{ij}^{2,2} & i \neq j \\ -\Sigma_{j, j \neq i} K_{ij}^{2,0} & i = j \end{cases} \tag{258}$$

where $K_{ij}^{1,1}, K_{ij}^{1,2}, K_{ij}^{2,1}$ and $K_{ij}^{2,2}$ are submatrices of $K_{ij}$ obtained by deleting some selected elements as follows:

- For $K_{ij}^{1,1}$ we delete the elements having row and column indexes specified by $\{1, ..., n - m \backslash \bar{I_x^i}\}$ and $\{1, ..., n - m \backslash \bar{I_x^j}\}$, respectively.

- For $K_{ij}^{1,2}$ we delete the elements having row and column indices specified by $\{1,...,n-m\backslash\bar{I}_x^i\}$ and $\bar{I}_x^j$, respectively.

- For $K_{ij}^{2,1}$ we delete the elements having the row and column indexes specified by $\bar{I}_x^i$ and $\{1,...,n-m\backslash\bar{I}_x^j\}$, respectively;

- For $K_{ij}^{2,2}$ we delete the elements having the row and column indexes specified by $\bar{I}_x^i$ and $\bar{I}_x^j$, respectively;

It turns out that matrix $\Phi^{'}$ is cogredient to $\Phi$ in (249).

Let us consider such $\tilde{h}^{ij} \geq 0, i, j = 1, ..., N$, that (A.4) is satisfied, and look at the submatrix $\Phi^{22}$ (which depends on $\tilde{h}^{ij}$ and not on $\bar{h}^{ij}$). Let $\mathsf{G}_\nu$ be the digraph as defined in the Section 3. Assumption (A.4) implies that each digraph $\bar{\mathsf{G}_\nu}$, opposite to $\mathsf{G}_\nu, (\nu = 1, ..., n-m$, has only one closed strong component. Therefore, those submatrices of $\Phi^{'}$ which represents the Laplacian matrices of $\mathsf{G}_\nu, (\nu = 1, ..., n-m)$, are cogredient to lower-block-triangular matrices with two diagonal blocks, where the first is an irreducible Metzler matrix which has one eigenvalue at the origin and the remaining ones in the left-half plane, and the second is a diagonally dominant Meztler matrix, which is therefore stable [5]. The center node of $\mathsf{G}_\nu$ must belong to the set of nodes of the unique closed strong component of $\bar{\mathsf{G}_\nu}$,. Therefore, $\Phi^{22}$ is composed by submatrices of $\Phi^{'}$ that are obtained from the irreducible Metzler matrices by deleting their rows and columns with indices corresponding to the nodes of the strong component of $\bar{\mathsf{G}_\nu}$. These irreducible submatrices are, in general, cogredient to

$$L_\nu^D = \begin{bmatrix} -\sum_{j,j\neq 1}\alpha_{1j} & \alpha_{12} & \cdots & \alpha_{1\tilde{N}} \\ \alpha_{21} & -\sum_{j,j\neq 2}\alpha_{2j} & \cdots & \alpha_{1\tilde{N}} \\ \ddots & \ddots & \ddots & \ddots \\ \alpha_{\tilde{N}1} & \alpha_{\tilde{N}2} & \cdots & -\sum_{j,j\neq \tilde{N}}\alpha_{\tilde{N}j} \end{bmatrix} \tag{259}$$

where $\tilde{N} \leq N, \alpha_{ij} \geq 0, \alpha_{21} > 0$. Deleting the first row and column of the block matrix $L_\nu^D$ we obtain a block matrix in which the first row is strictly diagonally dominant, being that $\alpha_{21} > 0$ as a consequence of the irreducibility of $L_\nu^D$. Consequently, this matrix is Metzler and quasi-dominant diagonal, which implies that it is Hurwitz [11]. Therefore, the whole matrix $\Phi^{22}$ is Hurwitz, having in mind assumptions (A.5) and (A.6).

Assuming that $h_{ij} = 0, \forall i, j$; we obtain that $\Phi^{12} = 0$, and $\Phi^{11}$ is asymptotically stable, having in mind that the matrices $\bar{A}_{11}^i - L^i\bar{C}_1^i, (i = 1, ..., n-m)$ are Hurwitz by the assumption A.3. This implies that the whole matrix $\Phi$ is Hurwitz. Now consider some $\tilde{h}_{ij}$ as above and choosing such $\bar{h}_{ij} \geq 0$ in accordance with (A.4) we can directly conclude that there exists such $\varepsilon > 0$

that the system keeps asymptotically stable as long as $\bar{h}_{ij} < \varepsilon$, having in mind the continuous dependence of the eigenvalues of $\Phi$ on the values of $\bar{h}_{ij}$. Theorem 1 is proved. $\triangleright$.

# References


[1] F.J. Bejarano, L. Fridman and A. Poznyak. “Exact state estimation for linear systems with unknown inputs based on hierarchical super-twisting algorithm”. Int. J. Robust Nonlinear Control, 17, pp. 1734–1753, 2007.

[2] F.J. Bejarano, L. Fridman and A. Poznyak. “Unknown Input and State Estimation for Unobservable Systems”. SIAM J. Contr. Optimizat., 48, 2, pp. 1155–1178, 2009.

[3] C. Edwards, S.K. Spurgeon. Sliding Mode Control: Theory And Applications. Taylor and Francis (1998)

[4] C. Edwards, S. K. Spurgeon, and R. J. Patton,“Sliding mode observers for fault detection and isolation,” *Automatica*, vol. 36, pp. 541–553, 2000.

[5] P. Ferguson and J. How, “Decentralized estimation algorithms for formation flying spacecraft”. Proc. AIAA Guid. Nav. Control Conf, 2003.

[6] T. Floquet, C. Edwards, S.K. Spurgeon “On sliding mode observers for systems with unknown inputs”. International Journal of Adaptive Control and Signal Processing, 21, pp. 638–656, 2007.

[7] L. Fridman, A. Levant, and J.Davila, “Observation of Linear Systems with Unknown Inputs via High-Order Sliding-Mode,” *International Journal of Systems Science*, vol. 38, no. 10, pp. 773–791, 2007.

[8] H.R. Hashimpour, S. Roy and A.J. Laub “Decentralized structures for parallel Kalman filtering”, IEEE Trans. Aut. Control, vol. 33, pp 88-93, 1988

[9] M. L. J. Hautus. “Strong detectability and observers” Linear Algebra and its Applications, 50, pp. 353-368, 1983.

[10] B.P. Molinari. “A strong controllability and observability in linear multivariable control”, IEEE Transaction on Automatic Control, 21(5), pp. 761-764, 1976.

[11] R. Olfati-Saber, “Distributed Kalman filter with embedded consensus filters”. Proc. 2005 IEEE Conf. Decision and Control CDC’05.

[12] W. Ren, R.W. Beard and D.B. Kingston "Multi-agent Kalman consensus with relative uncertainty". Proc. 2005 American Control Conference ACC'05.

[13] C.W. Sanders, E.C. Tacker and T.D. Linton "Specific structures for large scale state estimation algorithms having information exchange". IEEE Transactions on Automatic Control, 23, pp. 255-260, 1978.

[14] E.C. Tacker and C.W. Sanders "Decentralized structures for state estimation in large scale systems". Large Scale Systems, 1, pp. 39-49, 1980.

[15] L. Moreau "Stability of multiagent systems with time-dependent communication links". IEEE Transactions on Automatic Control, vol. 50, pp 169- 182, 2005.

[16] Pillosu, S., Pisano, A., Usai, E.(2010). " Consensus based decentralized estimation for linear plants with unknown inputs ". 49th IEEE Conference on Decision and Control (CDC) 6857-6862 (2010)

[17] Pillosu, S., Pisano, A., Usai, E.(2011). " Desentralized State Estimation for Linear Systems with Unknown Inputs. A Consensus-based Approach". IET Control Theory Appl 4(1):498-506 (2010).

[18] W. Ren and R.W. Beard "Consensus seeking in multi-agent systems using dynamically changing interaction topologies". IEEE Transactions on Automatic Control, vol. 50, pp 655-661, 2005.

[19] S. Simani, C. Fantuzzi and R.J. Patton, "Model-based fault diagnosis in dynamic systems using identification techniques", Advances in Industrial Control, Springer–Verlag, London, 2002.

[20] A. Speranzon, C. Fischione and K. H. Johansson, " Distributed and collaborative estimation over wireless sensor networks". Proc. 2006 IEEE Conf. Decision and Control CDC'06.

[21] F. Xiao, L. Wang " Consensus problems for high-dimensional multi-agent systems". IET Control Theory & Applications, vol. 1, No.3. May 2007.

# Part II

# Applications

# 6 Unknown-input observation techniques in Open Channel Hydraulic Systems

## 6.1 Preliminaries

This chapter addresses the problem of state and disturbance estimation for an open-channel hydraulic system. Particularly, a cascade of $n$ canal reaches, joined by gates, is considered. The underlying Saint-Venant system of PDEs is managed by means a collocation-based finite-dimensional approximation. The resulting nonlinear systems' dynamics are linearized around a subcritical uniform flow condition, and an estimation algorithm is designed by combining a conventional linear Unknown-Input Observer (UIO) and a nonlinear Disturbance Observer (DO) based on the sliding mode approach. By using measurements of the water level in three points per reach, the suggested algorithm is capable of estimating, both, the time varying infiltration and the discharge variables in the middle point of the reaches. The UIO and DO design procedures are constructively illustrated throughout the chapter. Simulation results are discussed to verify the effectiveness of the proposed schemes.

## 6.2 Introduction

Most open-channel hydraulic systems are currently manually operated by flow control gates or by means of self regulating gates that maintain either the flow or the level. Current goals in this field include their automatic operation in order to improve water distribution efficiency and safety [21]. A key problem is to reconstruct the information needed for control or monitoring diagnosis purposes (water levels, discharges, and infiltrations), some of which are intrinsically impossible or difficult to measure, by limiting the number of required sensors in the field.

Flow sensors, in particular, are expensive devices, and it would be desirable to accomplish the estimation and control tasks by using water level measurement sensors only. From the perspective of designing model-based control or diagnosis systems, this calls for easily tractable reduced-order numerical models that can accurately reflect the nonlinear behavior of water flow. The problem of deriving simple yet accurate models of the open channel systems dynamics is still an open and active area of investigation.

Open channel hydraulic systems are described by two nonlinear coupled partial differential equations (Saint-Venant Equations). It is widely recognized that relatively low-order approximation of the Saint-Venant Equation can provide sufficiently accurate information (see [21]). A number of finite-difference and finite elements approximation techniques have been suggested in the literature (see[27],[6]). Research efforts have been made to develop models' computationally simple yet accurate enough to be used for model-based observer or controller design purposes [18, 12, 31]. The collocation method is a special case of the so-called weighted-residual method, commonly used in computational physics for solving partial differential equations [13, 9].

For the considered open channel hydraulic systems it has been shown (see [29]) that a three-point orthogonal collocation model can be used to design a model-based nonlinear controller with guaranteed properties of closed loop stability. It has also been shown that the response of the reduced order collocation model is close enough to that obtained using high-accuracy solvers of commercial dedicated software packages [10, 4].

In [4] a three-point collocation-based nonlinear model of a single-reach irrigation canal was developed considering the canal begin, middle, and end points, respectively, and a constant uncertain infiltration was also taken into account. An observer capable of reconstructing the level variables and the constant infiltration was designed in [4] by measuring the level in the middle of the reach and the upstream and downstream flows. In [5], the observer considered in [4] was used to design a stabilizing feedback level control law.

In [26], a Kalman filter approach was developed for reconstructing unmeasured level variables by employing a discrete-time model of a the canal. In [1], an $H_\infty$ observer was developed to minimize the effect of the unknown inputs on the accuracy of the estimates. In [1], it was considered a system representation including time-varying delays, and a continuous-time full-order Unknown Input Observer (UIO) was suggested and used for the detection of certain faults in the irrigation canal actuators.

### 6.2.1 Aim, Contribution and structure of the chapter

The aim of this chapter is the development of new observation and estimation algorithms for a cascade of canal reaches subject to unmeasurable disturbances. Starting from the collocation-based model presented in [4] for a single-reach canal, here we **consider a cascade of n pools**, then the model described in [4] is properly generalized and manipulated. We also relax some of the standing assumptions made in [4], namely in the present chapter we :

i.) dispense with the need of flow rate measurements by allowing only level measurements;

ii.) consider a time varying infiltration, and

iii.) reduce the number of level sensors in the case of negligible infiltration.

Property ii. is important since the infiltration term also models seepage and evaporation phenomena, which are subject to seasonal and/or diurnal fluctuations. Section 6.2 recalls the Saint-Venant equations. Section 6.3 reviews the three-point collocation-based nonlinear model of a single reach presented in [4]. Additionally, in Subsection 6.3.1, the model is extended to a cascade of $n$ canal reaches, and in Subsection 6.3.2 the model linearization around the uniform flow condition is performed.

In Section 6.4, the two estimation problems(named "**Problem 1**" and "**Problem** 2") addressed in this chapter are stated. Only level measurement are permitted in both. Problem 1 involves the simplifying assumption of no (or negligible) infiltration, and, as a counterpart, it allows certain level variables to be not measured. The flow variables are wanted to be estimated, and the unmeasured level variables are wanted to be reconstructed as well.

Problem 2 deals with the general non-zero infiltration case but requires the measurement of the level variables in three points per reach. The flow variables are wanted to be estimated too, along with the unknown infiltrations, after a **finite-time estimation transient**.

Problem 1 and Problem 2 give rise to distinct observation problems for certain Linear Time-Invariant System with Unknown Inputs (LTISUI). In Section 6.5 a method for state estimation and unknown input reconstruction in general LTISUI is recalled. The approach is based on the assumption of "**Strong Observability**" [10, 15, 3], a structural condition on the LTISUI mathematical model. Such restriction has different, although equivalent, formulations, the simplest of which establishes that a certain matrix pair should be observable. Overall, the proposed scheme combines a linear Unknown-Input Observer (UIO) and a nonlinear Disturbance Observer (DO) based on the sliding mode approach.

In the Section 6.6, a case study of a canal with rectangular section and three reaches is illustrated. The parameters of the linearized models previously developed are computed on the basis of realistic input data.

In the successive Subsections 6.6.1 and 6.6.2 the techniques described in the Section 6.5 are applied to solve the estimation Problem 1 and Problem 2, respectively. It is shown, in both cases, that the structural requirement of strong observability holds for the resulting models, and corresponding simulation results are illustrated and commented, which will confirm the expected performance of the suggested observers. Section 6.8 states some final conclusion and draws possible lines for next researches on the topic.

## 6.3 Formulation of the Problem

Water flow dynamics in an open channel are governed by the system of Saint-Venant partial differential equations [2]

$$\begin{aligned} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} &= w & (260)\\ \frac{\partial Q}{\partial t} + \frac{\partial (Q^2/S)}{\partial x} + gS\left(\frac{\partial H}{\partial x} - I + J\right) &= \frac{1}{2}(w - |w|)\frac{Q}{S} & (261) \end{aligned}$$

where $x \in [0, L]$ is the spatial variable ($L$ being the channel length), $t$ is the time variable, and $S(x,t)$, $Q(x,t)$ and $H(x,t)$ being the wet section, water flow rate and relative water level, respectively. The term $w = w(x,t)$ in the right-hand side of (260), (261) represents the infiltration. $J$ represents the friction term, which has the following expression

$$J = \frac{Q|Q|}{Di^2}, \qquad Di = kS\left(\frac{S}{P}\right)^{\frac{2}{3}}. \tag{262}$$

where $k$ is the Strickler friction coefficient, $P(x,t)$ being the transversal wet length and $I$ is the canal slope. We refer to canals with rectangular section, so if $B$ is the constant canal width one has that:

$$S = BH, \quad P = B + 2H. \tag{263}$$

Thus, on the basis of (263) model (260)-(261) can be rewritten in terms of the $Q$ and $H$ variables only, and, in particular, eq. (260) modifies as follows

$$\frac{\delta H}{\delta t} = -\frac{1}{B}\frac{\delta Q}{\delta x} + \frac{1}{B}w. \tag{264}$$

If the canal slope is sufficiently low, as it is the case, e.g., in irrigation channels, it can be assumed subcritical flow condition. This makes it needed to complement (260)-(261) with two boundary conditions (BCs), one upstream and one downstream (cfr. [2], sect. 2.1.4). Moreover, since we are going to consider (a cascade of) water channels joined by hydraulic gates, which allow to deliver a given upstream and downstream discharge, it appears an appropriate choice that of imposing the discharge at the upstream and downstream boundaries as BCs, rather than the water depth (cfr. [2], sect. 2.1.4). Then, we complement (264)-(261) with

$$Q(0,t) = Q_A(t), \quad Q(L,t) = Q_B(t), \tag{265}$$

and with initial conditions

$$H(x,0) = H^0(x), \quad Q(x,0) = Q^0(x). \tag{266}$$

compatible with the considered BCs (265). It shall be noticed that in more complex conditions (e.g. an intermediate situation where the flow along the channel is partly subcritical and partly supercritical) one might need to specify more than two BCs and/or to refer to more involved weak formulations of the BCs (see e.g. [28]), whose treatment appears out of the scope of this chapter.

## 6.4 Collocation-based finite-dimensional model

In [10] it was shown that the Saint Venant equation can be effectively approximated by ordinary differential equations of finite dimension using a collocation point Galerkin method. It has been also shown that three collocation points placed at the canal upstream, middle, and downstream points (say points $A$, $M$, and $B$, respectively) leads to a sufficiently accurate representation for observation and control purposes. Consider the channel depicted in the Fig. 28, interconnecting the upstream and downstream reservoirs through the adjustable gates #1 and #2, and subject to uniform water infiltration and withdrawal concentrated in correspondence with the gate. By evaluating equation (264) at the collocation points, and discretizing by finite-differences the resulting spatial derivatives of the flow variable, let us recall the following dynamics ([10]) for the state variables $H_A(t)$, $H_M(t)$ and $H_B(t)$:

$$\begin{array}{rcl} \dot{H}_A(t) & = & \frac{1}{BL}[-4Q_M(t) + 3Q_A(t) + Q_B(t)] + \frac{w(t)}{B} \\ \dot{H}_M(t) & = & \frac{1}{BL}[Q_A(t) - Q_B(t)] + \frac{w(t)}{B} \\ \dot{H}_B(t) & = & \frac{1}{BL}[4Q_M(t) - Q_A(t) - 3Q_B(t)] + \frac{w(t)}{B}. \end{array} \tag{267}$$

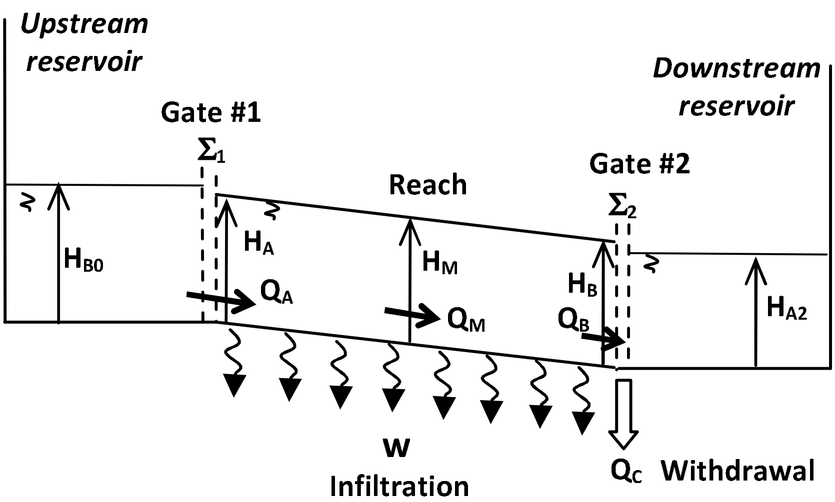



Figure 28: Single canal reach with infiltration loss.

The model can be simplified by expressing the upstream and downstream flow variables $Q_A$ and

$Q_B$ as functions of the water levels and of the opening section of the gates. In the permanent flow regime the next relations hold [7]:

$$Q_A = \eta_1 \Sigma_1 \sqrt{2g(H_{B0} - H_A)} \tag{268}$$

$$Q_B = Q_C + \eta_2 \Sigma_2 \sqrt{2g(H_B - H_{A2})} \tag{269}$$

where a withdrawal $Q_C$ from the users is taken into account in the downstream flow balance, $\eta_1$ and $\eta_2$ are the discharge coefficients of the upstream and downstream gates, respectively, $\Sigma_1$ and $\Sigma_2$ are the opening sections of the two gates, and $H_{B0}$ and $H_{A2}$ are, respectively, the water levels within the upstream and downstream reservoirs.

Let us derive the dynamic relationship between the middle point flow variable $Q_M$ and the remaining system variables. Such a relation can be derived by applying a collocation based discretization method to the second Saint Venant equation (261) (see [10, 4]).

$$\dot{Q}_M = \psi_q(Q_A, Q_B, Q_M, H_A, H_M, H_B, w) \tag{270}$$

The form of the nonlinear function $\psi_q$ is [10, 4]

$$\begin{aligned} \psi_q(\cdot) &= gBH_M\left(I + \frac{H_A - H_B}{L}\right) + \left(\frac{2(Q_A - Q_B)}{BL}\right)\frac{Q_M}{H_M} \\ &+ \left(\frac{H_B - H_A}{BLH_M^2} - \frac{g}{K^2 BH_M(\frac{BH_M}{B+2H_M})^{\frac{4}{3}}}\right) Q_M^2 \end{aligned} \tag{271}$$

The above model (267)-(271) is going to be generalized in the next subsection 6.4.1 to represent a **cascade** of canal reaches.

### 6.4.1 $n$-reaches cascade modeling

Now let us consider a cascade of $n$ canal reaches connecting the two upstream and downstream reservoirs, separated by $n+1$ adjustable gates, and subject to infiltration losses, spatially uniform along each canal, as represented in the Figure 29.

By choosing three collocation points for each channel, and using the same notation as before to denote the resulting points $A_i$, $M_i$, $B_i$ $(i = 1, 2, ..., n)$ the model (267)-(269) can be generalized as follows

$$\begin{aligned} \dot{H}_{Ai} &= \tfrac{1}{B_i L_i}[-4Q_{Mi} + 3Q_{Ai} + Q_{Bi}] + \tfrac{w_i}{B_i} \\ \dot{H}_{Mi} &= \tfrac{1}{B_i L_i}[Q_{Ai} - Q_{Bi}] + \tfrac{w_i}{B_i} \\ \dot{H}_{Bi} &= \tfrac{1}{B_i L_i}[4Q_{Mi} - Q_{Ai} - 3Q_{Bi}] + \tfrac{w_i}{B_i} \end{aligned} \tag{272}$$

$$Q_{Ai} = \eta_i \Sigma_i \sqrt{2g(H_{Bi-1} - H_{Ai})} \tag{273}$$

$$Q_{Bi} = Q_{Ci} + Q_{Ai+1} = Q_{Ci} + \eta_{i+1}\Sigma_{i+1}\sqrt{2g(H_{Bi} - H_{Ai+1})} \tag{274}$$

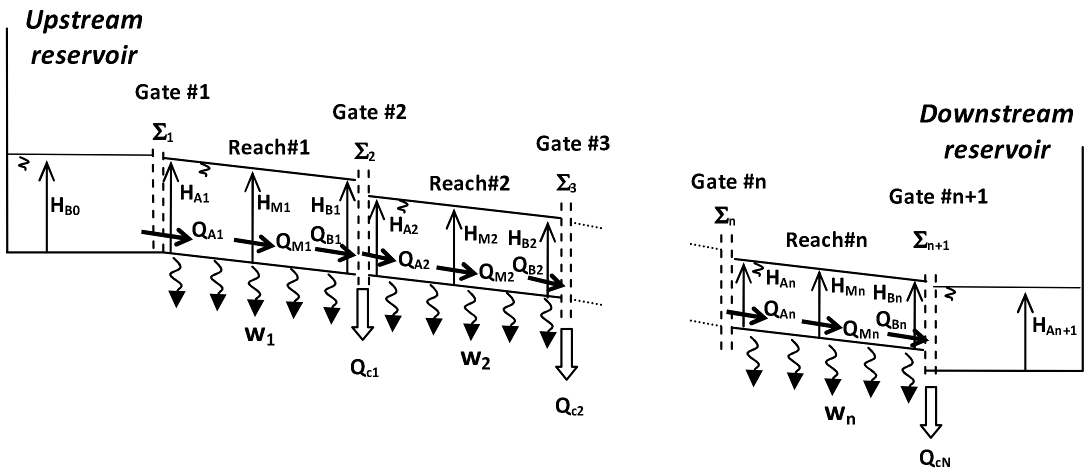



Figure 29: Cascade of $n$ canal reaches with infiltration losses.

where $H_{Ai}$, $H_{Mi}$ and $H_{Bi}$ $(i = 1, 2, ..., n)$ are the state variables, $Q_{Ai}$, $Q_{Mi}$ and $Q_{Bi}$ $(i = 1, 2, ..., n)$ denote the flow at the collocation points, $w_i$ $(i = 1, 2, ..., n)$ is the infiltration in the $i-$th reach, $\eta_j$ and $\Sigma_j$ $(j = 1, 2, ..., n+1)$ are the discharge coefficients and the opening sections of the $i$-th gate, and $Q_{Ci}$ is the widthdrawal request from the users. $H_{B0}$ and $H_{An+1}$ represent the **constant** levels in the upstream and downstream reservoirs.

Considering (273) and (274) into (272) one obtains a more compact expression for the systems' nonlinear dynamics:

$$\begin{array}{rcl} \dot{H}_{Ai} & = & \frac{1}{B_iL_i}[f_A(H_{Bi-1}, H_{Bi}, H_{Ai}, H_{Ai+1}, \Sigma_i, \Sigma_{i+1}) - Q_{Mi} + Q_{Ci}] + \frac{w_i}{B_i} \\ \dot{H}_{Mi} & = & \frac{1}{B_iL_i}[f_M(H_{Bi-1}, H_{Bi}, H_{Ai}, H_{Ai+1}, \Sigma_i, \Sigma_{i+1}) - Q_{Ci}] + \frac{w_i}{B_i} \\ \dot{H}_{Bi} & = & \frac{1}{B_iL_i}[f_B(H_{Bi-1}, H_{Bi}, H_{Ai}, H_{Ai+1}, \Sigma_i, \Sigma_{i+1}) + 4Q_{Mi} - 3Q_{Ci}] + \frac{w_i}{B_i} \end{array} \tag{275}$$

with implicit definition of the nonlinear functions $f_A(\cdot)$, $f_M(\cdot)$ and $f_B(\cdot)$.

The dynamic relationship between the middle point flow variable $Q_{Mi}$ and the remaining system variables can be derived by generalizing the previously given relationship (270), together with (273)-(274) as follows:

$$\dot{Q}_{Mi} = \psi_{iq}(Q_{Ai}, Q_{Bi}, Q_{Mi}, H_{Ai}, H_{Mi}, H_{Bi}, w_i) \tag{276}$$

$$\begin{array}{rcl} \psi_{iq}(\cdot) & = & gB_iH_M\left(I + \dfrac{H_{Ai} - H_{Bi}}{L_i}\right) + \left(\dfrac{2(Q_{Ai} - Q_{Bi})}{BL}\right)\dfrac{Q_{Mi}}{H_{Mi}} \\ & + & \left(\dfrac{H_{Bi} - H_{Ai}}{B_iL_iH_{Mi}^2} - \dfrac{g}{K^2B_iH_{Mi}(\frac{B_iH_{Mi}}{B_i+2H_{Mi}})^{\frac{4}{3}}}\right)Q_{Mi}^2 \end{array} \tag{277}$$

### 6.4.2 Linearized Model

The nonlinear model (275) can be linearized in a vicinity of the uniform flow condition (see[7]). Let $\overline{Q}_i$ $(i = 1, 2, ..., n)$, denote the flow in the i-th pool in the uniform flow condition. Let also $\overline{H}_i$ $(i = 1, 2, ..., n)$ be the corresponding water levels, and $\overline{\Sigma}_j$ $(j = 1, 2, ..., n+1)$ be the corresponding gates opening sections.

Define the corresponding deviation variables

$$\begin{aligned} h_{Ai} &= H_{Ai} - \overline{H}_i, \quad h_{Mi} = H_{Mi} - \overline{H}_i, \quad h_{Bi} = H_{Bi} - \overline{H}_i, && (278) \\ q_{Mi} &= Q_{Mi} - \overline{Q}_i, \quad \sigma_j = \Sigma_j - \overline{\Sigma}_j. && (279) \end{aligned}$$

where $i = 1, 2, ..., n$ and $j = 1, 2, ..., n + 1$.

Relation (273)and (274) can be linearized as follows in a vicinity of the uniform flow condition [7]

$$Q_{Ai} = \overline{Q}_i + a_i\sigma_i(t) + b_i[h_{Bi-1}(t) - h_{Ai}(t)], \tag{280}$$

$$Q_{Bi} = Q_{Ci} + \overline{Q}_{i+1} + a_{i+1}\sigma_{i+1}(t) + b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)], \quad i = 1, ...n \tag{281}$$

with the coefficients

$$a_i = \eta_i\sqrt{2g(\overline{H}_{i-1} - \overline{H}_i)}; \quad b_i = \frac{\eta_i\overline{\Sigma}_i\sqrt{2g}}{\sqrt{2(\overline{H}_{i-1} - \overline{H}_i)}} \tag{282}$$

and

$$h_{An+1} = 0, \quad h_{B0} = 0. \tag{283}$$

as a consequence of the fact that the water level in the upstream and downstream reservoirs is supposed to keep constant.

Substituting (280)-(281) into (272)-(274), and considering the continuity conditions

$$\overline{Q}_i = \overline{Q}_{i+1} + Q_{Ci}, \qquad i = 1, ...n \tag{284}$$

one obtains the linearized dynamics of the deviation level variables:

$$\begin{aligned} \dot{h}_{Ai} &= \tfrac{1}{B_iL_i}\{-4q_{Mi} + 3a_i\sigma_i(t) + 3b_i[h_{Bi-1}(t) - h_{Ai}(t)] \\ &+ \; a_{i+1}\sigma_{i+1}(t) + b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)]\} + \tfrac{w_i}{B_i} \\ \dot{h}_{Mi} &= \tfrac{1}{B_iL_i}\{a_i\sigma_i(t) + b_i[h_{Bi-1}(t) - h_{Ai}(t)] \\ &- \; a_{i+1}\sigma_{i+1}(t) - b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)]\} + \tfrac{w_i}{B_i} \\ \dot{h}_{Bi} &= \tfrac{1}{B_iL_i}\{4q_{Mi} - a_i\sigma_i(t) - b_i[h_{Bi-1}(t) - h_{Ai}(t)] \\ &- \; 3a_{i+1}\sigma_{i+1}(t) - 3b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)]\} + \tfrac{w_i}{B_i} \end{aligned} \tag{285}$$

Now defining vectors

$$\begin{aligned} h &= [h_{A1}\ h_{M1}\ h_{B1}\ h_{A2}\ ...h_{An}\ h_{Mn}\ h_{Bn}]^T, \quad h \in R^{3n} && (286)\\ \sigma &= [\sigma_1\ \sigma_2 ... \sigma_{n+1}]^T, \quad \sigma \in R^{n+1} && (287)\\ q_M &= [q_{M1}\ q_{M2} ... q_{Mn}]^T, \quad q \in R^n && (288)\\ w &= [w_1\ w_2 ... w_n]^T, \quad w \in R^n && (289)\end{aligned}$$

it is possible to rewrite the system (285) in the compact state space form

$$\dot{h} = Ah + M_\sigma \sigma + M_q q_M + M_w w \tag{290}$$

with implicitly defined constant matrices $A$, $M_\sigma$, $M_q$ and $M_w$ of appropriate dimension. Vector $q_M$ is governed by the nonlinear differential equation

$$\dot{q}_M = \psi(h, q_M, \sigma, w) \tag{291}$$

where

$$\psi(h, q_M, \sigma, w) = [\psi_{1q}(\cdot),\ \psi_{2q}(\cdot),\ \ldots\ , \psi_{nq}(\cdot)]^T \tag{292}$$

and the functions $\psi_{iq}(\cdot)$ $(i = 1, 2, ..., n)$ given in (277). Note that the nonlinear dynamics (291)-(292) **need not to be linearized** since vector $q_M$ is going to be treated as an unknown input of the systems, rather than as a part of the system state.

Since the first and last equation in (285) are not affected by the level variables $h_{Mi}$, it is possible to consider a reduced-order version of system (290) where the state vector $h$ is replaced by the reduced-order version

$$\tilde{h} = [h_{A1}\ h_{B1}\ h_{A2}\ ...h_{An}\ h_{Bn}]^T, \quad \tilde{h} \in R^{2n} \tag{293}$$

The corresponding reduced-order state space model is given by

$$\dot{\tilde{h}} = \tilde{A}\tilde{h} + \tilde{M}_\sigma \sigma + \tilde{M}_q q_M + \tilde{M}_w w, \tag{294}$$

whose matrices $\tilde{A}, \tilde{M}_\sigma, \tilde{M}_q, \tilde{M}_w$ are easily obtained by removing selected rows and columns with index $(2+3k), k = 0, 1, ..., n-1$, from matrix A and rows with the same index from the matrices $M_\sigma, M_q, M_w$ of the full-order model (290).

## 6.5 Flow and infiltration estimation problem statement

In this section we make reference to the linearized dynamics (290) and its reduced-order form (294), and we address two distinct state and disturbance estimation problems under the common constraint that **only level measurements are allowed**. Vector $\sigma$ is supposed to be **known**, while vectors $w$ and $q_M$ are both **unmeasurable**. We cast the next problems to be tackled:

**Problem 1.** By measuring only a portion of the **reduced**-order state vector $\tilde{h}$, and assuming no infiltrations ($w = 0$), asymptotically estimate the flow vector $q_M$ and the unmeasured elements of vector $\tilde{h}$.

**Problem 2.** By measuring the full vector $h$, reconstruct the infiltration vector $w$ and the flow vector $q_M$ in finite time.

Both Problems 1 and 2 will be solved by making use of unknown-input observers (UIO) under the requirement of "strong observability" [10, 15] for certain subsystems that shall be specified later on. An UIO design procedure for general, strongly observable, linear time-invariant systems with unknown inputs (LTISUI) is used in the next Section 6.6. The successive subsections 6.6.1 and 6.6.2 address the Problem 1 and 2, respectively, by exploiting the presented UIO design framework.

## 6.6 Case study and simulation results

In this section we'll apply statements of chapter 2 to solve the problem formulated in section 6.5. The vector state $x$ of chapter 2 is now the vector $h$ whose components are the deviations of the water level from the steady state value.

We shall consider a test canal with rectangular section and three reaches with the next realistic parameters:

- Number of reaches: $n = 3$;

- Lengths: $L_1 = 4000m; L_2 = 5000m; L_3 = 2000m$;

- Widths: $B_1 = B_2 = B_3 = 2m$;

- Discharge coefficient: $\eta = 0.6$;

- Roughness coefficient: $K_s = 50\frac{m^{\frac{1}{3}}}{s}$;

- Slope: $I = 0,001$;

- Water level in upstream reservoir: $H_{B0} = 3m$;

- Water level in downstream reservoir: $H_{A4} = 0.5m$;

- Withdrawals ($\frac{m^3}{s}$): $Q_{C1} = 2,\ Q_{C2} = 2,\ Q_{C3} = 1,$

The opening section of the $4-th$ gate is kept constant :

$$\Sigma_4 = 0.538m^2 \tag{295}$$

The uniform flow condition is characterized by the following values:

- Flow rates ($\frac{m^3}{s}$): $\overline{Q}_1 = 6.017,\ \overline{Q}_2 = 4.007,\ \overline{Q}_3 = 1.966.$

- Levels ($m$): $\overline{H}_1 = 2,40$, $\overline{H}_2 = 1.72$, $\overline{H}_3 = 0.99$

- Opening sections ($m^2$): $\overline{\Sigma}_1 = 2.923,\ \overline{\Sigma}_2 = 1.829,\ \overline{\Sigma}_3 = 0.866;$

The opening sections of the gates 1, 2 and 3 are adjusted according to

$$\begin{array}{ll} \Sigma_1 & = \overline{\Sigma}_1 + 0.8sin[(2\pi/1000)t] \\ \Sigma_2 & = \overline{\Sigma}_2 + 0.5sin[(2\pi/1000)t] \\ \Sigma_3 & = \overline{\Sigma}_3 + 0.3sin[(2\pi/1000)t] \end{array} \tag{296}$$

and the infiltration variables are set as

$$w_1 = w_2 = w_3 = 0.1e^{-0.001t}; \tag{297}$$

Considering the above mentioned parameters, the matrices of the "full order" linearized model (290) turn out to be

$$A = 10^{-3} \begin{bmatrix} -1.9 & 0 & 0.4 & -0.4 & 0 & 0 & 0 & 0 & 0 \\ -0.6 & 0 & -0.4 & 0.4 & 0 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & -1.1 & 1.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.9 & -0.9 & 0 & 0.1 & -0.1 & 0 & 0 \\ 0 & 0 & 0.3 & -0.3 & 0 & -0.1 & 0.1 & 0 & 0 \\ 0 & 0 & -0.3 & 0.3 & 0 & -0.4 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0.3 & -0.3 & 0 & -0.3 \\ 0 & 0 & 0 & 0 & 0 & -0.3 & 0.3 & 0 & -0.8 \end{bmatrix} \tag{298}$$

$$M_\sigma = 10^{-3} \begin{bmatrix} 0.8 & 0.3 & 0 & 0 \\ 0.3 & -0.3 & 0 & 0 \\ -0.3 & -0.8 & 0 & 0 \\ 0 & 0.7 & 0.2 & 0 \\ 0 & 0.2 & -0.2 & 0 \\ 0 & -0.2 & -0.7 & 0 \\ 0 & 0 & 1.7 & 0.5 \\ 0 & 0 & 0.6 & -0.5 \\ 0 & 0 & -0.6 & -1.4 \end{bmatrix} \tag{299}$$

$$M_q = 10^{-4}\begin{bmatrix} -5 & 0 & 0 \\ 0 & 0 & 0 \\ 5 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -10 \\ 0 & 0 & 0 \\ 0 & 0 & 10 \end{bmatrix}, \quad M_w = \begin{bmatrix} 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0.5 \end{bmatrix}, \tag{300}$$

The performance of the observer is tested by means of simulations made in the Matlab-Simulink environment. The system and the observers are integrated by fixed step Runge-Kutta method, with the integration step $T_s = 10^{-4}s$. The actual level error variables are initialized to the value 0.1 and all the observer's initial conditions are set to zero.

The actual $Q_M(t)$ profiles are generated by solving the corresponding system of nonlinear differential equations (291)-(292), with the initial conditions $Q_M(0) = [6.017, 4.007, 1.966]$.

### 6.6.1 Flow estimation with partial level measurements and no infiltration (Problem 1)

Consider the reduced-order linearized dynamics (294) by assuming no infiltration (i.e., $w = 0$) according to the statement of Problem 1 (see Section 4):

$$\dot{\tilde{h}} \;=\; \tilde{A}\tilde{h} + \tilde{M}_\sigma \sigma + \tilde{M}_q q_M, \quad \tilde{h} \in \mathbb{R}^6. \tag{301}$$

The matrices $\tilde{A}$, $\tilde{M}_q$, $\tilde{M}_\sigma$ and $\tilde{M}_w$ of the reduced order model (294) are obtained removing the rows and column with index 2, 5 and 8 from the matrix $A$ and the rows with the same indexes from the matrices $M_q$, $M_\sigma$ and $M_w$ given in (298)-(300). The third element $\tilde{h}_{A2}$ of vector $\tilde{h}$ is supposed to be not measured, according to the next output equation

$$\tilde{y} = [\tilde{h}_{A1}, \tilde{h}_{B1}, \tilde{h}_{B2}, \tilde{h}_{A3}, \tilde{h}_{B3}] = \tilde{C}\tilde{h}, \tag{302}$$

It is worth noting that system (301)-(302) is a special case of the general dynamics (86) with the modified notation $\tilde{h} = x$, $\sigma = u$, $q_M = \xi$.

The state and output transformation matrices take the following form:

$$\tilde{T} = 10\begin{bmatrix} -0.05 & -0.05 & 0.05 & 0.05 & 0 & 0 \\ 0.04 & 0.04 & 0.04 & 0.04 & 0.05 & 0.05 \\ -0.04 & -0.04 & -0.04 & -0.04 & 0.05 & 0.05 \\ -100 & 100 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 250 & 0 & 0 \\ 0 & 0 & 0 & 0 & -50 & 50 \end{bmatrix} \tag{303}$$

$$\tilde{U} = 10 \begin{bmatrix} -0.05 & -0.05 & 0 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0 & 0.05 & 0.05 \\ -100 & 100 & 0 & 0 & 0 \\ 0 & 0 & 250 & 0 & 0 \\ 0 & 0 & 0 & -50 & 50 \end{bmatrix} \tag{304}$$

By computing the matrices of the transformed system dynamics, it can be readily verified that $(\bar{A}_{11}, \bar{C}_1)$ is an observable pair, which confirms that the matrix triplet $(\tilde{A}, \tilde{M}_q, \tilde{C})$ is strongly observable, hence the proposed design method can be applied to reconstruct, both, the unknown vector $q_M$ and the **unmeasured** level variable $h_{A2}$. The suggested scheme (94),(98), (104)-(106), has been implemented with the observer gain matrix

$$\tilde{L} = 10^{-3} \begin{bmatrix} -1.7 & 2.1 \\ -0.6 & -3.0 \\ -3.3 & -0.9 \end{bmatrix} \tag{305}$$

that assign the observation error dynamics the desired spectrum of eigenvalues [-0.05,-0.05,-0.005]. The gain parameters $\alpha$ and $\lambda$ of the unknown input reconstruction algorithm are set according to (106) as

$$\alpha = 1.5\sqrt{5}, \quad \lambda = 5. \tag{306}$$

The next Figures 41 show the actual and estimated profiles of the unknown flow variable $q_{M1}$ during the TEST 1, respectively. The left and right plot focus on the transient and long term behaviour. After a transient of less then twenty seconds, the estimated flow converges towards the actual profile. The estimation performance for the flow variables $q_{M2}$ and $q_{M3}$ is almost equivalent and it is not depicted for brevity. The reconstruction of of the unmeasured level variable $h_{A2}$ is shown in the Figure 31, whose left and right plot show the transient and long term behaviour, respectively.

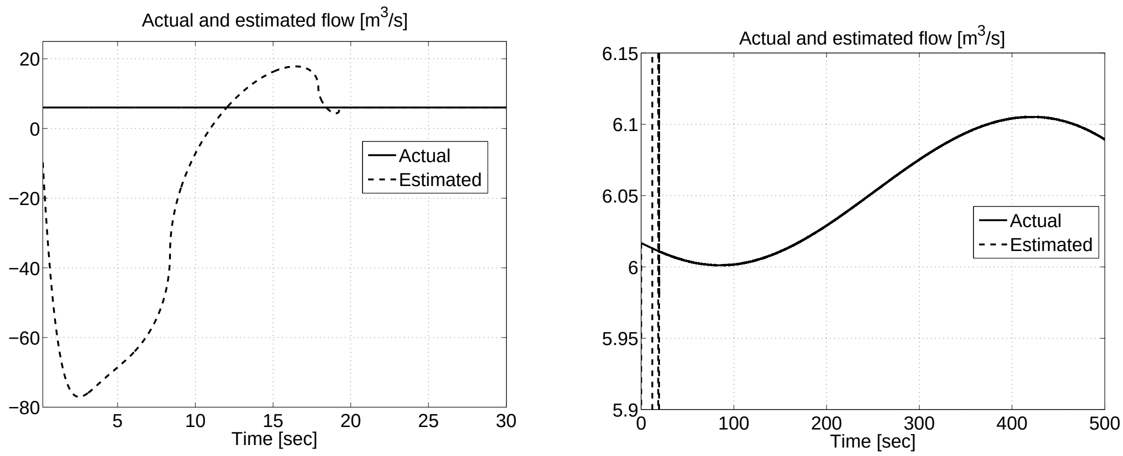



Figure 30: Actual and estimated flow variable $q_{M1}$ in the TEST 1

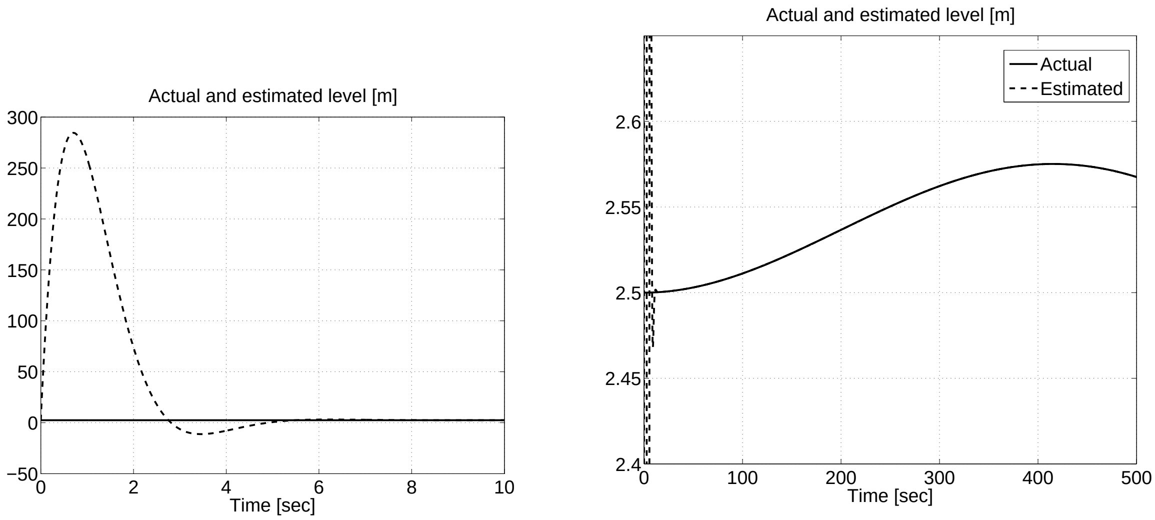



Figure 31: Actual and estimated level variable $h_{A2}$ in the TEST 1

### 6.6.2 Flow and infiltration estimation with full level measurements (Problem 2)

We consider in this section the full-order model (290), reported as follows

$$\dot{h} \quad = \quad Ah + M_\sigma \sigma + M_q q_M + M_w w \tag{307}$$

and we assume the availability for measurement of the full state vector $h$, (i.e. the considered output is $y = h$). The problem here is to reconstruct the unknown flow vector $q_M$ and the infiltration vector $w$ after a **finite** observation transient time. The matrices of system (307) were given in (298)-(300). System (307) along with the considered output equation $y = h$ can be rewritten as

$$\begin{aligned} \dot{h} &= Ah + M_\sigma \sigma + F \cdot \begin{bmatrix} q_M \\ w \end{bmatrix}, \quad F = [M_q \;\; M_w] \\ y &= h \end{aligned} \tag{308}$$

Since the state vector is supposed to be fully available, a simplified version of the design methodology previously described can be applied to reconstruct the unknown vectors $q_M$ and $w$ (see Remark 2). The state and output transformation matrices $T$ and $U$ are now coinciding and taking the form:

$$T = U \quad = \quad \begin{bmatrix} F^\perp \\ \left[F^T F\right]^{-1} F^T \end{bmatrix} \tag{309}$$

which leads to

$$T = \begin{bmatrix} 0.2 & -0.3 & 0.2 & -0.3 & 0.7 & -0.3 & 0.2 & -0.3 & 0.2 \\ 0.2 & -0.5 & 0.2 & -0.1 & 0.1 & -0.1 & -0.3 & 0.7 & -0.3 \\ 0.3 & -0.6 & 0.3 & 0.2 & -0.4 & 0.2 & 0.2 & -0.3 & 0.2 \\ -1000 & 0 & 1000 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.7 & 0.7 & 0.7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1250 & 0 & -1250 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.7 & 0.7 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -500 & 0 & 500 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.7 & 0.7 & 0.7 \end{bmatrix} \qquad (310)$$

By computing the matrices of the transformed system dynamics, it can be readily verified that $(\bar{A}_{11}, \bar{C}_1)$ is an observable pair, that confirms that the matrix triplet $(A, [M_q, M_w], I)$ is strongly observable.

Since the state vector is already available, only the observer (110), (104)-(106),(100),(103) should be implemented for reconstructing the unknown input, which provides the **finite time estimation** condition (111). The gain parameters $\alpha$ and $\lambda$ of the observation algorithm are set as

$$\alpha = 1.5\sqrt{5}, \qquad \lambda = 5. \qquad (311)$$

The next Figures 32 shows the actual and estimated profiles of the unknown flow variable $q_{M2}$ during the TEST 2, of duration 100 seconds. The left and right plot show the transient and long term behaviour. After a transient of about half a second, the estimated flow converges towards the actual one. The estimation performance for the flow variables $q_{M1}$ and $q_{M3}$ is almost equivalent and it is not shown for brevity. The reconstruction of of the unknown infiltration

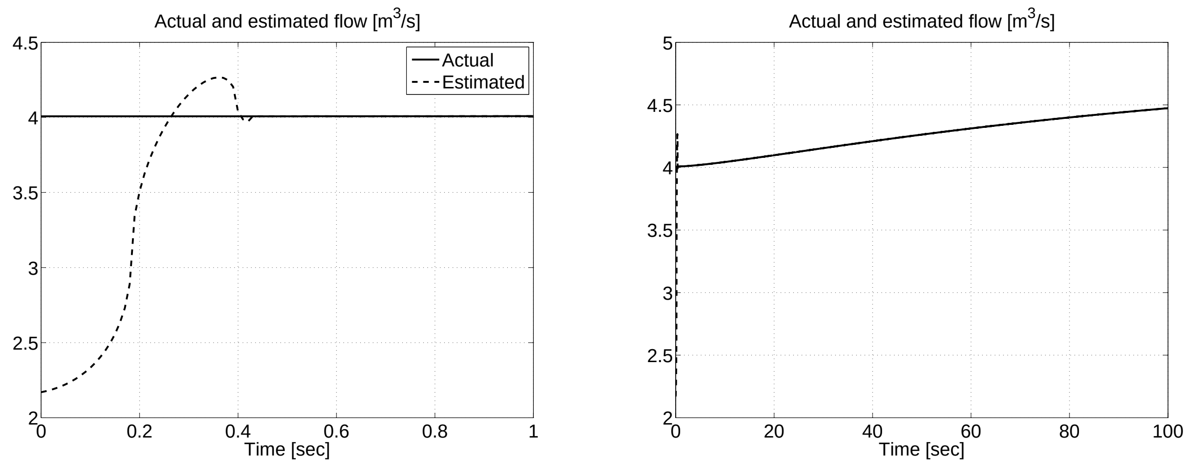



Figure 32: Actual and estimated flow variable $q_{M2}$ in the TEST 2

variable $w_3$ is shown in the Figure 33. The left and right plot show the transient and long term behaviour, respectively.

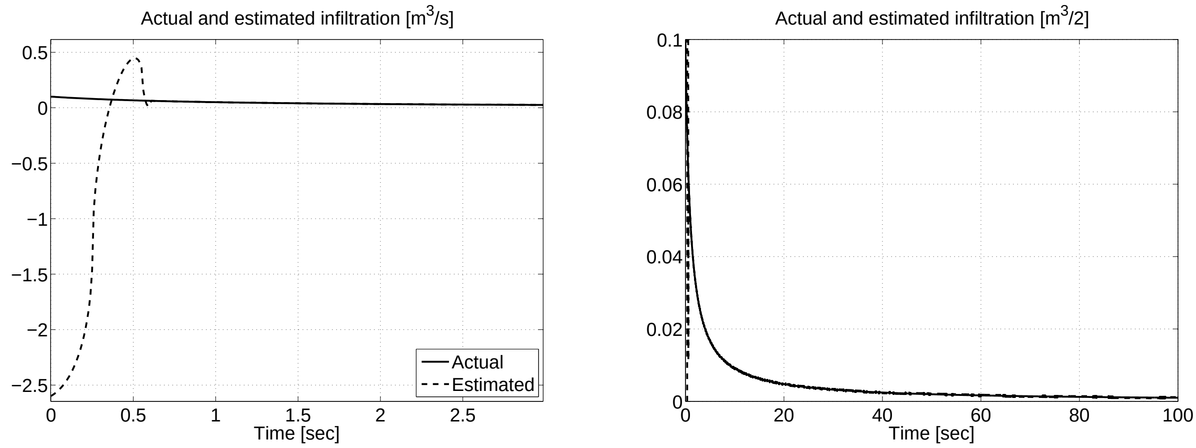



Figure 33: Actual and estimated infiltration variable $w_3$ in the TEST 2

## 6.7 Tests using high accuracy solvers

In the next series of tests (TEST 3), the level measurements have been generated by solving the original Saint Venant system of PDEs using the Preissmann implicit finite-difference solution scheme. Implicit schemes which can use large time steps without any stability problem are in fact more widely applied than more traditional finite difference solvers. The Preissmann scheme is the most widely applied implicit finite difference method because of its simple structure with both flow and geometrical variable in each grid point. Following the description of the method made in [2]) , the solution and its spatial and temporal derivatives are approximated by means of the next expressions

$$f(x,t) = \theta\left[\phi f_{j+1}^{k+1} + (1-\phi) f_j^{k+1}\right] + (1-\theta)\left[\phi f_{j+1}^{k} + (1-\phi) f_j^{k}\right] \tag{312}$$

$$\frac{\partial f(x,t)}{\partial t} = \phi \frac{f_{j+1}^{k+1} - f_{j+1}^{k}}{\Delta t} + (1-\phi)\frac{f_j^{k+1} - f_j^{k}}{\Delta t} \tag{313}$$

$$\frac{\partial f(x,t)}{\partial x} = \theta \frac{f_{j+1}^{k+1} - f_{j}^{k+1}}{\Delta x} + (1-\theta)\frac{f_{j+1}^{k} - f_j^{k}}{\Delta x} \tag{314}$$

where $f(x,t)$ is the hydraulic variable of concern (water level or discharge), $\Delta t$ and $\Delta x$ are the time and space discretization steps, $f_j^k = f(j\Delta x, k\Delta t)$ and $\theta, \phi$ are weighting coefficients. Parameters $\phi$ and $\theta$ were both set to 0.5 in our resolution model. The time step was set as $0.1s$, and the space discretization step was chosen separately for each canal in order to have ten spatial solution nodes per canal.

In the TEST 4, the performance of the suggested scheme has been compared with that of an extended Kalman filter that considers the unmeasurable infiltrations as constant parameters described by a fictitious dynamic relations $\dot{w}_i = 0$. The resulting nonlinear dynamic model has

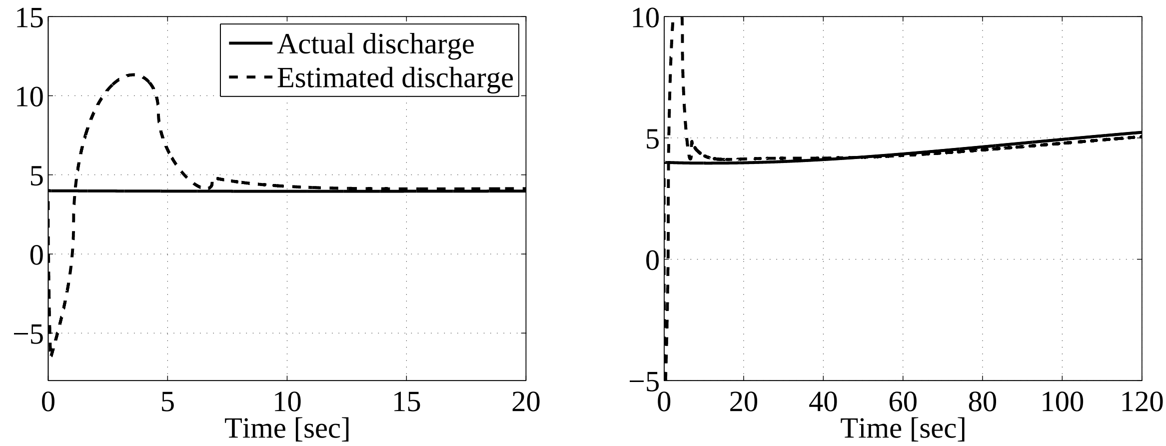



Figure 34: Actual and estimated flow variable $q_{M2}$ in the TEST3. Left plot: zoom on transient. Right plot: long term behaviour.

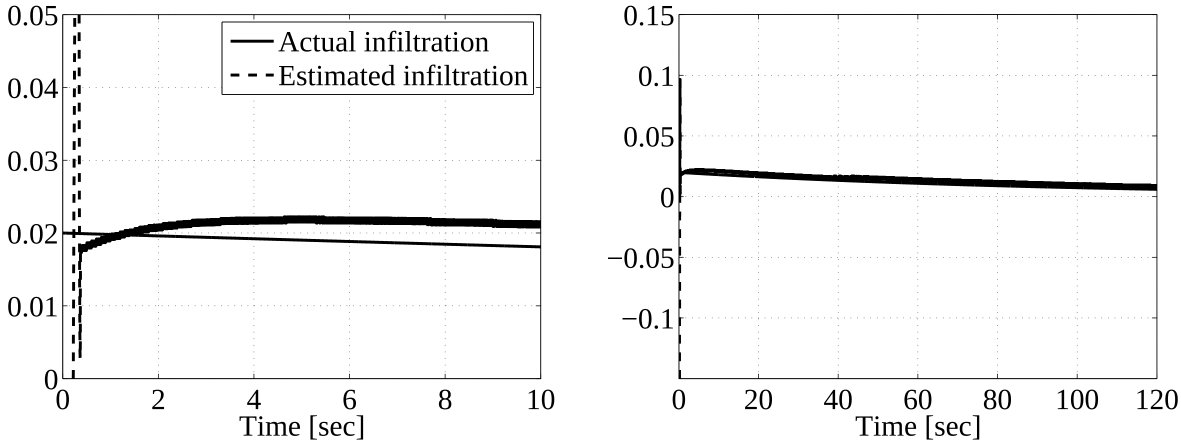



Figure 35: Actual and estimated infiltration variable $w_3$ in the TEST3. Left plot: zoom on transient. Right plot: long term behaviour.

been constructed by properly augmenting (275) and (276)-(277). The observer gain matrix has been selected coinciding with the constant matrix of the steady Kalman filter which considers the linearized approximation of the underlying dynamics. The implemented EKF has been supplied with the level measurements provided by the Preissmann solution scheme.

As shown in the figure 36, the flow and infiltration reconstruction performance are acceptable but slightly less accurate as compared to the previously obtained results with the proposed methodology. We do not claim, however, any superiority of our method over the EKF as this would require much deeper investigations. It can be just commented that our method is provably robust against the presence of time varying unmeasurable infiltration and discharges, while effectiveness of the EKF relies on their "slowly varying" nature.

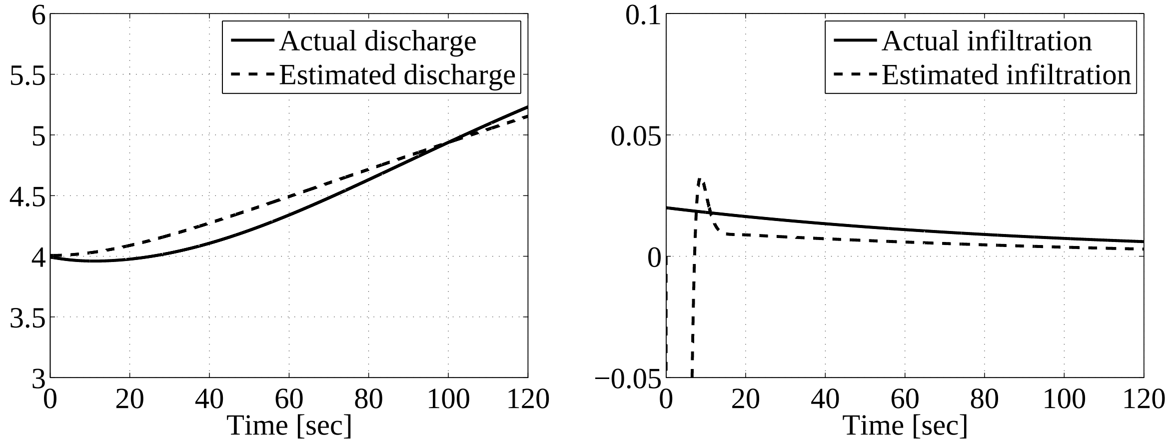



Figure 36: Actual and estimated flow variable $q_{M2}$ (left plot) and infiltration variable $w_3$ (right plot) in the TEST4

In the conclusive TEST 5, parameter uncertainty has been introduced by randomly corrupting all the parameters of the model so as to introduce a maximal percentage error of 10%. This has been done by multiplying each parameter by a random coefficient in the interval $[0.9, 1.1]$. A different coefficient for each model parameter has been used. Within the present TEST 5, the level measurements computed using the Preissmann solution scheme were used. Figure 37 shows the flow and infiltration reconstruction performance, which is deteriorated, as compared to the results of TEST 3 where no parameter mismatches were considered, but keep relatively acceptable.

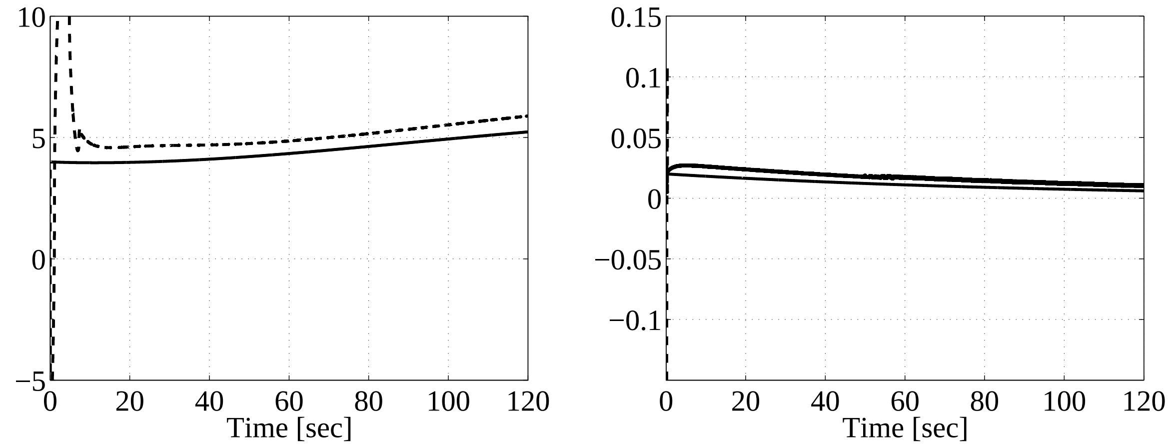



Figure 37: Actual and estimated flow variable $q_{M2}$ (left plot) and infiltration variable $w_3$ (right plot) in the TEST5

## 6.8 Conclusion

A linear UIO and a nonlinear sliding mode disturbance observer have been combined to reconstruct water level, discharge and infiltration variables in open channel irrigation canals connected in cascade and subject to unknown time-varying infiltrations. The underlying, collocation based, nonlinear dynamics are linearized around the subcritical uniform flow condition. The UIO and DO design procedures are constructively illustrated along this chapter. Simulation results using realistic data are discussed to verify the effectiveness of the proposed schemes.

The linearization of the model could be possibly avoided by generalizing the the strong-observability based UIO and SMDO design method to the nonlinear case. This, however, needs further investigation. Other interesting tasks for next research are the decentralization of the schemes (e.g. by consensus-based methodologies), and/or their use to address observer-based controller design problems. Preliminary results were recently presented (see [5]).

# References


[1] Bedjaoui, N., X. Litrico, D. Koenig and P.O. Malaterre (2006). $H_\infty$ observer for time-delay systems Application to FDI for irrigation canals *Proc. of 45th IEEE Conf. Dec. Contr. (CDC '06)*, 532-537, San Diego, US

[2] Begovic, O., Felipe, J.C. and Ruiz (2007). V.M."Real-time implementation of a decentralized control for an open irrigation canal prototype." *Asian Journal of Control*, 9(2), 170-179.

[3] F.J. Bejarano, L. Fridman and A. Poznyak (2007). "Exact state estimation for linear systems with unknown inputs based on hierarchical super-twisting algorithm". *Int. J. Robust Nonlinear Control*, 17(18),1734-1753.

[4] Besanon, G.,Dulhoste, J.-F. and Georges, D. (2001). "Non-linear observer design for water level control in irrigation canals." *Proc. of 40th IEEE Conf. Dec. Contr. (CDC '01)*, 4968-4973.

[5] Besanon, G.; Dulhoste, G.F. and Georges, D.(2008)."Nonlinear Observer-Based Feedback for Open-Channel Level Control." *Journal of Hydraulic Engineering*, 134(9), 1267-1274.

[6] Colley, R.L. and Moin, S.A. (1976). "Finite element solution of Saint-Venant equations", *Journal of Hydraulical Engineering. Division ASCE,* 102(6), 759-775.

[7] Corriga,G., Sanna,S., Usai,G.(1983). "Sub-optimal constant volume control for open-channel networks", *Applied Mathematical Modelling*, 7(4), 262-267.

[8] Corriga,G., Sanna,S., Usai,G.(1989). "Estimation of uncertainty in an open-channel network mathematical model", *Applied Mathematical Modelling*, 13(11), 651-657.

[9] D. Dochain, J.P. Babary and N. Tali-Maamar(1992). "Modelling and adaptive control of nonlinear distributed parameter bioreactors via orthogonal collocation", *Automatica*, 28(5), 873-883.

[10] Dulhoste, J.-F., Besanon, G., and Georges, D. (2001). "Non-linear control of water flow dynamics by inputoutput linearization based on a collocation method model". *European Control Conf.*

[11] Edwards, C., Spurgeon, S. K., and Patton, R.J.(2000). "Sliding mode observers for fault detection and isolation", *Automatica*, 36, 541-553.

[12] Euren, K., and E. Weyer (2007).System identification of irrigation channels with undershot and overshot gates. *Control Engineering Practice*, 15(7), 813-824.

[13] Fletcher C.A.J, (1984). "Computational Galerkin methods". *Springer Series in Computational Physics.*

[14] Fridman, L., Levant, A. and Davila, J.(2007). "Observation of Linear Systems with Unknown Inputs via High-Order Sliding-Mode", *International Journal of Systems Science*, 38(10), 773-791.

[15] Hautus, M. L. J.(1983). "Strong detectability and observers" Linear Algebra and its Applications, 50, 353-368.

[16] Levant, A.(1993)." Sliding order and sliding accuracy in sliding mode control", *Int. J. Contr.*, 58, 1247-1263.

[17] Li, Y., Cantoni, M and Weyer, E. (2005)."On Water-Level Error Propagation in Controlled Irrigation Channels." *Proc of the 44th IEEE Conference on Decision and Control, and the European Control Conference Seville*, 2101-2016.

[18] Litrico, X. and Fromion, V.(2004)."Simplified Modeling of Irrigation Canals for Controller Design", *Journal of irrigation and drainage engineering*, 130(5), 373-383.

[19] Litrico, X., Fromion, V., Baumea, J.-P., Arranja, C., and Rijo, M.(2005). "Experimental validation of a methodology to control irrigation canals based on Saint-Venant equations" *Control Engineering Practice*, 13(11), 1425-1437.

[20] Litrico, X., Fromion, V., (2009). Modeling and Control of Hydrosystems, Springer, Berlin. ISBN 978-1-84882-623-6.

[21] Mareels, I., Weyer, E., Ooi, S. K., Cantoni, M., Li, Y. and Nair,G. (2005). "Systems engineering for irrigation systems: Successes and challenges". *Annual Reviews in Control*, 29, 191-204.

[22] Molina, L.S., and Miles, J.P.(1996)."Control of an irrigation canal". *Journal of Hydraulic Engineering*, 403-410.

[23] Ooi, S.K., Krutzen, M. P. M. and Weyer, E.(2005). "On physical and data driven modelling of irrigation channels". *Control Engineering Practice* 13(4), 461-471.

[24] Pillosu, S., Pisano, A., Usai, E.(2011a). " Unknown-input observation techniques in Open Channel Hydraulic Systems". 2011 IFAC World Congress, September 2011, Milan (Italy)

[25] Pillosu, S., Pisano, A., Usai, E.(2011b). " Unknown-input observation techniques for Infiltration and Water Flow Estimation in Open-Channel Hydraulic Systems ". In press in the Control Engineering Practice.

[26] J. Mohan Reddy (1995). "Kalman filtering in the control of irrigation canals". Appl. Math. Modelling, 19,201-209.

[27] Strelkoff, T (1970). "Numerical solution of Saint-Venant equation", Journal of Hydraulical Engineering. Division ASCE, 96(1), 223-252.

[28] Strub, I, and Bayen, A (2006). "Weak formulation of the boundary condition for scalar conservation laws: an application to highway traffic modeling", Int. J. Rob. Nonlin. Contr., 16, 733-748.

[29] Villadsen J.V., and Michelsen M.L.,(1978), "Solution of differential equations models by polynomial approximation". *Prentice Hall. Englewood Cliffs, N.J.*.

[30] Weyer, E., Bastin, G.(2008), "Leak detection in open water channels", *Proc. of the 17th IFAC World Congress*, 7913-7918.

[31] Zhuan, X. and Xia, X.(2007), "Models and control methodologies in open water flow dynamics: A survey", *IEEE Africon, Windhoek, Namibia.*

# 7 Fault Detection and Compensation in Open-Channel Hydraulic Systems

## 7.1 Introduction

The main aim of this chapter is the development of a new actuator fault-detection algorithm for a cascade of canal reaches with control gates. The possibility of compensating the faults is also investigated. By measuring level variables only, and by assuming a negligible infiltration, we address the following tasks

i.) finite-time reconstruction of additive gate actuator faults;

ii.) feedback compensation of the faults in the control gates.

iii.) finite-time estimation of the system discharge variables.

In [24], the problem of estimating the discharge and the (possibly time-varying) infiltration was addressed by a UIO and sliding mode scheme conceptually equivalent to the one presented in this chapter. In the above work, however, faults in the control gates were not allowed.

This chapter is organized as follows. Section 7.2 recalls the collocation-based model for the considered cascade of $n$ canal reaches. In Section 7.3 the estimation and fault detection problem is formulated and a methodology for the finite-time estimation of the discharge variable and reconstruction of additive faults affecting the control gates is described, along with its application for compensating the effect of the faults. In Section 7.4 a case study of a canal with rectangular section and three reaches is illustrated, and the parameters of the corresponding linearized model are computed on the basis of realistic input data. Simulation results are discussed. Section 7.5 states some final conclusion and draws possible lines for next researches.

## 7.2 Open channel hydraulic system modeling

Under the assumption of negligible infiltration, water flow dynamics in an open channel are governed by the system of Saint-Venant partial differential equations [2], as we saw in the previous chapter.

Let us recalled the linearized dynamics of the deviation level variables, stated in chapter 6:

$$
\begin{array}{rcl}
\dot{h}_{Ai} & = & \frac{1}{B_i L_i}\{-4q_{Mi} + 3a_i\sigma_i(t) + 3b_i[h_{Bi-1}(t) - h_{Ai}(t)] \\
 & + & a_{i+1}\sigma_{i+1}(t) + b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)]\} \\
\dot{h}_{Mi} & = & \frac{1}{B_i L_i}\{a_i\sigma_i(t) + b_i[h_{Bi-1}(t) - h_{Ai}(t)] \\
 & - & a_{i+1}\sigma_{i+1}(t) - b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)]\} \\
\dot{h}_{Bi} & = & \frac{1}{B_i L_i}\{4q_{Mi} - a_i\sigma_i(t) - b_i[h_{Bi-1}(t) - h_{Ai}(t)] \\
 & - & 3a_{i+1}\sigma_{i+1}(t) - 3b_{i+1}[h_{Bi}(t) - h_{Ai+1}(t)]\}
\end{array} \tag{315}
$$

Now, let us assume that the control gate deviation variables $\sigma_i$ are composed of a nominal, desired, value $\sigma_i^{des}$ (normally zero) and an additive unmeasurable fault signal $\sigma_i^{fault}$, i.e.

$$
\sigma_i = \sigma_i^{des} + \sigma_i^{fault} \tag{316}
$$

Defining vectors

$$
h = [h_{A1}\ h_{M1}\ h_{B1}\ h_{A2}\ ...h_{An}\ h_{Mn}\ h_{Bn}]^T, \quad h \in \mathbb{R}^{3n} \tag{317}
$$

$$
q_M = [q_{M1}\ q_{M2}...q_{Mn}]^T, \quad q \in \mathbb{R}^n \tag{318}
$$

$$
\sigma^{des} = [\sigma_1^{des}\ \sigma_2^{des}...\sigma_{n+1}^{des}]^T, \quad \sigma^{des} \in \mathbb{R}^{n+1} \tag{319}
$$

$$
\sigma^{fault} = [\sigma_1^{fault}\ \sigma_2^{fault}...\sigma_{n+1}^{fault}]^T, \quad \sigma^{fault} \in \mathbb{R}^{n+1} \tag{320}
$$

it is possible to rewrite the linearized system (315) in state space form

$$
\dot{h} = Ah + M_q q_M + M_\sigma \sigma^{des} + M_\sigma \sigma^{fault} \tag{321}
$$

with implicitly defined constant matrices $A$, $M_\sigma$ and $M_q$ of appropriate dimension. Vector $q_M$ is governed by the nonlinear differential equation

$$
\dot{q}_M = \psi(h, q_M, \sigma) \tag{322}
$$

where

$$
\psi(h, q_M, \sigma) = [\psi_{1q}(\cdot),\ \psi_{2q}(\cdot),\ \dots\ , \psi_{nq}(\cdot)]^T \tag{323}
$$

and the functions $\psi_{iq}(\cdot)$ $(i = 1, 2, ..., n)$ were given in (277). Note that the nonlinear dynamics (322)-(323) **need not to be linearized** since vector $q_M$ is going to be treated as an unknown input to the system, rather than as a part of the system state, hence the relation (322) is only used for numerical simulation purposes and the underlying parameters of the nonlinear functions $\psi_{iq}(\cdot)$ in the right hand side of definition (323) can be therefore assumed uncertain.

## 7.3 Problem statement and its solution

In this chapter we make reference to the linearized dynamics (321) of the deviation variables, which can be rewritten as

$$
\dot{h} = Ah + M_\sigma \sigma^{des} + F\xi, \qquad F = [M_q\ \ M_\sigma] \tag{324}
$$

where vector

$$\xi = \begin{bmatrix} q_M \\ \sigma^{fault} \end{bmatrix} \in \mathbb{R}^{2n+1} \tag{325}$$

collects all the **unknown** inputs to the system. We make the next Assumptions

**Assumption 1** *The state vector $h$ is measured, and the matrices $A$, $M_q$, $M_\sigma$ of system (324) are known.*

**Assumption 2** *Vector $\sigma^{des}$ is supposed to be known, while vectors $\sigma^{fault}$ and $q_M$ are both unmeasurable.*

**Assumption 3** *Matrix $F$ is full rank.*

**Assumption 4** *There exist known a-priori constants $\Sigma_d$ and $Q_{Md}$ such that*

$$\|\dot{\sigma}^{fault}\|_1 \leq \Sigma_d^{fault}, \qquad \|\dot{q}_M\|_1 \leq Q_{Md} \tag{326}$$

*where $\|\cdot\|_1$ denotes the 1-norm of the respective vector argument.*

The task is to reconstruct the fault vector $\sigma^{fault}$ and the discharge variables vector $q_M$ in finite time under the above statements and assumptions.

For the generic matrix $\Delta \in \mathbb{R}^{n_r \times n_c}$ with $rank\ \Delta = r$, we define $\Delta^\perp \in \mathbb{R}^{n_r - r \times n_r}$ as a matrix such that $\Delta^\perp \Delta = 0$ and $rank \Delta^\perp = n_r - r$. Matrix $\Delta^\perp$ always exists and, furthermore, it is not unique. Let $\Delta^+ = [\Delta^T \Delta]^{-1} \Delta^T$ denote the left pseudo-inverse of $\Delta$ such that $\Delta^+ \Delta = I_{n_c}$, with $I_{n_c}$ being the identity matrix of order $n_c$.

Define

$$\bar{h}_2 \;\; = \;\; F^+ h, \qquad \bar{h}_2 \in \mathbb{R}^{2n+1}; \tag{327}$$

The next unknown-input estimator is proposed

$$\dot{z}_2 = F^+ h + F^+ M_\sigma \sigma^{des} + v_1 + v_2 \tag{328}$$

with the output injection terms $v_1$ and $v_2$ defined as

$$v_1 \;\; = \;\; k_1(\bar{h}_2 - z_2) + k_2 \left|\bar{h}_2 - z_2\right|^{1/2} sign\ (\bar{h}_2 - z_2) \tag{329}$$

$$\dot{v}_2 \;\; = \;\; k_3\ sign\ (\bar{h}_2 - z_2), \qquad v_2(0) = 0. \tag{330}$$

where $k_1, k_2, k_3$ are proper design tuning constants. The convergence properties of the above unknown-input estimator are established in the next

**Theorem 15** *Consider system (324), satisfying the Assumptions 1-4. and the estimator (328)-(330), based upon the transformations (327), and with the gain parameters $k_1, k_2, k_3$ satisfying the next inequalities*

$$k_1 > 2\sqrt{\rho} \qquad k_2 > 0 \qquad k_3 > \rho\sqrt{k_1} \qquad \rho = \Sigma_d^{fault} + Q_{Md} \tag{331}$$

*Then there is a finite $T^*$ such that exact reconstruction of the unknown input vector $\xi$ is achieved by means of the proposed estimator according to*

$$v_2(t) = \xi(t), \qquad \forall t \geq T^*. \tag{332}$$

**Proof of Theorem 15**

Consider the error variable $\epsilon = \bar{h}_2 - z_2$. By considering the plant and observer dynamics (324), (328)-(330), the error time derivative takes the form

$$\dot{\epsilon} = \xi(t) - v_1(t) - v_2(t) = \xi(t) - k_1\epsilon - k_2|\epsilon|^{1/2}\text{sign}(\epsilon) - v_2(t) \tag{333}$$

By making the change of variable

$$\pi(t) = \xi(t) - v_2(t) \tag{334}$$

one can augment and rewrite (333) as

$$\begin{aligned} \dot{\epsilon} &= -k_1\epsilon - k_2|\epsilon|^{1/2}\text{sign}(\epsilon) + \pi(t) && (335)\\ \dot{\pi} &= -k_3\text{sign}(\epsilon) + \frac{d}{dt}\xi(t) && (336) \end{aligned}$$

Stability of (335)-(336) is more easily analyzed by considering separately the decoupled second-order dynamics

$$\begin{aligned} \dot{\epsilon}_i &= -k_1\epsilon_i - k_2|\epsilon_i|^{1/2}\text{sign}(\epsilon_i) + \pi_i(t) && (337)\\ \dot{\pi}_i &= -k_3\text{sign}(\epsilon_i) + \frac{d}{dt}\xi_i(t) && (338) \end{aligned}$$

where the subindex $i = 1, 2, ..., 2n+1$ denotes the $i$-th entry of the corresponding vector. The derivative term $\frac{d}{dt}\xi_i(t)$ can be bounded as follows by virtue of the Assumption 4

$$\left|\frac{d}{dt}\xi_i(t)\right| \leq \Sigma_d^{fault} + Q_{Md} \tag{339}$$

Stability of a class of systems including the dynamics (337)-(339) above was already investigated in the literature (cfr. [3], Th. 5), where, particularly, the global finite time stability of the system was demonstrated by means of the radially-unbounded non-smooth Lyapunov function

$$V_i = \xi_i^T H \xi_i, \qquad i = 1, 2, ..., 2n+1 \tag{340}$$

$$\xi_i = \begin{bmatrix} |\epsilon_i|^{1/2}\text{sign}(\epsilon_i) \\ \epsilon_i \\ \pi_i \end{bmatrix} H = \begin{bmatrix} (4k_3 + k_2^2) & k_1k_2 & -k_2 \\ k_1k_2 & k1^2 & -k_1 \\ -k_2 & -k_1 & 2 \end{bmatrix} \tag{341}$$

where $\epsilon_i$ and $z_i$ denote the $i$-th entry of vector $\epsilon$ and $z$,respectively. It turns out after the appropriate computations (cfr. [3], Proof of Th. 5) that the tuning conditions (331) imply the existence of two positive constants $\gamma_1$ and $\gamma_2$ such that

$$\dot{V}_i \leq -\gamma_1 V_i - \gamma_2\sqrt{V_i}, \qquad \gamma_1, \gamma_2 > 0 \tag{342}$$

The inequality (342) guarantees the global **finite time** convergence of $V_i$ to zero, and, hence, $\epsilon_i$ and $\pi_i$ to zero starting from some finite moment $T_i^* > 0$. By the definition (334) of $\pi(t)$, it directly follows the condition (332) with

$$T^* = \max_i T_i^* \tag{343}$$

Theorem 15 is proven. □

It follows from Theorem 15 that vector $v_2$, which after a finite time matches the unknown input vector $\xi$, should be understood in terms of the decomposition

$$v_2 = \begin{bmatrix} \hat{q}_M \\ \hat{\sigma}^{fault} \end{bmatrix}, \qquad \hat{q}_M \in \mathbb{R}^n, \qquad \hat{\sigma}^{fault} \in \mathbb{R}^{n+1} \tag{344}$$

Thus the first $n$ entries of vector $v_2$ represent the estimates of the discharge variables at the middle points, and the successive $n+1$ elements represent the estimates of the actuator additive faults. Theorem 15 ensures the finite time convergence of both the estimates $\hat{q}_M$ and $\hat{\sigma}^{fault}$ towards the corresponding actual variables $q_M$ and $\sigma^{fault}$, respectively.

### 7.3.1 Compensation of the actuator faults

The effect of nonzero actuator faults is that of steering the system away from the uniform flow condition. We exploit the possibility of compensating the detected faults by adjusting the desired reference opening deviations $\sigma^{des}$ in such a way that the effect of the fault is canceled, i.e.

$$\sigma^{des} = -\hat{\sigma}^{fault} \tag{345}$$

where $\hat{\sigma}^{fault}$ is the estimated fault provided by the previously proposed scheme. The compensated dynamics is

$$\dot{h} = Ah + M_\sigma(\sigma^{fault} - \hat{\sigma}^{fault}) + M_q q_M \tag{346}$$

The unknown input estimator (328) being specified with the relation (345), it turns out that Theorem 15 is still is force. In fact, $\sigma^{des}$ cancels out from the error time derivative (333), and thus the convergence properties of Theorem 15 hold irrespectively of $\sigma^{des}$. Thus the convergence $\hat{\sigma}^{fault} \rightarrow \sigma^{fault}$ is achieved within a finite time interval, after which the compensated system becomes

$$\dot{h} = Ah + M_q q_M \tag{347}$$

being insensitive to the faults.

## 7.4 Case study and simulation results

We shall consider a test canal with rectangular section and three pools with the realistic parameters of the previous chapter:

The opening sections of the four gates are selected as

$$\Sigma_i \quad = \overline{\Sigma}_i + \sigma_i^{des} + \sigma_i^{fault} \tag{348}$$

The additive fault vector is set as

$$\begin{aligned} \sigma_1^{fault} &= \sigma_3^{fault} = \sigma_4^{fault} = 0; && (349) \\ \sigma_2^{fault} &= \begin{cases} 0 & t < 250 \\ 0.5 - 0.5cos\left(\frac{2\pi}{1000}(t-250)\right) & t \geq 250 \end{cases} && (350) \end{aligned}$$

and the desired deviations $\sigma_i^{des}$ preliminarily set to zero.

Considering the above mentioned parameters, the matrices of the linearized model (324) turn out to be the same of Chapter 6 (298)-(299)-(300).

Matrix $F^+$ is given by

$$F^+ = \begin{bmatrix} -624.9 & 3041.4 & 1375.1 & -0.3 & -0.2 & -0.3 & -10.3 & 18.5 & -10.3 \\ 1405.7 & -2342.9 & 1405.7 & -2953.3 & -1362.6 & -453.3 & 42.0 & -75.6 & 42.0 \\ 1438.5 & -2397.6 & 1438.5 & -3487.4 & -2789.9 & -3487.4 & -1083.0 & -850.5 & -83.0 \\ 625.1 & 2291.5 & 625.1 & -0.2 & -0.1 & -0.2 & -6.1 & 11.0 & -6.1 \\ 624.9 & -1041.5 & 624.9 & 0.2 & 0.2 & 0.2 & 7.1 & -12.8 & 7.1 \\ 624.6 & -1041.0 & 624.6 & -1514.2 & -1211.4 & -1514.2 & 30.2 & -54.4 & 30.2 \\ 758.1 & -1263.5 & 758.1 & -1837.9 & -1470.3 & -1837.9 & -650.3 & -829.4 & -650.3 \end{bmatrix} \tag{351}$$

The performance of the observer is tested by means of simulations made in the Matlab-Simulink environment. The system and the observers are integrated by fixed step Runge-Kutta method, with the integration step $T_s = 10^{-3}s$. The actual level error variables are initialized to the value 0.1 and all the observer's initial conditions are set to zero.

The actual $Q_M(t)$ profiles are generated by solving the corresponding system of nonlinear differential equations (322)-(323), with the initial conditions $Q_M(0) = [6.017, 4.007, 1.966]$.

The gain parameters $\alpha$ and $\lambda$ of the observation algorithm are set as

$$k_1 = 4.24, \quad k_2 = 1, \quad k_3 = 9.27. \tag{352}$$

The next Figure 38 shows the actual and estimated profile of the unknown fault variable $\sigma_2^{fault}$ during the TEST, of duration 750 seconds. The reconstruction of of the unknown flow variable

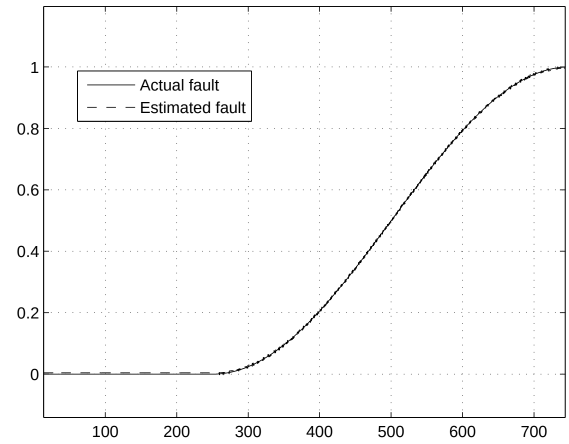



Figure 38: Actual and Estimated fault of $\sigma_2^{fault}$

$Q_{M2}$ is demonstrated in the Figure 39.

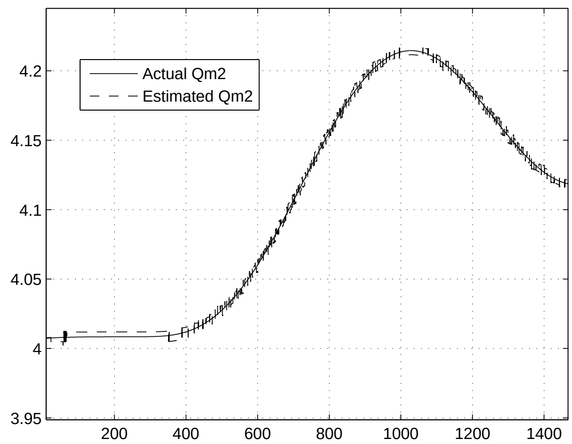



Figure 39: Actual and Estimated Flow Variable $Q_{M2}$

After the appearance of the fault, the uniform flow condition is lost.

By using the compensation strategy (345) the capability of reconstructing the fault is maintained, as shown in Figure (40), The system operating point now stays in the uniform regime condition also after the appearance of the fault, as shown in Figure (41).

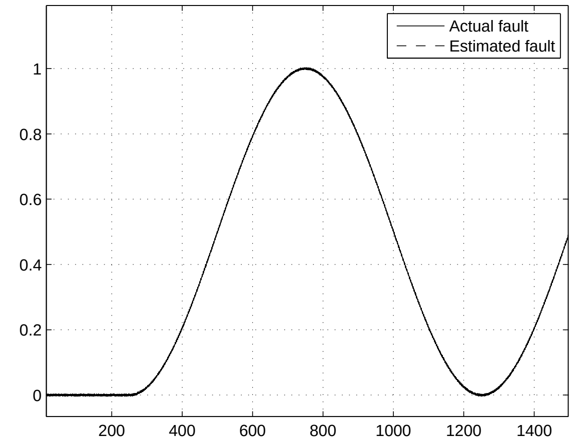



Figure 40: Actual and Estimated fault of $\sigma_2^{fault}$ after the application of compensation strategy

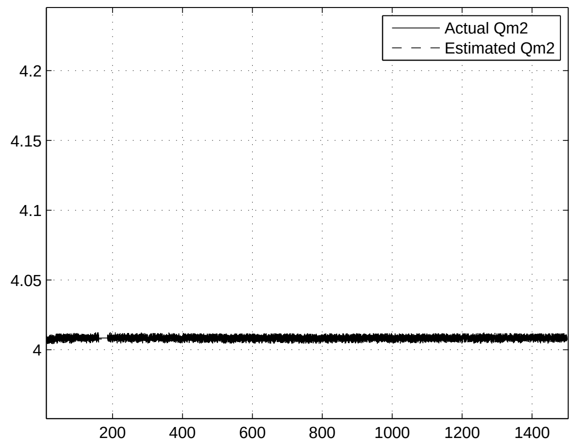



Figure 41: Actual and Estimated flow variable $Q_{M2}$ after the application of compensation strategy

## 7.5 Conclusion

A linear UIO and a nonlinear sliding mode disturbance observer have been combined to reconstruct discharge and faults of gate variables in open channel irrigation canals connected in cascade. The underlying, collocation based, nonlinear dynamics are linearized around the subcritical uniform flow condition. The UIO and DO design procedures are constructively illustrated along this chapter. Simulation results using realistic data are discussed to verify the effectiveness of the proposed schemes.

The linearization of the model could be possibly avoided by generalizing the the strong-observability based UIO and SMDO design method to the nonlinear case. This, however, needs further investigation. Other interesting tasks for next research are the decentralization of the schemes (e.g. by consensus-based methodologies), and/or their use to address observer-based controller design problems.

## References


[1] Koenig, D., Bedjaoui, N. and Litrico, X.(2005)."Unknown Input Observers Design for Time-Delay Systems Application to An Open-Channel." *Conference on Decision and Control, and the European Control Conference*, 12-15.

[2] Litrico, X., Fromion, V., (2009). Modeling and Control of Hydrosystems, Springer, Berlin. ISBN 978-1-84882-623-6.

[3] Moreno, J., Osorio, M (2008). "A Lyapunov approach to second-order sliding mode controllers and observers". *Proc of the 47th IEEE Conference on Decision and Control CDC 2008, Cancun*, 2101-2016.

[4] Pillosu, S., Pisano, A., Usai, E.(2011). " Actuator Fault Detection and Compensation in Open-Channel Hydraulic Systems". In press in the Asian Journal of Control.

[5] Pillosu, S., Pisano, A., Usai, E.(2011). " Decentralized State Estimation for Linear Systems with Unknown Inputs. A Consensus-based Approach". *IET Journal of Control Theory and Applications*, vol. 5, n. 3, 498-506.

# 8 Conclusions

The first part of the thesis, Chapter 1 - 3, has been devoted to introduce some indispensable concepts inherent the Unknown Input Observers and Consensus Algorithm. The main method for estimation in presence of Unknown Input have been described with particular attention to the concept of Strong Observability, well known in literature. Chapter 4 provides a solution of the problem of the decentralized estimation using Consensus Algorithm with Distributed Kalman Filters. Particular attention has been given to the connectivity properties for a system of subsystems to be stabilizable. In order to be complete from the optimization point of view, the chapter focuses on those techniques that enable the choice of consensus gains by solving a Lyapunov equation.

In Chapter 5 a new algorithm for the decentralized estimation in presence of deterministic Unknown Input was presented. It was proved that in the case in which the property of Strong Observability for the subsystems and Quasi-Strongly Connectiveness for the Interconnection Graph, it is possible reconstruct both state and unknown input. The procedure was illustrated by same numerical illustration.

In Chapter 6, we have focused our attention into the problem of state estimation in Open Channel Hydraulic System. Firstly, considering a single reach canal, with the measure of the flow up stream and downstream, supposing a constant infiltration, an algorithm for the estimation of the flow in the middle of the canal and da infiltration is proposed. After, still in chapter 6, we have the same problem but with more stringent condition: a cascade of n-canal reach is considered, with no flow measurements and with time varying infiltration. Here the theory of UIO has helped us to discover new algorithm in which is possible to estimate flow and infiltration. We have also showed how is possible to reconstruct unmeasured water levels in the case of absence of infiltration. Some numerical example has been proposed and simulation results have been shown.
In Chapter 7, in Open Channel Hydraulic Systems, the UIO theory has been used to design the reconstruction of actuator fault of gates. The proposed observers, which makes use of a sliding mode controller algorithm, permits the direct reconstruction of the fault and it's been shown how the injection of the reconstructed fault determine compensation and allows to maintain the desired flow rate. Experimental results made on simulation confirm the effectiveness of the proposed schemes.

Among possible lines for future investigations it appears interesting to embed the suggested approaches into novel, consensus based, fault tolerant control systems for nonlinear uncertain processes.

More significant and practically relevant industrial applications of the developed approaches are devised in the framework of the previously mentioned PRODI project.