UNIVERSITÀ CAMPUS BIO-MEDICO DI ROMA
DOCTORAL PROGRAMME IN BIOMEDICAL ENGINEERING
CYCLE XXV

NEW METHODOLOGIES FOR THE FILTERING AND STATE ESTIMATION PROBLEMS IN BIO-MEDICAL APPLICATIONS

Doctoral Dissertation of:
Valerio Cusimano

Supervisor:
Prof. Alfredo Germani

2012

Abstract

BIOMEDICAL engineering (BME) is the application of engineering principles and design concepts to medicine and biology. This field seeks to close the gap between engineering and medicine: It combines the design and problem solving skills of engineering with medical and biological sciences to advance healthcare treatment, including diagnosis, monitoring, treatment and therapy. Biomedical engineering has only recently emerged as its own discipline, compared to many other engineering fields. Such an evolution is common as a new field transitions from being an interdisciplinary specialization among already-established fields, to being considered a field in itself. Much of the work in biomedical engineering consists of research and development, spanning a broad array of subfields (tissue and cellular engineering, bioinformatics, robotics, etc.).

In this field, the control theory is becoming increasingly important in medical applications because the use of specific techniques (state observer, parameter estimation, Kalman filter etc.) allows to obtain information not available with traditional engineering methods. For this reason, in this thesis have been developed new techniques for filtering and state estimation for the solution of biomedical problems.

In particular in Chapter 1, it has been described a method of discretization and a state observer capable of tackling the problem of tumor growth. A state observer is a mathematical tool that allows tracking the behavior of the unknown state variables of a system starting from the value of a measurable subset. A state observer is based on the knowledge of the dynamics of the system and its evolution from an initial estimate of the state variables converges asymptotically to the true value.

In the field of molecular biology, the study of gene expression, meaning by this term the series of events that after the activation of transcription of a gene, leading to the production of the corresponding protein, is of considerable interest. Gene expression is regulated internally to the cell mainly by transcription factors, specific proteins that determine which genes are being expressed and which are not.

In this thesis in Chapter 2, has been developed an original mathematical model for the

gene expression between co-regulated genes. The dynamics of the adjustment made by the transcription factor was modeled as a white noise and the model parameters were estimated using the Kalman filter and identified through the maximum likelihood estimation.

Finally, it was studied two classical problems of robotic field (planar tracking and perspective vision) and their solutions have been addressed through the use of the new concept of virtual measurement map, in the Chapter 3.

Contents

Contents

CHAPTER 1

Tumor Growth

1.1 Introduction

Cancer represents one of the leading causes of death in the developed countries. In US, in 2009, almost 1.5 million new cancer cases have been diagnosed and, in the same year, more than half a million deaths due to cancer have been registered [42]. This has huge effects and implications in the health state, that calls for innovative treatment solutions: within this framework, the availability of algorithms to improve treatment infusion processes design might concretely contribute to reduce morbidity and economic burden, making the novel care-strategy more efficient and with smaller side-effects.

In the last decades, many mathematical models have been proposed, to describe the untreated [2] and treated tumors growth [22] at different level of complexity, specifically from the macroscopic to the microscopic molecular scale of analysis. These models have been aimed at improving the knowledge of the tumor growth mechanism and/or at acquiring useful hints of chemotherapy and radiotherapy outcomes. On the other hand, they have been only marginally used to implement a therapy strategy and, more in general, in clinical applications, mainly due to the intrinsic complexity of phenomena related with tumor growth and to the difficulties in identifying the model parameters from a limited set of available measurements. This problem is dramatically true for *in vivo* applications, since in this case the available information is provided just by means of biomedical image techniques (TAC, RNM, US) [84].

Generally speaking, detailed mathematical models for tumor growth must take into account several factors and, specifically, the nutrient concentration profiles into the tumor

mass. To accomplish this task, even in the unrealistic assumption of perfectly spherical symmetry, partial differential equations must be solved, to include the space and time dependence for all variables of interest.

In the early stage of the invasion, called the avascular phase, a solid tumor exhibits a spheroidal form: in this phase, nutrients enter in the tumor mass only by means of diffusion. As the tumor grows, its inner cells are progressively less supplied with these nutrients, eventually switching to a quiescent state, further followed by death, if the nutrients falls below a critical concentration value C_C. The formation of this necrotic core, including dead and quiescent cells, is a critical stage in tumor invasion, since quiescent cells produce signaling molecules, called Tumor Angiogenic Factors (TAF's), responsible for the tumor angiogenesis. Thus, the forecast of the necrotic core onset is a remarkable tool to predict the tumor invasion evolution.

Interesting results can be achieved by adopting models based on ordinary differential equations (ODEs), aimed at describing the evolution over time of the number of proliferant, quiescent and dead cells; even neglecting the nutrient concentration variation in the tumor mass, they are able to reproduce with reasonable accuracy the typical tumor growth. These models generally rely on an empirical definition of the dynamics of the tumor cell growth, based on parameters that characterize its kinetic features. Among them a relevant place is occupied by the Gompertz model, a model of the growth of a population in the case of limited resources, that has been seen chosen as a benchmark for the description of the tumor growth [14, 65].

1.2 Mathematical model

In this thesis, the Gompertz model was taken into account to describe the time evolution of the number $N(t)$ of tumor cells.

Gompertz model has been proposed in the early XIX century, for epidemiological purposes [34]; since then, it has been applied in many different fields, to explain the controlled growth of entities under limiting supply conditions.

In the biological area, this model was applied efficiently to represent the cell population growth in the case of limited nutrients and also for the tumor cell growth in the avascular phase [70], recently, it has been revised in terms of an energetic interpretation of its kinetic parameters [14].

The practical applicability of it is limited by two factors, namely:

1. Parameter identification: there is no straightforward method to infer the value of the parameters contained in the model from biological evidence.

2. Discretization: as many other models based on nonlinear ODEs, the Gompertz model cannot be immediately applied when only discrete-time measurements are available.

As for the second point, it must be noticed that the time interval between available measurements is usually much larger than the threshold allowed for simple linearization techniques, thus more sophisticated methodologies are required in order to analyze experimental data.

The approach, proposed in this work to overcome these issues, is a combination of a novel discretization schema and a nonlinear discrete-time state observer. In control theory, a state observer is a mathematical tool that allows to track the behavior of the unknown state variables of a system starting from the value of a measurable subset. A state observer is based on the knowledge of the dynamics of the system and its evolution from an initial estimate of the state variables converges asymptotically to the true value. Even though linear systems observers, such as the Luenberger observer, are standard tools in control applications, observers for nonlinear cases [18] are more difficult to design.

It was extended the model to take into account the action of a chemotherapy drug, so the corresponding dynamical system describes the time evolution of the number of neoplastic cells $N(t)$ and the plasma drug concentration $A(t)$.

The discrete-time version of the Gompertz's model of the tumor growth, together with the state observer, represents a valuable tool in practical applications, for example:

1. it is possible to estimate the unknown parameters of the model for specific tumor cells, starting from *in vivo* or *in vitro* measurements taken at large time intervals;
2. it is possible to quantify the impact of a chemotherapy agent from experimental measurements through the estimation of the parameters of the combined tumor-drug dynamical system;
3. the knowledge of the model parameters can be used to devise more efficient strategies of chemotherapy infusion;
4. an efficient chemotherapy infusion can be devised even when the initial knowledge of the tumor and drug dynamics are uncertain.

It was validated this approach *in silico* on the first two practical applications sketched above, and, specifically, it was tested the ability of the discrete-time model to approximate the dynamics of the continuous-time original model, as well as to identify the unknown parameters of the model starting from measurements of the tumor size and chemotherapic concentration after the application of different drug infusion sessions. These results comply satisfactorily with *in vivo* literature data, collected in similar conditions.

1.2.1 Gompertz model

The Gompertz model is defined as solution of the equation:

$$\dot{N}(t) = \gamma \cdot N(t) \cdot \log\left(\frac{N_\infty}{N(t)}\right), \tag{1.1}$$

where $N(t)$ is the size of the cell population at the time t, γ is the kinetic constant of the growth rate, N_∞ is the theoretical saturation value that depends on the kind of cells and the environmental properties.

In the case of solid tumors, $N(t)$ can also be expressed in terms of the cell growth kinetics r_N that is supposed to be a linear function of the number of proliferant cells $P(t)$

(the ones that are able to replicate) [14].

$$\dot{N}(t) = r_N(t) = \text{const} \cdot P(t) \tag{1.2}$$

where the *const* does not depend on N. Comparing (1.1) and (1.2) it was derived *const*$= \gamma$, and

$$P(t) = N(t) \cdot \log\left(\frac{N_\infty}{N(t)}\right) \tag{1.3}$$

The fraction of proliferant cells is therefore

$$f_P = \frac{P(t)}{N(t)} = \log\left(\frac{N_\infty}{N(t)}\right). \tag{1.4}$$

Since it must be $0 \leq f_P \leq 1$, the two following conditions on $N(t)$ are obtained:

$$\begin{aligned} f_P \geq 0 \quad &\rightarrow \quad N \leq N_\infty \\ f_P \leq 1 \quad &\rightarrow \quad N \geq N_{min} = \frac{N_\infty}{e}. \end{aligned}$$

The first condition is trivial: the overall cell number must not exceed the asymptotic saturation value. The second one is a lower bound to the cell population number, since for $N < N_{min}$ the cells are very few and they behave independently. In this case, nutrients supply is far larger than the amount strictly required to cell maintenance and replication, so all the N cells are vital and the kinetic expression can be approximated by an exponential growth.

Thus, the unconstrained growth law, *i.e.* without any chemotherapy action, is described in terms of the function

$$\begin{cases} \dot{N}(t) = \gamma \cdot N(t) & N(t) < N_{min} \\ \dot{N}(t) = \gamma \cdot N(t) \cdot \log\left(\frac{N_\infty}{N(t)}\right) & N(t) \geq N_{min}. \end{cases} \tag{1.5}$$

In this work only the case with $N(t) \geq N_{min}$ will be considered, due to its significance in health issues (only larger cell spheroids are detectable by the common diagnostic techniques).

The chemotherapy effect of one drug acting on the proliferant cells P was introduced :

$$\dot{A}(t) = u(t) - A(t)\left(\lambda_A + \beta_2 P(t)\right) = u(t) - A(t)\left(\lambda_A + \beta_2 N(t) \log\left(\frac{N_\infty}{N(t)}\right)\right) \tag{1.6}$$

where $A(t)$ is the drug concentration in blood, λ_A is the kinetic constant referred to its metabolic decay rate, β_2 is the kinetic constant, referred to the drug kill rate; $u(t)$ is the input function for the drug injection into the body, in this case represented by a periodic finite impulse of equation

$$u(t) = \begin{cases} a_c & \text{if } \mathrm{mod}(t, T) > \tau \\ 0 & \text{otherwise} \end{cases} \tag{1.7}$$

of amplitude a_c and period T ; (1.7) models a chemotherapic infusion that takes place for a fraction τ of the chemoterapic cycle, and mimics a typical infusion protocol strategy.

Assuming that the drug concentration can kill of a fraction of tumor cells proportional to $A(t)$ through some constant β_1, the resulting dynamical system becomes

$$\begin{cases} \dot{N}(t) = N(t) \log\left(\frac{N_\infty}{N(t)}\right)(\gamma - \beta_1 A(t)) \\ \dot{A}(t) = u(t) - A(t)\left(\lambda_A + \beta_2 N(t) \log\left(\frac{N_\infty}{N(t)}\right)\right) \end{cases} \tag{1.8}$$

Since it concerned with the estimation of the possibly unknown parameters of the model, a common approach in the parameter identification field was used that consists of considering any unknown parameter θ_i as a *state variable* with equation $\dot{\theta}_i = 0$, that reflects the fact that they are actually constant over time; in this way, any state estimation technique may be used to estimate also the parameters of the system. Here it is presented how to evaluate the parameters λ_A e β_1, that is, the metabolic decay rate of the chemotherapic drug and its effect on the tumor cells. The complete dynamical model for the tumor growth is therefore

$$\begin{cases} \dot{N}(t) &= N(t) \log\left(\frac{N_\infty}{N(t)}\right)(\gamma - \beta_1 A(t)) \\ \dot{A}(t) &= u(t) - A(t)\left(\lambda_A + \beta_2 N(t) \log\left(\frac{N_\infty}{N(t)}\right)\right) \\ \dot{\beta}_1 &= 0 \\ \dot{\lambda}_A &= 0. \end{cases} \tag{1.9}$$

1.3 Mathematical tools

1.3.1 Discretization

The oncological therapy and disease follow-up rely typical on periodic assessments, resulting in a time evolution of tumor size. These data are, then, discrete in nature, all the more often not distributed in an even fashion over time.

Furthermore, the time spanning between two measurements is not negligible, when compared to the time scale of interest. For all these reasons, within this field of application, it is mandatory to build-up mathematical tools able to manage discrete-time data. However, the usual discretization methods for continuous-time models are not suitable for this application, because they only work for much smaller discretization intervals.

Considering eq. (1.9), it is clear that the variables that can be measured, such as the tumor size $N(t)$, are available only at time intervals spanning over days – in the case of real tumors – or at best over hours – in the case of data collected from *in vitro* experiments. In any case, these time intervals are quite long with respect to the the dynamics of the system, and this make usual numerical integration methods based on linearization techniques, such as Eulero or Runge-Kutta, completely inapplicable.

These considerations have encouraged the development of the novel discretization approach that it is presented in this thesis. This approach can be applied, as well, to a large

class of continuous-time nonlinear systems. The basic idea is known as *Carleman linearization* [54], and it consists of building an infinite system of linear equations that is equivalent to a given finite system of nonlinear equations. The resulting linear system, truncated to attain the desired approximation, can be discretized exactly. In the following the procedure in the case of the system (1.9) is described.

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth nonlinear map. From now on, the symbol $\otimes$ denotes the Kronecker matrix product, the notation $A^{[i]}$ is used for the Kronecker power of matrix A (that is $A \otimes A \otimes \cdots \otimes A$, repeated i times). The standard Jacobian of f can be formally written as $\nabla_x \otimes f$, where ∇_x denotes the operator $[\partial/\partial x_1, \ldots, \partial/\partial x_n]$; higher-order Jacobians of f can thus be expressed as $\nabla_x^{[i]} \otimes f$.

The system (1.9) can be written as

$$\dot{x}(t) = f(x(t), u(t)) \tag{1.10}$$

where $x(t) = [N(t)\ A(t)\ \beta_1\ \lambda_A]^T \in \mathbb{R}^4$ is the state variables vector. In the forced case, the input $u(t)$ is assumed to be constant in the time interval $t \in [k\Delta, (k+1)\Delta]$, thus in this time interval the system is autonomous. The aim of the proposed discretization approach is to find an expression for $x((k+1)\Delta)$, given $x(k\Delta)$ and $u(k\Delta)$. The system (1.10) can be re-written, for the time interval of interest, as

$$\dot{x}(t) = f(x(t), u(t)) = \bar{f}(x(t), u(k\Delta)), \qquad t \in [k\Delta, (k+1)\Delta] \tag{1.11}$$

Let $\varphi(t) = x(t) - x(k\Delta)$ and consider the ν-th order Taylor expansion of $f(x, u)$ with respect to x at the initial point $x(k\Delta)$:

$$\dot{x}(t) = \bar{f}(x(t), u(k\Delta)) = \dot{\varphi}(t) = \sum_{i=0}^{\nu} A_i^1 \varphi^{[i]} + r_{\nu+1}(\varphi(t)) \tag{1.12}$$

with

$$A_i^1 = \frac{\nabla^{[i]} \otimes \bar{f}|_{x(k\Delta)}}{i!} \in \mathbb{R}^{n \times n^i} \tag{1.13}$$

and

$$\| r_{\nu+1}(\varphi) \| \leqslant \frac{\Gamma_{\nu+1}}{(\nu+1)!} \| \varphi(t)^{\nu+1} \| \tag{1.14}$$

where $\Gamma_{\nu+1}$ denotes an upper bound for the $\otimes^{[\nu+1]} \bar{f}|_{x(k\Delta)}$. The function $\varphi(t)$ for $k\Delta \leq t \leq (k+1)\Delta$ represents the evolution of the system starting at $x(k\Delta)$, in the discretization interval $[k\Delta,\ (k+1)\Delta]$. The case $\nu = 1$ correspond to the classical linearization procedure.

For the model of interest, the value $\nu = 2$ provides a satisfactory approximation of the state evolution; as a matter of fact, neglecting the remainder in the Taylor's series and using the same symbols for both the original and the approximated state variables, eq. (1.12) can be written

$$\dot{x}(t) \ = \ \sum_{i=0}^{2} \frac{\nabla^{[i]} \otimes \bar{f}|_{x(k\Delta)}}{i!} \varphi(t)^{[i]} = \sum_{i=0}^{2} A_i^1 \varphi^{[i]} = A_0^1 + A_1^1 \varphi + A_2^1 \varphi^{[2]} \tag{1.15}$$

Let

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} \varphi \\ \varphi^{[2]} \end{bmatrix} \tag{1.16}$$

be a new vector of state variables, made up by the original variables φ and by their Kronecker square $\varphi^{[2]}$; then, $\phi_1 \in \mathbb{R}^4$, $\phi_2 \in \mathbb{R}^{16}$, and $\phi \in \mathbb{R}^{20}$.

The dynamics of the system with respect to ϕ can be written therefore:

$$\dot{\phi} = \begin{bmatrix} \dot{\phi}_1 \\ \dot{\phi}_2 \end{bmatrix} = \begin{bmatrix} A_0^1 + A_1^1\varphi + A_2^1\varphi^{[2]} \\ A_0^2\varphi + A_1^2\varphi^{[2]} + A_1^3\varphi^{[3]} \end{bmatrix} = \begin{bmatrix} A_0^1 + A_1^1\phi_1 + A_2^1\phi_2 \\ A_0^2\phi_1 + A_1^2\phi_2 + A_1^3\phi_1 \otimes \phi_2 \end{bmatrix} \tag{1.17}$$

where $A_i^2 = A_i^1 \otimes I_4 + I_4 \otimes A_i^1$, $A_i^2 \in \mathbb{R}^{4^2 \times 4^{i+1}}$, and I_n is the identity matrix of size n. A solution for this linear system can be easily obtained by neglecting the term $\phi_1 \otimes \phi_2$. In this case it is

$$\dot{\phi} = L + M\phi \tag{1.18}$$

with

$$L = \begin{bmatrix} A_0^1 \\ 0 \end{bmatrix}, \quad M = \begin{bmatrix} A_1^1 & A_2^1 \\ A_0^2 & A_1^2 \end{bmatrix}$$

Specifically for the model of interest, $L \in \mathbb{R}^{20}$ is defined as:

$$L = \begin{bmatrix} f(x(k)) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} N \log\left(\frac{N_\infty}{N}\right)(\gamma - \beta_1 A) \\ -A\left(\lambda_A + \beta_2 N \log\left(\frac{N_\infty}{N}\right)\right) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and

$$A_1^1 = \begin{bmatrix} (\gamma - \beta_1 A)(\log\left(\frac{N_\infty}{N}\right) - 1) & -\beta_1 N \log\left(\frac{N_\infty}{N}\right) & -AN \log\left(\frac{N_\infty}{N}\right) & 0 \\ (-\beta_2 \log\left(\frac{N_\infty}{N}\right) + \beta_2)A & -\lambda_A - \beta_2 N \log\left(\frac{N_\infty}{N}\right) & 0 & -A \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} (\gamma - x_3\, x_2)(\log\left(\frac{N_\infty}{x_1}\right) - 1) & -x_3\, x_1 \log\left(\frac{N_\infty}{x_1}\right) & -x_1\, x_2 N \log\left(\frac{N_\infty}{x_1}\right) & 0 \\ (-\beta_2 \log\left(\frac{N_\infty}{x_1}\right) + \beta_2)A & -x_4 - \beta_2\, x_1 \log\left(\frac{N_\infty}{x_1}\right) & 0 & -x_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The system is linear and can be discretized exactly, for $t \in [k\Delta, (k+1)\Delta]$, in the form:

$$\begin{aligned} x_a(k+1) &= x_a(k) + [I_n \mid 0_{m-1}] \int_0^\Delta \exp^{M(k)\tau}\, d\tau L(k) \\ &= x_a(k) + [I_n \mid 0_{m-1}](M(k)^{-1}[\exp^{M(k)\Delta} - I]L(k)) \end{aligned} \tag{1.19}$$

where $x_a(k)$ stands for the approximate value of the $x(t)$ calculated at $t = k\Delta$. If the matrix exponential is evaluated by means of a h-th order truncated series, the discrete system becomes

$$x_d(k+1) = F(x(k), u(k)) = x_d(k) + \sum_{i=0}^{h} \left(\frac{(M(k))^i \Delta^i}{i!} L(k) \right) \tag{1.20}$$

1.3.2 Model parameters identification via a state nonlinear observer

Drug pharmacokinetics and pharmacodynamics are key issues in oncological pharmacology. The knowledge of the molecular mechanism for the drug action, yet substantial, is just the very first step in the drug efficiency evaluation in terms of therapy outcome over time. It is of a capital significance, indeed, to determine quantitatively the time evolution of drug distribution in body and the kinetics of its pharmacological action.

A nonlinear observer can be an important tool to follow up the time evolution of complex systems, such as the biological ones. Moreover, observers can also be used to estimate the parameters of the model; this application differs considerably from the common regression procedure, since it allows a dynamical adaptation of the parameters according to the system measured output, possibly accounting also for the system perturbation or the onset of novel phenomena in the system time evolution process.

Specifically, in this work a procedure to evaluate the kinetic drug parameters, β_1 and λ_A, starting from the discrete-time data of $N(t)$ and $A(t)$, that is tumor size and drug plasma concentration, was implemented. This task may be accomplished representing the parameters as state variables, and by means of an extension of the nonlinear discrete-time observer [19] to the case of multiple measured variables. The observer provides a dynamical estimate of the parameters that converges to the true value and that can be used to define an optimal chemotherapy treatment strategy.

The discrete-time system described has the form:

$$\begin{cases} x(k+1) = F(x(k), u(k)) \\ y(k) = h(x(k)) \end{cases} \tag{1.21}$$

with state $x(k) \in \mathbb{R}^4$ and output $y(k) = [y_1(k)\ y_2(k)]^T = [x_1(k)\ x_2(k)]^T \in \mathbb{R}^2$, this latter represented by the discrete-time measurements of $N(t)$ and $A(t)$.

It was briefly outlined the theory underlying the design of the discrete-time observer for the system of interest. More details for single output systems can be found in [19]. Let us define the nonlinear transformation $\Phi_{u(k-1)} : R^4 \rightarrow R^4$

$$z(k) = \begin{pmatrix} y(k) \\ y(k-1) \end{pmatrix} \tag{1.22}$$

$$= \begin{pmatrix} h \circ F(x(k-1), u(k-1)) \\ h(x(k-1)) \end{pmatrix} = \tag{1.23}$$

$$= \Phi_{u(k-1)}(x(k-1)), \tag{1.24}$$

where $\circ$ denotes the composition operator. If $\Phi_{u(k-1)}$ admits an inverse, the state vector is:

$$x(k-1) = \Phi^{-1}_{u(k-1)}(z(k)) \tag{1.25}$$

The inverse function $\Phi^{-1}_{u(k-1)}(z(k))$ generally does not exist in closed form, it is therefore required to introduce an approximated solution.

To this aim, the Newton-Raphson algorithm can be used; starting from the $x^*(k-2)$, the estimate of $x(k-2)$, we have as predicted value of the state at the time instant $k-1$:

$$\bar{x}(k-1) = F(x^*(k-2), u(k-2))$$

Let us consider the first order Taylor series of (1.22), with starting point $\bar{x}(k-1)$:

$$z(k) = \phi(\bar{x}(k-1), u(k-1)) + \left.\frac{d\phi(x(k-1), u(k-1))}{dx}\right|_{\bar{x}(k-1)} (x(k-1) - \bar{x}(k-1)) + r \tag{1.26}$$

where r stands for the remainder of order two.

The algorithm is readily obtained by removing the remainder and by solving eq. (1.26), whose solution will be taken as $x^*(k-1)$, the updated state estimate at time $k-1$,

$$x^*(k-1) = \bar{x}(k-1) + \left(Q_{(x,u)}\right)^{-1} [z(k) - \phi(\bar{x}(k-1), u(k-1))] \tag{1.27}$$

where $Q_{(x,u)} = \left.\frac{d\phi(x(k-1),u(k-1))}{dx}\right|_{\bar{x}(k-1)}$.

Finally, the state observer $\hat{x}(k)$ is obtained as

$$\hat{x}(k) = F(x^*(k-1), u(k-1)) \tag{1.28}$$

In practice, it may be useful to introduce in (1.27) a multiplicative gain constant $c \in (0, 1]$ with the aim to speed up the convergence of the algorithm. Therefore, the state equation has the following form:

$$x^*(k-1) = \bar{x}(k-1) + c\ \left(Q_{(x,u)}\right)^{-1} [z(k) - \phi(\bar{x}(k-1), u(k-1))] \tag{1.29}$$

1.4 Results and discussion

1.4.1 Performance of the time-discrete model

After the introduction of the approach, an analysis of the performance of the discretization scheme, that is, its ability to reproduce the dynamics of the continuous-time model in presence of large discretization intervals, is required. Given the system (1.9), it was considered two possible scenarios, corresponding to (i) free growth, and (ii) periodic infusion of a chemotherapy agent.

In the case (i), the time interval $[0, 50]$ is chosen in order to allow the population to reach the steady value $N_\infty = 10^5$ at $\gamma = 0.2\,\mathrm{d}^{-1}$ (d=days).

The population starts at $N(0) = 1$; notice that the initial exponential growth of the population size (from about 0 to 10^4 in the interval $[0, 15]$) constitutes a challenging dynamics for any discretization scheme. In Figure 1.1 the dynamics of the continuous and discretized models are compared for a discretization interval of $\Delta = 0.75\,\mathrm{d}$. With reference to eq. (1.20), two cases are plotted: $\nu = 2,\ h = 2$, that corresponds to a second order Carleman linearization with a second order polynomial approximation of the matrix exponential; $\nu = 2,\ h = 5$, that introduces a fifth order polynomial approximation of the matrix exponential. It is noteworthy that in the second case, the evolution of the continuous model is well approximated by the discrete model.

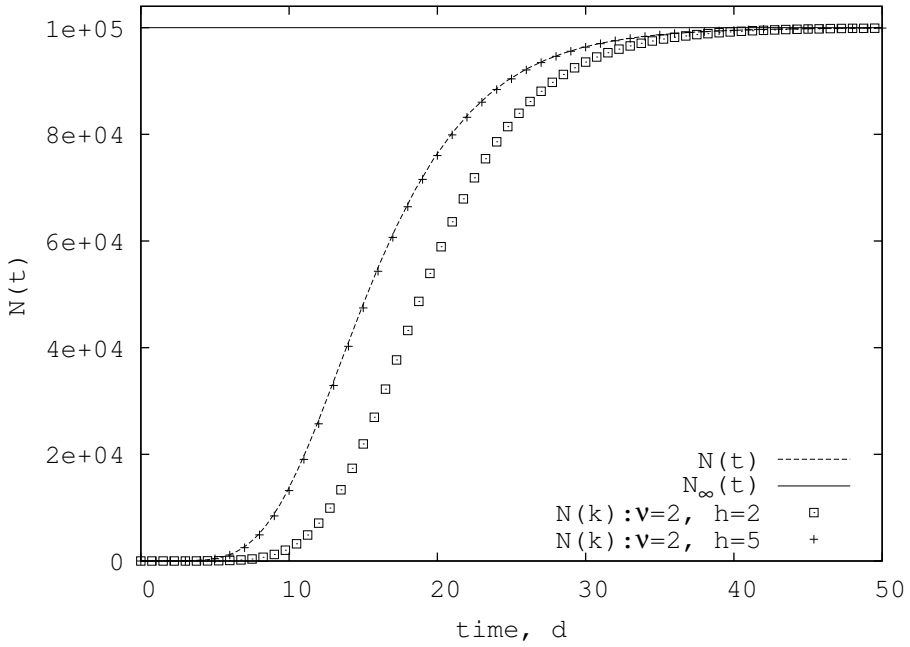

Figure 1.1: *Gompertz continuous and discrete models, the unforced case.*

In the scenario (ii) (Figure 1.2) the chemoterapic infusion function is modeled as in (1.7) with a finite impulse with amplitude $a_c = 0.2$, infusion time $\tau = 3\,\mathrm{d}$ and period $T = 10\,\mathrm{d}$. The plot reports the time interval corresponding to two infusion cycles, with $\Delta = 0.5\,\mathrm{d}$, $\beta_1 = 0.2\,\mathrm{d}^{-1}$, $\lambda_A = 0.1\,\mathrm{d}^{-1}$, $\beta_2 = 0.1\,\mathrm{d}^{-1}$. In this case, the tumor size at about $t = 7\,\mathrm{d}$ is approximately half the N_∞ before starting to decrease as an effect of the chemoterapic infusion. The function $A(t)$ is not smooth, due to the noncontinuous nature of $u(t)$; the discretized model uses $\nu = 2,\ h = 5$, as above. It can be remarked that, even in the case of a non-continuous input step function, the time-discrete model approximates well the reference Gompertz model.

The curves represented in Figure 1.2 adhere to the real evolution over time of treated neoplastic lesions. Moreover, it emerges a very good compliance between continuous and discrete-time systems. As a comparison, it may be noticed that a typical solver based on the Runge-Kutta method for nonlinear ODEs would have used an integration step of about

0.02 d. The proposed discretization scheme, instead, allows a discretization step about 25 times larger.

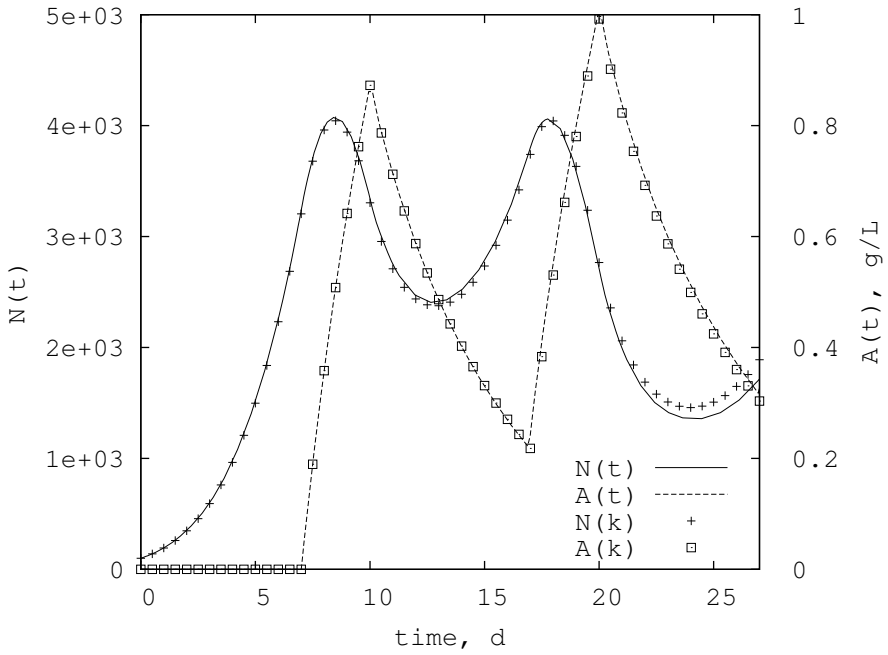

Figure 1.2: *Gompertz continuous and discrete models, the forced case.*

1.4.2 Parameter estimation

Now the use of the nonlinear observer to the problem of estimating the parameters of the system in presence of measurements with large discretization intervals was evaluated ; as in the previous section, the unforced and forced cases were considered.

In the unforced case (no chemotherapy) the observer is used to estimate the final tumor size N_∞, whereas in the forced scenario (infusion of the chemotherapic agent), the observer is used to estimate the parameters β_1 and λ_A, related to the drug dynamics and effect.

This choice may correspond to an *incremental estimation* strategy: the parameters related to the tumor are estimated, by means of the observer, starting from laboratory measurements of its free growth. The information acquired in this first step can be used to design a succeeding experiment including the drug infusion, with the aim at estimating the parameters related to the drug dynamics. At this point, all the model parameters are known, and the model can be used to predict the effect of the drug therapy or to optimize the drug infusion strategy.

In the free growth scenario, the parameter values are the same used before. As stated above, the aim is estimating N_∞, thus the observer is built for the following dynamical

systems

$$\begin{cases} \dot{N}(t) = N(t) \log\left(\frac{N_\infty}{N(t)}\right) \\ \dot{N}_\infty = 0, \end{cases} \tag{1.30}$$

and the measured output is $N(t)$.

Applying the model 1.30 on the evolution time interval of $[0, 50]$, with $N_\infty = 10^5$, $N(0) = 1$ and $\gamma = 0.2\,\mathrm{d}^{-1}$, whereas the initial estimate of the state variables is $N^*_\infty = 10^4$, $N^* = 1$, Figure 1.3 shows the dynamics of the discretized models and nonlinear observer, for a discretization interval of $\Delta = 0.5\,\mathrm{d}$; for the sake of an improved visualization, the values are plotted as continuous lines, although they are discrete.

It can be noticed that the observer converges asymptotically to the true value after less than one third of the time required for cancer cells (N) to reach the saturation condition (N_∞); this has the important practical implication that it is possible to predict the final maximum size of the tumor starting from a small set of initial measurements of its growth.

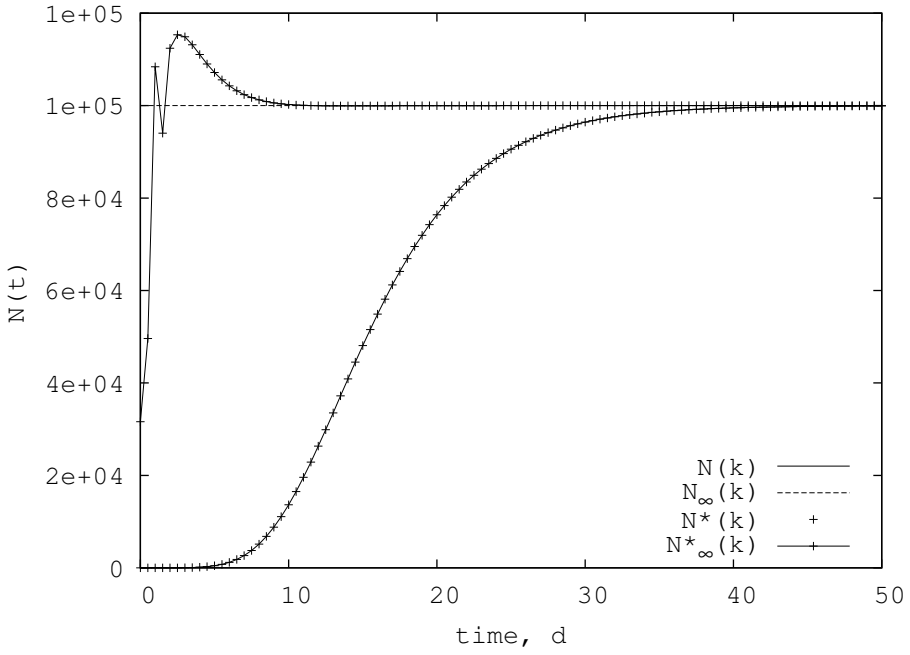

Figure 1.3: *Gompertz discrete models and nonlinear observer, the unforced case.*

In the forced case, it was built the discretized version and the observer for the system (1.9), and the measured system output is the pair $[N(k)\, A(k)]$. In the simulations, the initial conditions are the same as in the previous case. The true values of the parameters are $\beta_1 = 0.2\,\mathrm{d}^{-1}$, $\lambda_A = 0.1\,\mathrm{d}^{-1}$, and the initial estimate used by the observer is $\beta_1^* = 0.7\,\mathrm{d}^{-1}$, $\lambda_A^* = 1\,\mathrm{d}^{-1}$. Figure 1.4 shows that also in this case, the observer converges quickly to the real value, after a short transient phase. A good estimate for λ_A is obtained after 2 days (only 3 measurements are necessary), and a good estimate for β_1 after about 8 days, corresponding to 15 measurements.

Thus, the discrete and continuous-time observers can be used successfully to evaluate model parameters, with a fast and reliable convergence, even if the initial value is quite far from the final one.

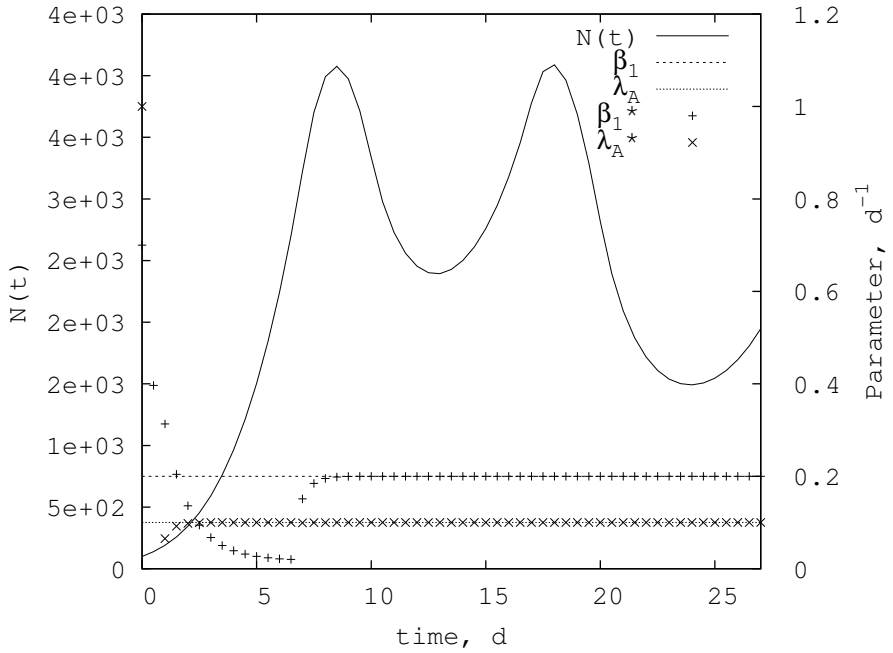

Figure 1.4: *Gompertz discrete models and nonlinear observer, the forced case: kinetic parameters (β_1, λ_A) estimation is also showed*

1.5 Conclusions

In this study, a classical dynamic model for solid tumor growth (Gompertz model) was presented; it was introduced also the effect of a chemotherapy agent, in terms of tumor size shrinking kinetics. The model has been used to develop a methodology relying on measured properties for the system (such as the tumor size), to infer information about other hidden variables. It was showed some preliminary applications of this methodology, in terms of capability to characterize tumor size dynamics, estimation of unknown parameters and design of the chemotherapy infusion strategy.

The main contribution of this work is to show the feasibility, in the context of anticancer treatment, of an approach based on a discrete-time systems and the use of nonlinear observers; this approach is amenable to be applied to a variety of biological systems and models.

On the side of the nonlinear observer, an useful extension has been to adopt a discrete measurement approach, rather than a continuous one, in order to reflect the discrete nature of the measurements of tumor growth, both *in vitro* and *in vivo*: this task has been accomplished by introducing a proper discretization procedure.

The simulation results show that the untreated and treated tumor growth predicted by the proposed discrete-time model complies satisfactorily with time evolution of real biological systems: the saw tooth form of $N(t)$ replicates the typical outcome of chemotherapy periodic applications to neoplastic lesions.

Further on, it is shown that these methods can be successfully extended to the parameters estimation; this application can be thought as an alternative to the classical nonlinear regression algorithms, that often rely on a small interval of convergence. Moreover, the parameters estimation via the nonlinear observer is fast and precise, even if the initial values are quite far from the true ones.

Future extensions of this work on the biological side include the development of more sophisticated models of the tumor growth, whereas on the methodological side the proposed approach can be extended to the stochastic case, through the use of a nonlinear filter in the place of the deterministic observer.

CHAPTER 2

mRNA Half-Lives

2.1 Introduction

Appropriate and timely changes in gene expression are essential for cell life. The transcriptome in a cell is finely regulated by a large number of molecular mechanisms able to adjust the balance between mRNA production and degradation. Every aspect of transcript life is subject to elaborate control but, traditionally, the focus of the research has been on transcriptional regulation [31]. However, whereas mRNA abundance results from the dynamic interplay between transcription and degradation, the speed by which cells can adjust their mRNA levels is critically dependent on the rate of mRNA turnover [64]. As a result, small changes in mRNA stability may dramatically drive rapid variations of transcript abundance. Efforts to understand the underlying principles of mRNA decay and transcription co-ordination are very important since the balance between transcription and decay influences most, if not all, the cell responses to endogenous and exogenous signals [53].

The current widespread interest in this work has been fostered by the finding of specific regulatory mechanisms of mRNA stability such as, for example, RNA binding proteins [32] and small RNAs [39]. Regulation of transcript stability cannot be considered a simple "disposal system" but a sophisticated tool for the proper orchestration of the global cell response to internal and external stimuli. Remarkably, a key role of mRNA stability has been reported in cancer, inflammatory diseases and Alzheimer's [16]. In recent years there has been a surge in empirical studies that measured, on a genome-wide scale in a variety of environmental conditions, messenger half-lives of many organisms, including plants, mammals and fungi. The discovery of such new regulatory layer has clarified that,

in order to obtain a clear picture of the underlying regulatory machinery, it is necessary to complement the traditional time-course experiment measuring the cell transcriptional response under certain conditions far from steady state with decay rates data under the *same* condition [73].

Experimental procedures for the evaluation of mRNA decay rates are based on measuring gene expression upon inhibition of transcription or on pulse-chase RNA labeling protocols. Such protocols are very critical (see Figure 2.1 for a comparison among different studies), since, for instance, transcriptional shut-off blocks growth and has a profound effect on cellular physiology, as well as on mRNA metabolism.

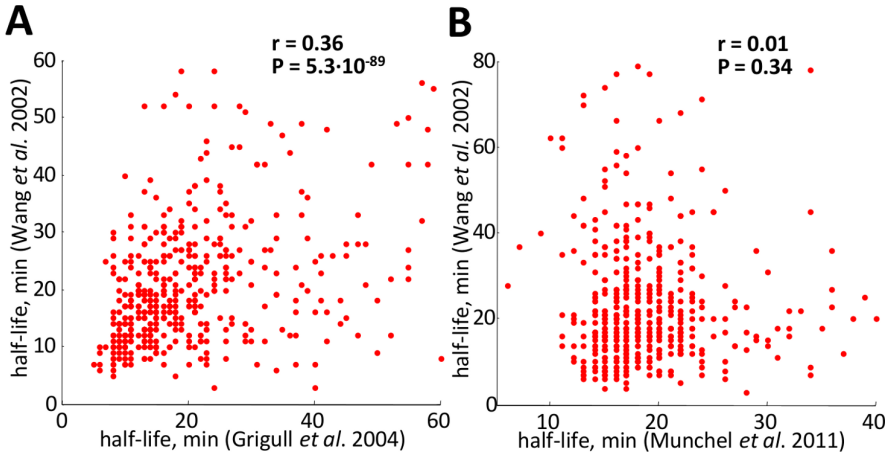

Figure 2.1: ***Basic comparison statistics among the yeast S. cerevisiae mRNA half-lives during asynchronous growth measured by three independent laboratories.*** *Three genome-wide studies considered are: Grigull et al, Wang et al and Munchel et al. (A) Scatterplot of Wang et al and Grigull et al datasets; (B) Scatterplot of Munchel et al and Wang et al datasets.*

In fact, Wang *et al.* [83] and Grigull *et al.* [36] datasets show a low value of the Pearson correlation ($r = 0.36$), and no correlation at all can be found ($r = 0.01$) between Munchel *et al.* [64] and Wang *et al.* datasets (see Figure 2.1A and Figure 2.1B respectively). Despite the same experimental conditions (asynchronous growth), the two half-life independent measurements obtained by Wang *et al.* and Munchel *et al.* are uncorrelated, probably due to differences in the shut off protocol (pulse chase for and thermal inactivation for , whereas Grigull *et al.* and Wang *et al.* appears significantly correlated, probably due to the same shut off protocol used (thermal inactivation).

It has been shown that genes having the same biological function are likely to share similar half-life values.

Measuring mRNA stability is a complex and expensive experiment and, given the condition-specific response of the degradation pathway, it would be desirable to take advantage of the large variety of expression experiments stored in public databases. To this end, it was developed a stochastic computational model of the expression kinetics to identify condition-specific mRNA stabilities which makes use only of experimental mRNA time profiles. It was also assumed that degradation rates are gene-specific but approximately

constant over the experiment time course. Predictions of algorithm, termed DRAGON (Decay RAtes from Gene expressiON), were validated on experimental mRNA abundance and turnover data, both collected during the Intraerythrocytic Developmental Cycle (IDC) of *Plasmodium falciparum*. The estimations were in line with the experimental measurements. Remarkably, the DRAGON estimated half-lives were consistent with the finding of a peculiar pattern of mean half-life values along the wave of sequentially induced genes in subsequent stages of *P. falciparum* development. It was also applied this methodology to public time-series datasets for which half-lives data, under the same experimental conditions, have not been experimentally measured. In particular the work was focused on budding yeast reproductive and metabolic cycle data. In fact, for the yeast *Saccharomyces cerevisiae*, only half-life data under asynchronous growth are publicly available [36,64,83] This study showed the presence of the same periodic pattern of mean half-life values in all datasets, thus suggesting that such behavior may be a general feature, not limited to the *Plasmodium falciparum* IDC.

2.2 mRNA Kinetics and Half-Life

Experimental evidence suggest that the majority of mRNAs are degraded with a first-order decay rate [71]. This allows to characterize mRNA disappearance time profiles by a first-order rate equation

$$\frac{dx(t)}{dt} = P(t) - kx(t) \tag{2.1}$$

where k is the decay rate (or half-life $t_{1/2}$, with $t_{1/2} = \ln(2)/k$), $x(t)$ is the mRNA concentration and $P(t)$ is the promoter activity (the rate of production of new mRNAs). It is worth noting that, the degradation rate k cannot be estimated from the concentration time profile $x(t)$ for a single gene, since the term $P(t)$ is not usually available in the typical time-course microarray experiment. The measurement of the promoter activity time profile would require additional experiments (such as those described in [30]) but, in this work, it is assumed that only mRNA abundance time-series data are available. At steady-state $x = x_{ss}$ and $P(t) = \bar{P}$ are constant so that $\frac{dx_{ss}(t)}{dt} = 0$ and, consequently,

$$x_{ss} = \bar{P}/k \tag{2.2}$$

From the above equation, it is clear that at steady-state an increase (decrease) in mRNA concentration can be produced either by an increase (decrease) of transcription or by a decreased (increased) value of the decay rate: the two regulatory strategies have therefore an equivalent outcome. As a result, from steady-states measurements, it is hopeless to reveal the relative contribution of transcription and degradation and, most importantly, their co-ordinated activity as well. By contrast, the whole kinetics of induction and relaxation, as measured by time-courses experiments, depends on the degradation and production rate in different ways: increasing (decreasing) the production rate results in a proportionally increased (decreased) mRNA abundance, whereas the rise time (*i.e.* the time required for the response to rise from 10% to 90% of its final value) is not affected. Increasing the

decay rate results in a faster rise time both in the induction and relaxation phases, whereas a decrease results in slower rise time [73]. This key point is illustrated in Figure 2.2A.

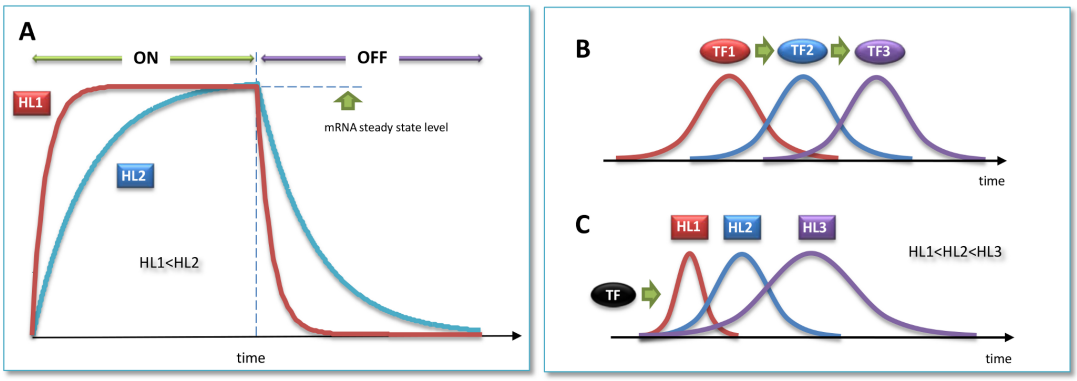

Figure 2.2: ***Kinetics of gene induction.*** *Panel A shows in silico experiments to illustrate some basic features of gene induction kinetics. The "ON" and "OFF" regions correspond to the turning "ON" or "OFF" of the promoter activity. (A) Induction kinetic of transcripts having the same steady-state concentration, but different half-lives and synthesis rates reaching the same steady state value. The time profile plotted in red corresponds to an unstable transcript and displays a fast induction and relaxation profile. By contrast, the blue one has an higher half-life value, resulting in a slower response. (B) Cascade of transcription factors resulting in waves of sequentially induced genes. The timing of expression peaks is modulated by transcriptional serial regulation. (C) Sequentially induced genes generated by a single transcription factor and a stability "gradient". The timing of expression peaks is modulated by post-transcriptional regulation. Early induced genes are those with a low half-life value, late induced ones are those with an high half-life value.*

Another important consequence of half-life specificity is the regulation of the timing of gene induction, as pointed out by Elkon *et al.* [25]. In fact, an expression wave, *i.e.* the sequential activation of genes, is usually interpreted as resulting from the corresponding activation of a multi-step transcription factors cascade (as illustrated in Figure 2.2B). Whereas such mechanism is certainly very important, there is also an alternative way to obtain an expression wave by means of a "stability gradient". As illustrated in Figure 2.2C, a single transcription factor may initiate transcription of a set of target genes and their peak of induction can be modulated by a stability "gradient", *i.e.* by specifically adjusted decay rates. More precisely, early induced genes would have short half-lives and late responding genes would have long half-lives. Clearly, both mechanism may well act in cells, thus generating a wide spectrum of responses.

Time-courses are a very common design for microarray analysis, which allows researchers to follow the dynamics of the cellular response to perturbations. Such data are available for a very large number of experimental conditions and organisms: only the Stanford Microarray Database includes to date 1545 time course data sets.

2.3 DRAGON

The goal of the DRAGON methodology is to derive a robust estimate of each mRNA species half-life starting from all available gene expression pairs. The rationale for the algorithm mainly draws on properties of pairs exhibiting a certain degree of common promoter activity (as in [26]). Besides, DRAGON infers common promoter activity using a statistical model that simulates both gene-specific and common effects.

The rate of change of mRNA concentration for a generic pair of genes, say gene j and gene h, is:

$$\begin{aligned} \frac{dx_j(t)}{dt} &= P_j(t) - k_j x_j(t) \\ \frac{dx_h(t)}{dt} &= P_h(t) - k_h x_h(t) \end{aligned} \tag{2.3}$$

where the symbols $x_j(t)$ and $x_h(t)$ represent the mRNA time profiles of the gene pair j and h, $P_j(t)$ and $P_h(t)$ are the promoters activity, and k_j and k_h are the degradation rate of mRNA of gene j and h, respectively. The terms $P_j(t)$, $P_h(t)$ are not known since it is considered the case in which only mRNA abundance is measured.

Promoter activities are modeled as the sum of two terms, the first one common to the pair and the other one specific for each gene:

$$\begin{aligned} P_j(t) &= b_j u(t) + w_j(t) \\ P_h(t) &= b_h u(t) + w_h(t) \end{aligned} \tag{2.4}$$

where $u(t)$ is the common part, scaled by constants b_j and b_h, whereas $w_j(t)$ and $w_h(t)$ are gene specific independent stochastic processes with zero mean, that is $w_j(t) = \sigma_j n_{x_j}(t)$, $w_h(t) = \sigma_h n_{x_h}(t)$. Equations (2.3)-(2.4) encompass the case of:

i) $P_j(t)$, $P_h(t)$ fully correlated (correlation $r = 1$) for which $b_j, b_h > 0$ and $\sigma_j = \sigma_h = 0$

ii) $P_j(t)$, $P_h(t)$ partially correlated (correlation $r < 1$) for which $b_j, b_h > 0$ and $\sigma_j, \sigma_h > 0$

iii) $P_j(t)$, $P_h(t)$ un-correlated (correlation $r = 0$) for which either $b_j = 0$ or $b_h = 0$.

Equations (2.3)-(2.4) can therefore be written for all available gene pairs; thus, for a set of N genes, there are N^2 pairs to analyze. For each gene pair DRAGON provides an estimate of the time profile of $u(t)$, of all the parameters, k_j, k_h, b_j, b_h, and the covariance matrix of the stochastic processes. For each gene therefore N estimates of the decay rate k are calculated, one for each pair containing that gene.

Notice that equations (2.3)-(2.4) yield a couple of linear stochastic differential equations. Since measurements of mRNA concentrations are available only at given time points, it is necessary to transform (2.3)-(2.4) in a couple of discrete stochastic equations. The exact discretization of (2.3)-(2.4) is possible since they are linear [47]. The *Kalman filter* [48] is used on the resulting discrete equations and a maximum likelihood algorithm is exploited to generate the best possible estimate of the parameters.

A complete description of the mathematical model and of the discretization and parameters estimation procedure is given in the following.

2.4 Stochastic modeling of expression kinetics and Kalman filtering

It is defined as $x_i(t)$ the time profile of the expression of gene i at time t. The underlying conservation equation simply stems from the observation that the rate of change of $x_i(t)$ with time, *i.e.* its time derivative $\dot{x}_i(t)$, must equal the difference between the production and degradation term. Based on experimental evidence [64], the degradation is well described by a first order term. The dynamics of the i-th transcript is therefore described by

$$\dot{x}_i\left(t\right) = -k_i x_i\left(t\right) + P_i(t) \tag{2.5}$$

where k_i is the mRNA decay rate of i-th messenger. This value is linked to the *half-life* h_i of the transcript by the relation $h_i = \ln(2)/k_i$. $P_i(t)$ is the i-th gene promoter activity regulated by transcription factors. Such regulation occurs by triggering or suppressing the transcription of the i-th gene, thus $P_i(t) \geq 0$. Moreover, the observed measure $y_i\left(t\right)$ is also a noisy time-series, so

$$y_i\left(t\right) = x_i\left(t\right) + \sigma_{y_i} n_{y_i}\left(t\right) \tag{2.6}$$

where σ_{y_i} is the standard deviation of measurements white noise n_{y_i}. Is is considered a generic pair of expression time profiles characterized by the presence of two terms: a stochastically correlated promoter activity $u(t)$ and a gene-specific term $n_{x_i}(t)$. Then the case is considered:

$$P_i(t) = b_i u(t) + \sigma_{x_i} n_{x_i}(t), \tag{2.7}$$

where b_i is a scaling factor accounting for the relative contribution to the overall promoter activity regarding gene i. The term $\sigma_{x_i} n_{x_i}$ models the part of the promoter activity which is not common to the pair. It is modeled this part by means of a noise term, n_{x_i} which is assumed to be a white noise. The common part $u(t)$ is modeled as a Wiener process:

$$\dot{u}(t) = \sigma_u n_u(t) \tag{2.8}$$

where $n_u\left(t\right)$ is white noise. Thus $u(t) = u(0) + \int_0^t \sigma_u n_u(\tau) d\tau$. The complete mathematical dynamic model for two transcripts j and h, together with their respective measurement equations, is

$$\begin{cases} \dot{x}_j\left(t\right) = -k_j x_j\left(t\right) + b_j u\left(t\right) + \sigma_{x_j} n_{x_j}\left(t\right) \\ \dot{x}_h\left(t\right) = -k_h x_h\left(t\right) + b_h u\left(t\right) + \sigma_{x_h} n_{x_h}\left(t\right) \\ \dot{u}(t) = \sigma_u n_u(t) \\ y_j\left(t\right) = x_j\left(t\right) + \sigma_y n_{y_j}\left(t\right) \\ y_h\left(t\right) = x_h\left(t\right) + \sigma_y n_{y_h}\left(t\right) \end{cases} \tag{2.9}$$

Now the linear dynamic system (2.9) can be rewrite using a compact matrix notation

$$\begin{cases} \dot{x}\left(t\right) = A\, x\left(t\right) + F\, N_x\left(t\right) \\ y\left(t\right) = C\, x\left(t\right) + G\, N_y\left(t\right) \end{cases} \tag{2.10}$$

where

$$x\left(t\right) = \begin{bmatrix} x_j\left(t\right) \\ x_h\left(t\right) \\ u\left(t\right) \end{bmatrix}, \qquad y\left(t\right) = \begin{bmatrix} y_j\left(t\right) \\ y_h\left(t\right) \end{bmatrix}$$

$$N_x\left(t\right) = \begin{bmatrix} n_{x_j}\left(t\right) \\ n_{x_h}\left(t\right) \\ n_u\left(t\right) \end{bmatrix}, \qquad N_y\left(t\right) = \begin{bmatrix} n_{y_j}\left(t\right) \\ n_{y_h}\left(t\right) \end{bmatrix}$$

and

$$A = \begin{bmatrix} -k_j & 0 & b_j \\ 0 & -k_h & b_h \\ 0 & 0 & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} \sigma_{x_j} & 0 & 0 \\ 0 & \sigma_{x_h} & 0 \\ 0 & 0 & \sigma_u \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad G = \begin{bmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{bmatrix}$$

Since the dynamic system (2.10) is linear, it can be exactly discretized (see [47]) for a given time interval Δ, corresponding to the time interval between two consecutive measurements. The k-th measurements corresponds to $t = k\Delta$, thus in the discretized system it can be used k in place of $k\Delta$, to keep the notation simple.

The solution of the linear dynamic system (2.10) is

$$x\left(t\right) = e^{A(t-t_0)} x\left(t_0\right) + \int_{t_0}^{t} e^{A(t-\tau)} F\left(\tau\right) N_x\left(\tau\right) d\tau \tag{2.11}$$

and its discretized form is

$$\begin{cases} x(k+1) = A_d\, x(k) + \Psi^{\frac{1}{2}}\, N_x(k) \\ y(k) = C\, x(k) + G\, N_y(k) \end{cases} \tag{2.12}$$

where

$$A_d = e^{A\Delta},$$

and Ψ is the covariance matrix defined by

$$\Psi = \int_0^{\Delta} e^{A\theta}\, F\, F^T\, e^{A^T\theta} d\theta. \tag{2.13}$$

The unknown parameters of the model to be estimated are k_i, b_i, σ_{x_i}, with $i = 1, 2$, σ_u and σ_y. The *state variables* of the system are $x_j(t)$, $x_h(t)$ and $u(t)$. For each given choice of the parameters it was used the Kalman filter [48] to estimate of the state variables.

The Kalman filter equation uses a feedback control strategy. It contains a *prediction term* for projecting forward (in time) the current state to obtain the *a priori* estimate, and a correcting term for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate

$$\hat{x}(k+1) = A_d\hat{x}\,(k) + K_P(\theta)\big[y(k+1) - CA_d\,\hat{x}(k-1)\big] \tag{2.14}$$

where $K_P(\theta)$ is the prediction Kalman gain that depends on the parameters θ of the stochastic equation.

For each choice of θ the Kalman filter is run. A probability value is associated to the resulting estimation. These values measures the probability that the current parametrization of the model generates the measured time series. Denoting by $\nu_o(k) = y(k) - CA_d\hat{x}(k-1)$ the *innovation* of the stochastic process, $\nu_o(k)$ is a sequence of independent gaussian random variables with covariance $\Psi_\nu = \left(CA_d^{-1}K_P + I\right)G^2$. The optimal set θ of parameters if therefore chosen according to a *maximum likelihood* criterion as the choice corresponding to the maximum of the *a priori* probability density of the innovation sequence. This corresponds to the minimum of the likelihood function

$$J\,(\theta) = n\ln\left(\det\left(\Psi_\nu\right)\right) + \sum_{k=1}^{n}\nu_o^T(k)\Psi_\nu^{-1}\nu_o(k).$$

where n is the number of samples. The half life $h_i = \ln(2)/k_i$ of the i-th messenger is the important information. To use all the available information and make the method robust with respect to measurement and estimation errors, the following algorithm was designed (as show in Figure 2.3):

1. Given a set of N mRNA time profiles, perform the maximum likelihood estimation for *every* pair (i, j) and compute the corresponding k_i and k_j.

2. For each pair (i, j) compute the ratio matrix M whose elements are the ratios between the half-lives of gene i and gene j. The matrix M is generally not symmetric due to the presence of outliers and numerical sensitivity. Thus it is defined the ratio estimation by row as $m^r_{i,j} = h^r_j/h^r_i$ and by column as $m^c_{j,i} = h^c_i/h^c_j$. The matrix M contains all the ratios h^r_j/h^r_i on the i-th row, and all the ratios h^c_i/h^c_j on the i-th column. Let us denote Σ^r_i the sum of the ith row, Σ^c_i the sum of the i column, and $< >$ the *mean* operator, that is,

$$\Sigma^r_i = \sum_{j=1}^{N}\frac{h^r_j}{h^r_i} = \frac{1}{h^r_i}\sum_{j=1}^{N}h^r_j, \qquad \Sigma^c_i = \sum_{j=1}^{N}\frac{h^c_i}{h^c_j} = h^c_i\sum_{j=1}^{N}\frac{1}{h^c_j}, \tag{2.15}$$

$$< \frac{1}{\Sigma^r_i} >= \frac{1}{N}\sum_{i=1}^{N}\frac{h^r_i}{\sum_{j=1}^{N}h^r_j} = \bar{h}\frac{1}{\sum_{j=1}^{N}h^r_j}, \tag{2.16}$$

$$< \Sigma^c_i >= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{h^c_i}{h^c_j} = \bar{h}\sum_{j=1}^{N}\frac{1}{h^c_j} \tag{2.17}$$

3. Given M, delete outliers to obtain a final matrix M^*. First, compute the probability density (using a smoothing kernel approach) of all the entries $m_{i,j}$ and delete those values below a probability of 10^{-3} of occurring in the distribution. Second, since ideally $m^r_{i,j} \cdot m^c_{j,i} = 1$, it is considered as outliers those pairs such that $m^r_{i,j} \cdot m^c_{j,i} \notin [\frac{1}{2},\ 2]$.

4. On the resulting M^* matrix compute for each transcript i two estimates of its half-life $\hat{h}^r_i$ and $\hat{h}^c_i$, using equations (2.15), (2.16)and (2.17). In this way it is obtained

$$\hat{h}^r_i = \bar{h}\frac{\frac{1}{\Sigma^r_i}}{<\frac{1}{\Sigma^r_i}>} \tag{2.18}$$

$$\hat{h}^c_i = \bar{h}\frac{\Sigma^c_i}{<\Sigma^c_i>}. \tag{2.19}$$

   This computation requires the value of $\bar{h}$. When this value is known for the group of transcripts under analysis the measured value can be used. Otherwise, letting $\bar{h} = 1$ one can obtain half-life values that are relative to the average half-life of the group. However, a third approach was used. All the results reported in this work have been obtained by replacing $\bar{h}$ with the geometric mean of $\hat{h}^r_i$ and $\hat{h}^c_i$, that is

$$\hat{\bar{h}} = \left(\prod_{i=1}^{N}\hat{h}^r_i \prod_{i=1}^{N}\hat{h}^c_i\right)^{\frac{1}{2N}}. \tag{2.20}$$

   The final estimate of the half-life $\hat{h}_i$ for the i-th gene is computed as the weighted average of $\hat{h}^r_i$ and $\hat{h}^c_i$ using as weights the respective variances $\sigma^2_{\hat{h}_r}$ and $\sigma^2_{\hat{h}_c}$ as follows

$$\hat{h}_i = \left(\frac{\hat{h}^r_i}{\sigma^2_{\hat{h}_r}} + \frac{\hat{h}^c_i}{\sigma^2_{\hat{h}_c}}\right) \cdot \left(\frac{1}{\sigma^2_{\hat{h}_r}} + \frac{1}{\sigma^2_{\hat{h}_c}}\right)^{-1} \tag{2.21}$$

   where $\sigma_{\hat{h}_r}$ and $\sigma_{\hat{h}_c}$ are the standard deviation of the $\hat{h}^r_i$ and $\hat{h}^c_i$, respectively.

5. It was considered as a quality index for each estimated half-life $\hat{h}_i$ the following:

$$C_i = \frac{1}{N}\sum_{j=1}^{N}\frac{\Psi_{ij}}{\sqrt{\Psi_{ii}\Psi_{jj}}}$$

   where Ψ_{ii}, Ψ_{jj} are noise variances of the discrete system (2.12) and Ψ_{ij} (see equation (2.13)) is the mutual covariance of the state noise between time series x_j and x_h. Thus, high values of C_i imply the presence of a correlation between $P_j(t)$ and $P_h(t)$ in equation (2.9). The half-lives having a C_i value smaller than the 10^{th} percentile of its distribution was removed.

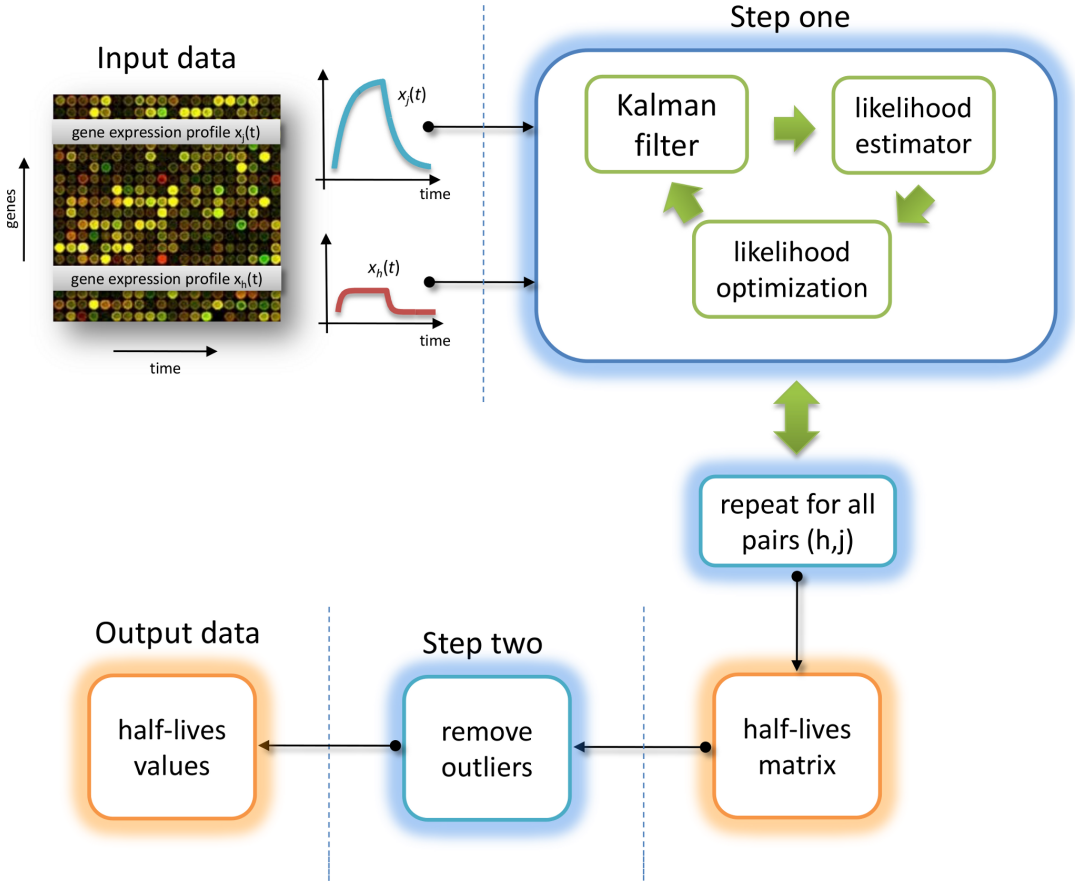

Figure 2.3: ***DRAGON algorithm pipeline.***

2.5 Results

2.5.1 Performance evaluation on Malaria IDC experimental data

The Intraerythrocytic Developmental Cycle is characterized by four morphologic stages: ring, trophozoite, schizont and late schizont. The cycle begins with the red cells invasion by merozoites followed by a remodeling of the host cell in the ring stage. The merozoites then develop into trophozoites. During the schizont stage, after a period of growth, the trophozoite undergoes an asexual dividing process and the parasite is ready for the next round of invasion by new merozoites (late schizont phase).

In [9] the authors, using microarrays, measured genome-wide mRNA abundance profiles across 48h during one cycle of *P. falciparum* IDC, collecting one sample per hour. Later on, Shock *et al.* [74] measured mRNA half-lives of 2774 transcripts of the IDC using chemical inhibitors to reach transcriptional shut-off.

The simultaneous availability of gene expression and decay data during the same biological process (IDC) represents a natural test bed for the validation of the DRAGON

algorithm. Therefore, DRAGON was applied on Bozdech *et al.* dataset to obtain mRNA stability estimations to be compared with Shock *et al.* measurements for performance evaluation. The resulting Pearson correlation between *in vitro* and *in silico* measures is $r = 0.6$ (P value $2.4 \cdot 10^{-61}$), and the first principal component accounts for 82% of the variability, thus showing a good performance for DRAGON algorithm (see Figure 2.4).

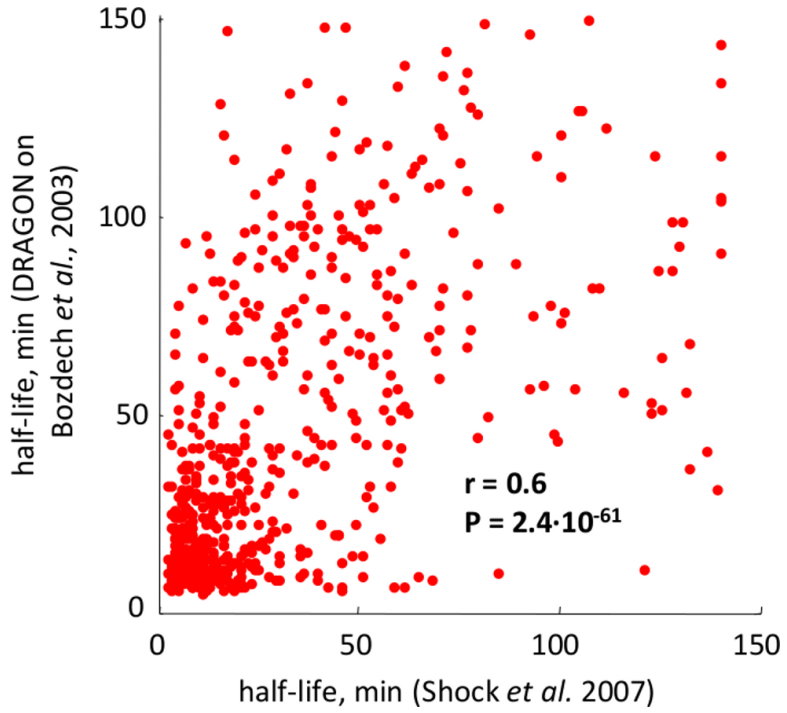

Figure 2.4: ***DRAGON algorithm validation using P. falciparum IDC data.*** *Scatterplot of mRNA half-lives for 616 genes estimated by DRAGON versus experimentally measured by Shock et al.*

However, since gene expression and decay data have been measured by different groups, it has been speculated that 18% of unexplained variability may be partly due to inherent biological variability and to transcriptional inhibition stress. As further analysis, it was computed average mRNA half-lives in both studies for functional categories (see Figure 2.5). Results show that the two studies are in better accordance when half-lives are averaged for all genes within any given functional category (Pearson correlation $r = 0.74$).

Remarkably, Shock *et al.*, found progressive stage-dependent average increases in mRNA stability and suggested such phenomenon to be a major determinant of mRNA accumulation (see Figure 2.6A). The same feature is also found using DRAGON estimated half-lives (see Figure 2.6B).

To investigate in further detail the behavior of average half-life of genes sequentially induced during IDC, it was computed for each gene the time point corresponding to its peak of expression and selected 48 groups of genes having peak of expression at each hourly time points over the 48 hours monitored by Shock *et al.*. For each gene group it was

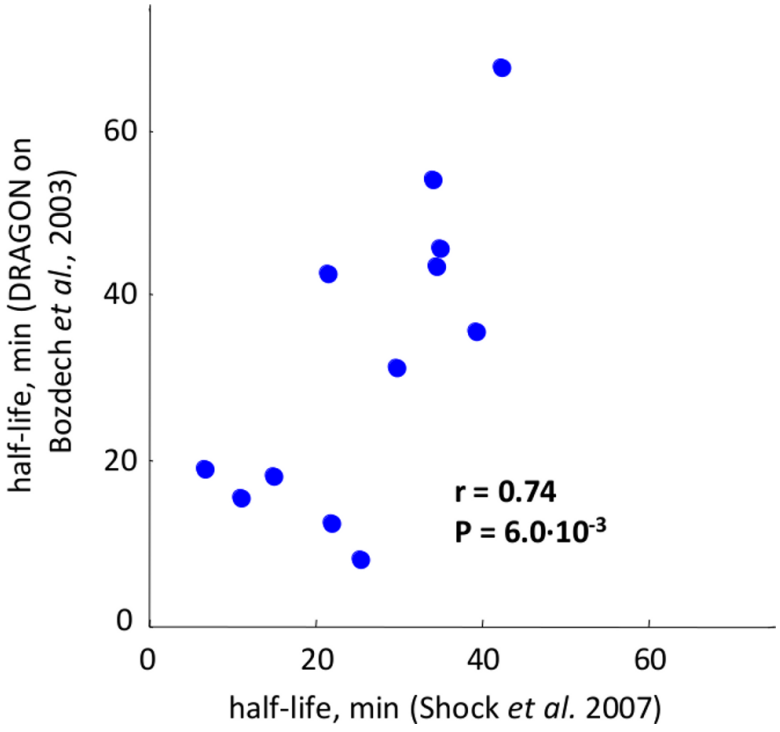

Figure 2.5: ***Functional categories analysis for DRAGON estimations using P. falciparum IDC data.*** *Average mRNA half-lives in both studies, DRAGON estimations versus and experimentally measured by Shock et al. half-lives, for 12 functional categories from the P. falciparum GO annotation database that are represented in the set of 616 transcripts by 5 or more members.*

computed half-lives mean and standard deviation and found a high correlation with the corresponding curve obtained using experimental data (Pearson correlation $r = 0.8$, P value $2.1 \cdot 10^{-10}$; see Figure 2.6C). Early responding genes are characterized by high instability, whereas late responders are more stable, as also reported by Elkon *et al.* in [25] when studying mammalian cells. A possible explanation for the presence of stable mRNAs at the schizont stage, suggested by Shock *et al.*, is that it may be important for the merozoite to receive a carefully regulated "starting package", that would allow rapid activation of the IDC following the next round of invasion. By contrast, the initial low mRNA stability values may be an indication of the fast dynamic remodeling after merozoite invasion. To evaluate the probability of obtaining such behavior by chance, the gene expression matrix was randomized and it was used DRAGON to estimate half-lives (see Figure 2.6D). Consistently, the estimation of half-lives using random data does not produce any correlation with experimental data (Pearson correlation $r = -0.17$).

2.5.2 Half-lives estimation during reproductive cycle in *S. cerevisiae*

Gene expression during yeast cell cycle has been recently measured by Pramila *et al.* [69] using alpha-factor synchronization and by Orlando *et al.* [66] using centrifugal elutriation

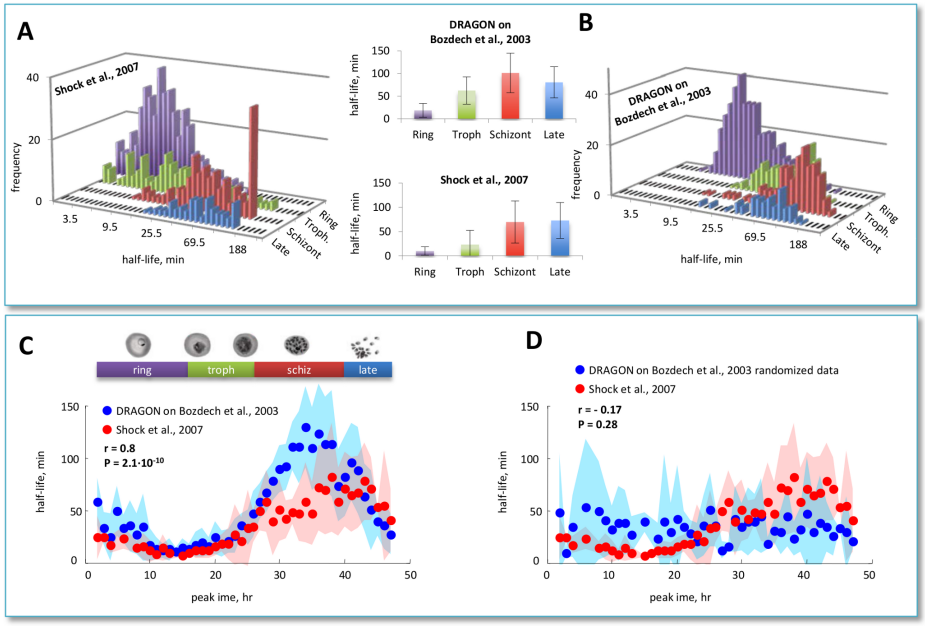

Figure 2.6: ***Periodic behavior of average half-lives of sequentially induced genes in P. falciparum IDC.*** *(A) Histograms of mRNA half-lives for genes induced at each stage of the P. falciparum IDC as experimentally measured by Shock et al. and (B) estimated by DRAGON algorithm. The inset panels show mean and standard deviation of half-lives during each stage. Both studies show an increase of average transcript stabilities of sequentially induced genes during P. falciparum IDC. (C) Average experimental and estimated half-life values (red and blue dots, respectively) corresponding to genes having the same expression peak timing, indicated on the x-axis. Standard deviations are drawn as pale blue and pale red stripes. The two curves both show a maximum peak of average half-life value during the schizont stage and a minimal value during the ring stage. A sharp increase of average half-life occurs during the trophozoite stage. The Person correlation between experimental and DRAGON estimated curve is , thus showing a good agreement between the two studies. (D) Effect of randomizing the gene expression matrix on DRAGON estimated half-lives.*

for synchronization. It was obtained a high consistency of DRAGON estimations using data for 569 transcripts over replicate datasets (Pearson correlation $r = 0.83$ for Pramila *et al.* dataset and Pearson correlation $r = 0.92$ for the Orlando *et al.* dataset; see Figure 2.7A-B).

The larger variability in half-lives estimations may be explained by the inconsistencies between replicate time-series in the Pramila *et al.* dataset with respect to the Orlando *et al.* dataset (see Figure 2.7C).

Notwithstanding significant differences in synchronization procedures, it was also found a high correlation of DRAGON half-lives estimations over the two datasets (Pearson correlation $r = 0.57$, P value $2.1 \cdot 10^{-39}$; see Figure 2.7D) where the first principal component accounts for 79% of the overall variability. It has been speculate that 21% of unexplained variability may be partly due to the different synchronization methods used. In fact, Orlando *et al.* obtained a cell cycle duration of about 2 hours, 8 samples per cycle, whereas

Pramila *et al.* obtained a cell cycle duration of about 1 hour, 12 samples per cycle. Consistently, most of the transcripts during the slower cycle display higher half-lives when compared to the fastest cycle (see Figure 2.7D).

2.5.3 GO annotations of genes with extreme half-lives in *S. cerevisiae*

In this paragraph functional annotations (done using GOrilla software [24]) of novel predicted half-lives provided by DRAGON algorithm using yeast reproductive and metabolic cycle time series are briefly discussed. For the yeast cell cycle it was normalized the half-life log-distribution (Z-score), for each dataset, and then computed the geometric mean to obtain a single half-life value for each gene. Notably, the averaging has the effect of reducing the impact of the different synchronization stress response.

Unstable genes are enriched with replication fork complex (p-value $5.95 \cdot 10^{-4}$) and stable genes (histones HA1-2,HB1-2) are enriched with nucleosome (p-value $7.4 \cdot 10^{-4}$). This is consistent with the need of producing a large number of histones during DNA replication process so that stable histone mRNAs contribute to a higher translation efficiency. Moreover, DNA replication timing requires first the formation of the replication fork, then the production of the needed histones for chromatin assembling: such temporal sequence of events is consistent with a rapid turnover of the replication complex genes and a slow turnover of the histone genes. Among unstable genes it was also found the G1/S transition cyclins and among stable ones we found G2/M transition cyclins. In this case, the temporal sequence of events is the progression of the cell cycle from DNA replication to mitosis.

For the yeast metabolic cycle many stable mRNA species involved in the organic acid and arginine metabolism and protein catabolic processes were found. Among unstable messengers we found genes involved in DNA repair (p-value $4.85 \cdot 10^{-8}$), DNA metabolism (p-value $7.5 \cdot 10^{-8}$) and chromatin silencing (p-value $4.29 \cdot 10^{-3}$).

2.6 Discussion

2.6.1 Periodic behavior of average half-lives of sequentially induced genes

The increasing pattern of average half-life found during *P. falciparum* IDC (shown in Figure 2.6A) motivated us to investigate whether a periodicity could be found also in other cyclical biological processes. The attention was focused on the reproductive cell cycle and the metabolic cycle in *Saccharomyces cerevisiae*, for which high resolution time series measurements are available on public repositories.

To study if a periodic pattern of average half-life of sequentially induced genes exists along the cell cycle progression, for each gene we computed the time points at which maximal expression is attained. Thus, it was obtained, for each time point, the list of genes having expression peak value at that time and computed the corresponding mean and variance of DRAGON estimated half-lives values. Indeed, a cyclic behavior along sequentially induced genes in both datasets was found (see Figure 2.8A for the Pramila *et al.* dataset and Figure 2.8B for the Orlando *et al.* dataset).

Synchronization methods, cell cycle duration and number of samples are different between the two cited studies, but, reassuringly, the phases of the cell cycle at which mean half-life is minimal or maximal is consistent. In fact, for both datasets it was observed a cyclical increase of mean half-life from G1 phase to M phase and a subsequent decrease back to G1. The figure clarifies that the minimal mean half-life is reached at the G1/S transition, whereas the maximal value correspond to the M/G2 phase for both cycles and datasets. The latter is consistent with the observation that, in higher eukaryotes, mitosis is accompanied by global repression of nuclear RNA synthesis, indicating that mRNAs must be stable to be inherited from daughter cells.

The yeast metabolic cycle has been recently studied by Tu *et al.* [80] using a continuous culture system, after a brief starvation period, the culture spontaneously began persistent respiratory cycles of about 5 hours. In the same study, a genome-wide microarray gene expression measurement was performed. Samples were taken every 25 minutes for 3 consecutive cycles. Using DRAGON algorithm it was estimated half-lives using data of 1043 transcripts. Surprisingly, also in this case a cyclical pattern for mean half-life of sequentially induced genes was found. The maximum peak is located at the RC phase and the minimum peak located at RB phase (see Figure 2.8C).

2.6.2 Integrated analysis–sequential waves of co-ordinated transcription and decay

Recently, the appearance of a number of studies has revealed the fundamental role of stability regulation in shaping appropriate cell response. A key point has been recently addressed by Shalem *et al.* [73], who have shown the dynamic co-ordinated interplay between transcription and degradation. They have found in yeast two basic regulatory strategies in response to stress. More precisely, they measured changes of mRNA abundance and decay rates in a yeast population subjected to oxidative and DNA damage stress. By grouping genes according to the time point at which the maximal (minimal) fold change is attained and combining normalized (mean and variance) mRNA abundance and decay rate data, they constructed a "stability *versus* folding" (SF) diagram where change in mRNA stability relative to a reference state (mean value in our case) is plotted against the maximal fold change. Using yeast expression time-course data obtained in response to an oxidative stress and a DNA damage, they were able to reveal two different strategies: a) a "counteracting regulation" strategy (see Figure 2.9A), characterized by genes in which an increase (decrease) in degradation rates counteracts a increase (decrease) in mRNA abundance, *i.e.* repressed genes are stabilized and induced genes are destabilized; b) a "synergistic regulation" strategy (see Figure 2.9B), characterized by genes in which an increase (decrease) in degradation rates is associated with an decrease (increase) in mRNA abundance, *i.e.* induced genes are stabilized and repressed genes are destabilized.

Shalem *et al.* also found that, progressing from early time points forward, the negative correlation (counteracting) was replaced with a positive correlation (synergistic). Such co-ordination strategy may permit crosstalk between different steps of mRNA biogenesis, providing a mechanism to control the order and timing of events. The work of Shalem *et al.* has shown the importance of combining expression data with decay rates under the

same experimental condition to reveal the underlying strategy of co-ordination of the two "regulatory arms", namely transcription and degradation. Uncovering such relationships is certainly a fundamental task, since the underlying reciprocal influences between mRNA production and degradation are largely unexplored. The DRAGON algorithm, by estimating half-lives directly from gene expression data under specific conditions, allows the computational integration of mRNA abundance and decay rates data, making this powerful combined analysis possible when experimentally measured half-lives are not available.

It was computed SF diagrams for *P. falciparum* IDC, yeast cell cycle (Pramila *et al.* dataset) and metabolic cycle (shown in Figure 2.108). In panels A,C and E each blue dot corresponds to a Pearson correlation of the SF diagram at the peak time point indicated on the x-axis, for the three datasets. In panels B,D and F the SF diagrams corresponding to the correlation values indicated by the arrows in panels A,C and E are displayed. The arrows point to maximal negative (red dots in panels B,D,F and red arrows in panels A,C,E) and maximal positive correlation values (green dots in panels B,D,F and green arrows in panels A,C,E).

Strikingly, in all cases the same conclusions of Shalem *et al.* were reached, namely it was found that early induced genes show counteracting regulation, whereas late induced genes show a synergistic regulation.

2.6.3 Advantages and disadvantages of the method

The main advantage of the DRAGON algorithm consists in the estimation of the mRNA half-lives directly from gene expression time-course during condition-specific experiments. Moreover it estimates the correlation among promoter activities between pairs of genes. Another advantage of the algorithm lies in its robustness. Specifically, it was observed that even if the accuracy of the absolute values of the estimated half-lives can be influenced by many factors (such as the number of points in the time series, the accuracy of the measurements, the time interval between samples, the choice of the thresholds for the outliers, etc.), the ranking of half-lives is insensitive to the factors mentioned above.

The main disadvantages are the following: DRAGON can work only with time-series under the same experimental condition and cannot handle steady-state values under different conditions. As a general rule a reliable estimate requires at least 10-12 time samples, *i.e.* a number significantly larger than the number of parameters to be estimated (this rule is not obviously always applicable as the required number of points depends strongly on the signal to noise ratio) and a sampling time not larger than the expected average half-life. If no information is available about the correlation of promoter activities, as a rule of thumb, a set of at least 50-100 time series must be processed together in order to have reliable half-lives estimates. One basic hypothesis is that the half-life of a transcript is approximately constant during the time course of the experiment, thus a substantial change of its value would yield an unreliable estimate. These problems can be handled by performing more measurements using a shorter sample time, or by considering moving time windows. The computational overhead can be significant: for a sample of 1000 time series there are 10^6 pairs to analyze, requiring a computation of about 150 hours on a medium-speed

single-processor machine capable of analyzing 2 pairs per second.

2.7 Conclusion

This analysis supports and strengthens Shalem *et al.* conclusions about the coordination of transcriptional and mRNA degradation in the cell in response to stress. it was demonstrated that during periodic processes, such as the *P. falciparum* IDC, the reproductive cell cycle and the metabolic cycle, the alternative interplays between changes in mRNA stability and changes in mRNA abundance are activated by periodically switching from a counteracting to a synergistic regulation. In light of these results, the classical vision of periodic processes as the result of serial transcription factor sequential activation, should be re-considered from a broader point of view by including post-transcriptional regulation and coordination.

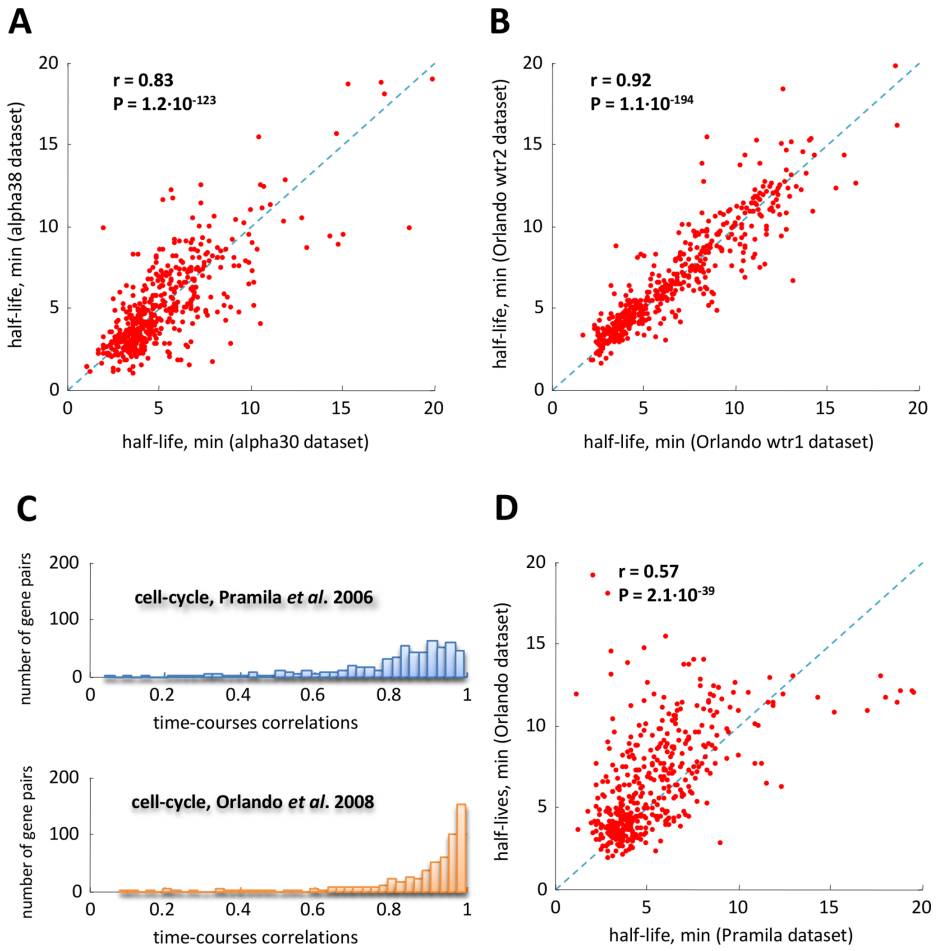

Figure 2.7: ***DRAGON performance over 569 cell cycle regulated genes from the Pramila and Orlando datasets.*** *Scatterplot of the DRAGON estimated half-lives using two replicates taken from (A) Pramila et al. dataset (denoted by alpha30 and alpha38) and from (B) Orlando et al. dataset (denoted by Orlando wtr1 and Orlando wtr2) and using (D) the Pramila et al. and the Orlando et al. dataset. (C) Histograms of the Pearson correlation values between time series relative to each gene in two replicates. Orlando et al. dataset shows a better consistency between replicates with respect to the Pramila et al. dataset. The half-lives obtained by replicate datasets have been averaged. The half-lives estimated using Orlando dataset show slightly higher values with respect to those obtained using Pramila dataset, as shown by the deviation from the bisector line (dashed blue line). This is consistent with the slower cell cycle in Orlando experiment (2 hours) compared to that of Pramila experiment (1 hour).*

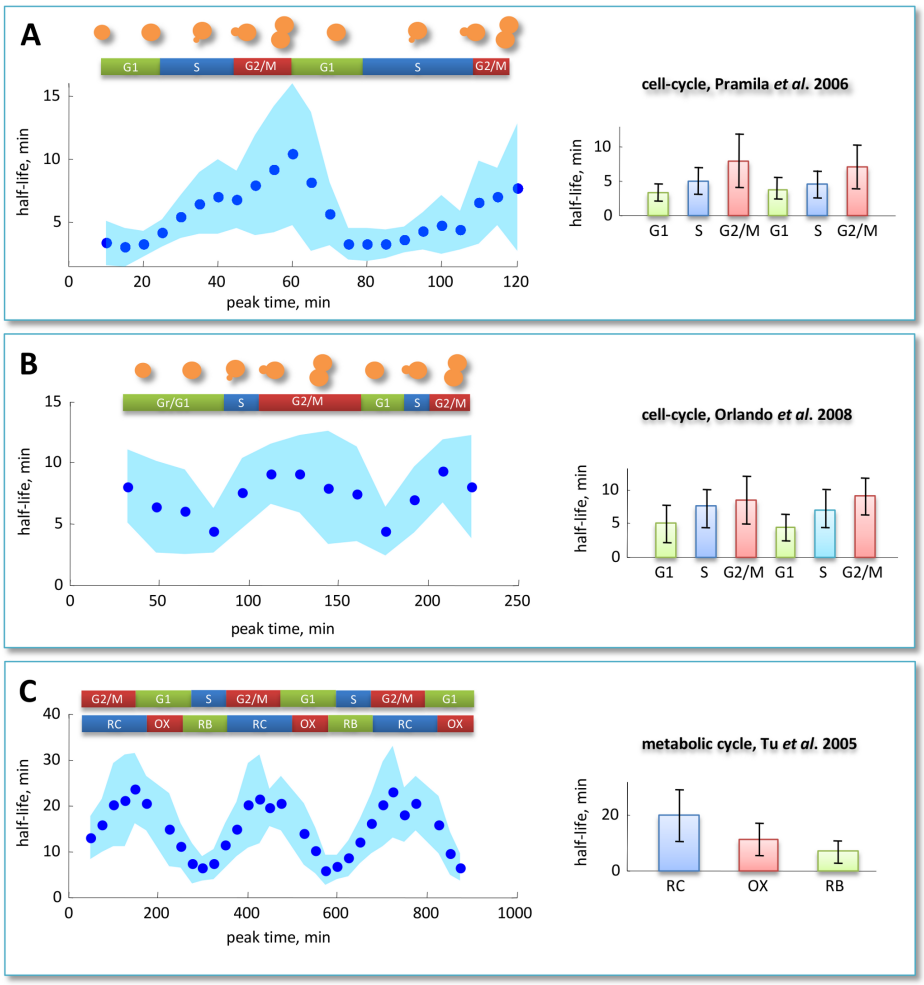

Figure 2.8: ***Periodic behavior of average half-life during the reproductive and metabolic cycle.*** *Average DRAGON estimated messenger half-life values corresponding to genes having the same expression peak timing, indicated on the x-axis (dark blue points). Standard deviations are drawn as pale blue stripes. (A) yeast cell cycle using Pramila dataset, (B) yeast cell cycle using Orlando dataset and (C) yeast metabolic cycle. Strikingly, in all datasets the maximal average half-life is attained for genes induced during G2/M phase, including the metabolic cycle. The minimal average half-life, in all datasets, is attained at the G1/S transition phase. The bar charts show mean and standard deviation of half-lives during each stage.*

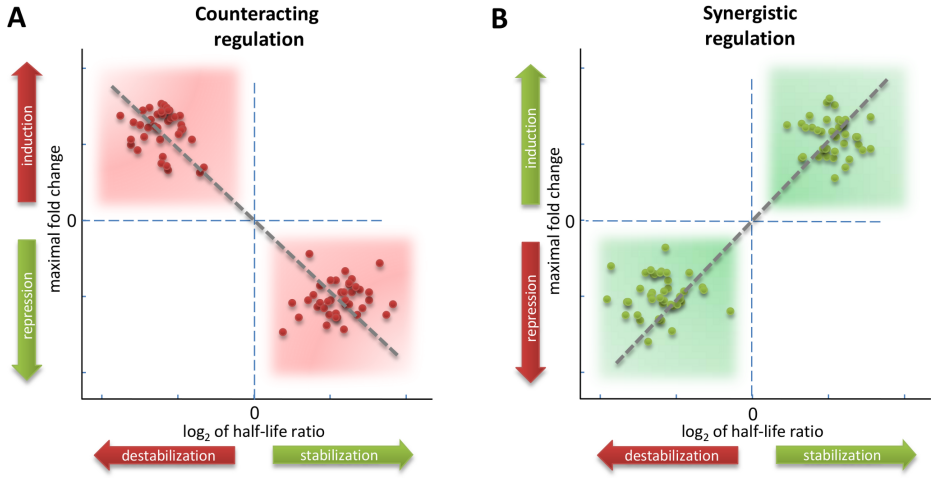

Figure 2.9: ***Illustration of the counteracting and synergistic regulatory strategy.*** *The stability/folding diagrams (SF), introduced by Shalem et al., show the change in mRNA stability relative to the average value plotted against the maximal fold change. (A) Counteracting strategy (negative correlation): induced genes are destabilized and repressed genes are stabilized. (B) Synergistic strategy (positive correlation): induced genes are stabilized and repressed genes are destabilized.*

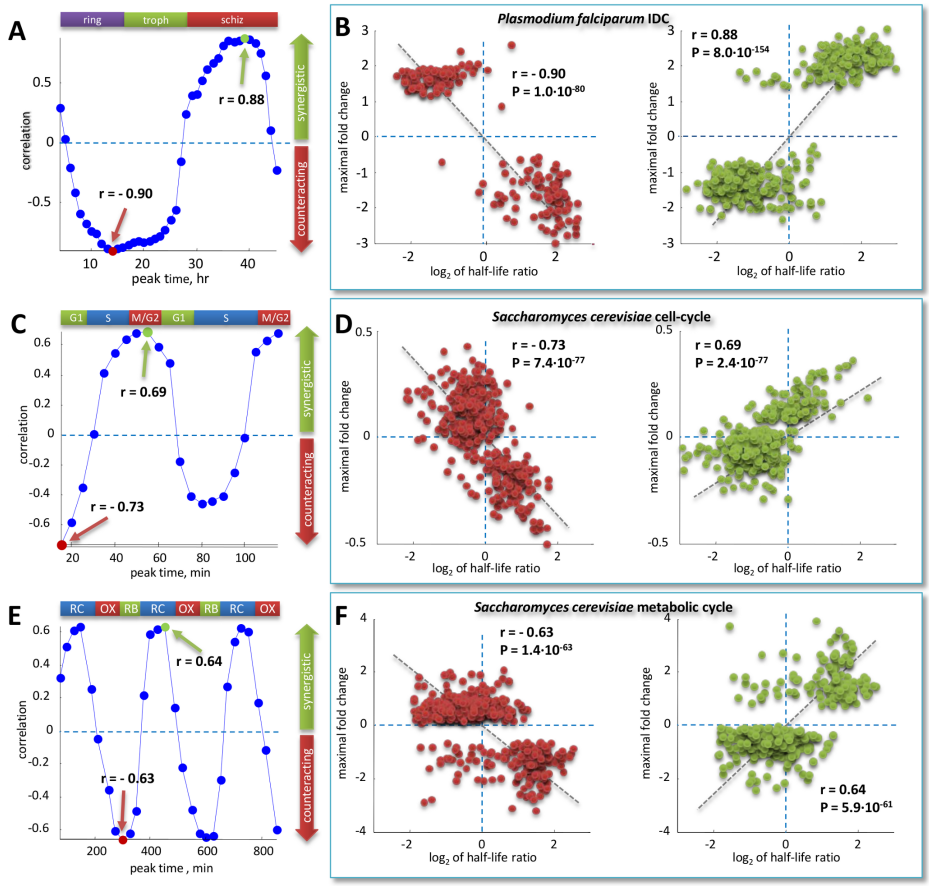

Figure 2.10: ***Temporal progress on the stability/folding (SF) diagrams.*** *In all cases under study, a progressive shift from an inverse (counteracting) to a direct (synergistic) relationship was found.*

CHAPTER 3

Virtual Output

In this chapter the solutions of two classical estimation problems for deterministic and stochastic systems is introduced through the use of the new concept of *virtual measurement map*. This concept is formalized by the following definitions:

Definition 3.0.1 (Stochastic system)**.** *Let*

$$x(k+1) = Ax(k) + Bu(k) + Fw(k), \tag{3.1}$$
$$y(k) = h(x(k), u(k), v(k)), \tag{3.2}$$

describe a dynamical system whose state $x(k)$ evolves in $\mathbb{R}^n$, $y(k) \in \mathbb{R}^q$ is the available measurement vector, and $u(k) \in \mathbb{R}^p$, $w(k) \in \mathbb{R}^r$, and $v(k) \in \mathbb{R}^s$ are the deterministic input, the state noise, and the output measurement error vectors, respectively. Such a system is said to admit a virtual measurement map if there exist a completely observed sequence $\{y_v(k)\}$, named virtual output, a $q \times n$ matrix $C(k)$, and a random white sequence $\{n(k)\}$, not necessary Gaussian, endowed with all moments up to a certain order 2ν, and uncorrelated with the state sequence $\{x(k)\}$, such that, for any $k \geq 0$, one has:

$$y_v(k) = C(k)x(k) + n(k). \tag{3.3}$$

Definition 3.0.2 (Deterministic system)**.** *Let*

$$x(k+1) = Ax(k) + Bu(k), \tag{3.4}$$
$$y(k) = h(x(k), u(k)), \tag{3.5}$$

describe a dynamical system whose state $x(k)$ evolves in $\mathbb{R}^n$, $y(k) \in \mathbb{R}^q$ is the available measurement vector, and $u(k) \in \mathbb{R}^p$ is the control input. Such a system is said to admit a virtual measurement map if there exist a completely observed sequence $\{y_v(k)\}$, named virtual output, and a $q \times n$ matrix $C(k)$, such that, for any $k \geq 0$, one has:

$$y_v(k) = C(k)x(k). \tag{3.6}$$

3.1 Planar Tracking

3.1.1 Introduction

Passive localization and target tracking are common problems of nonlinear filtering. The general aim is to estimate the time-varying state of a moving object basing on the noisy nonlinear measurements of radius and angle. Because of these nonlinearities, the problem is intrinsically infinite dimensional so that no implementable optimal algorithm can be defined. As a consequence, the problem has been classically solved through the suboptimal extended Kalman filter (EKF) algorithm [4,5] which is based on the first-order linearization of the nonlinear system. Unfortunately, the linearization may introduce large errors in the state estimate that sometimes cause divergence of the filter [13, 82].

A first attempt to reduce the linearization error is suggested in [77], but it is not useful for the target tracking system. In [37] a modified covariance EKF (MVEKF) algorithm is developed, where it is suggested to re-linearize the nonlinear measurement equations at the filtered state in order to reduce the linearization approximation errors.

An alternative strategy to improve the EKF performances is given by the iterated extended Kalman filter (IEKF) [5] which implements a repeated correction procedure to the standard EKF. A crucial point of this method is the choice of the number of repetitions since it is strongly connected with the convergence of the filter.

A substantial improvement is obtained by the unscented Kalman filter (UKF) [44, 45], which is based on the key idea of approximating the probability distribution instead of the nonlinear system equations. This intuition leads up to a filtering algorithm able to overcome the EKF in a wide range of applications [46, 81] with comparable computational complexity. In [85] an iterated version of the UKF (IUKF) is proposed. This extension is similar to that given by the IEKF for the EKF, but the introduced algorithm has better convergence properties and a consistent termination criterion. Obviously, this is paid in terms of computational complexity which grows linearly with the state dimension.

For real-time applications the additive computational load given by extra iterative procedures (e.g IEKF, IUKF) may be considered a significant limitation. Indeed, the simplicity of EKF and UKF is often preferred to the higher accuracy of more complex methods.

The mentioned methods belong to the wide class of filters that use any analytical approximation. A second class is composed by numerical Monte Carlo methods [7,29,35,40, 43,63], also known as particle filters (PF) since they are based on point mass (or "particle") representations of probability densities [3], without introducing any linearization. As well known, the estimation accuracy is improved with the increase of the number of particles, at expanse of a higher computational complexity.

In this work, the linearization procedure has been avoided by following a different philosophy; the main idea is to use range and angle measurements as time-varying parameters of the output matrix and of a "virtual" output vector. Therefore, no linearization needs to be introduced on the treatment of the measurement data at expense of losing the gaussianity of the measurement noise even if the original one is Gaussian. However, it is worth to stress that the whiteness of the transformed measurement noise is preserved. This allows standard Kalman filtering to be adopted for the state estimation. Such an algorithm constitutes, as it is well known, the linear least square estimator given the defined virtual measurements. As far as the nongaussianity of the transformed noise is concerned, the use of a polynomial Kalman like algorithm [11, 12] allows improving the performances of the linear one. Of course, in this case, the complexity of the algorithm could become less negligible.

3.1.2 Problem Formulation

The stochastic description of the kinematic 2-D movement of an object has received a great attention in the relevant scientific literature and several mathematical models have been proposed [5, 8, 28, 49, 55, 56, 58, 67, 68, 75]. In this work, the coordinates of the target at the time k will be denoted by $(p_1(k), p_2(k))$. Most models can be described by the following discrete-time stochastic linear system:

$$x(k+1) = Ax(k) + Fa(k), \qquad x(0) = x_0, \tag{3.7}$$

where

$$x(k) = \left[\begin{array}{cccccc} p_1(k) & \dot{p}_1(k) & \ddot{p}_1(k) & p_2(k) & \dot{p}_2(k) & \ddot{p}_2(k) \end{array} \right]^T, \tag{3.8}$$

and $\{a(k)\}$ represents the dynamical stochastic forcing process assumed to be an independent stationary sequence of zero mean random vectors in $\mathbb{R}^2$. Moreover, the statistical moments of $\{a(k)\}$ are supposed to be known up to a given order 2ν and summarized by vectors

$$\psi_a^{(i)} := \mathbb{E}\left[a^{[i]}(k)\right], \quad i = 1, 2, \ldots, 2\nu, \tag{3.9}$$

where $[i]$ stands for the the Kronecker i-th power (see Definition A.0.2). A and F are matrices of proper dimension whose structure depends on the chosen model. The initial state x_0 is supposed to be a Gaussian random vector with mean value $\bar{x}$ and covariance matrix Ψ_{x_0}.

As usual, the measurement process, at each time instant k, is given by the noisy values $\rho_m(k)$ and $\theta_m(k)$ of the polar coordinates $\rho(k)$ and $\theta(k)$ of the target:

$$\rho_m(k) = \rho(k) + n_\rho(k) = \sqrt{x_1^2(k) + x_2^2(k)} + n_\rho(k), \tag{3.10}$$

$$\tan(\theta(k)) = \tan(\theta_m(k) - n_\theta(k)) = \frac{x_2(k)}{x_1(k)}, \tag{3.11}$$

where $n_\rho(k)$ and $n_\theta(k)$ denote the measurement errors. Both the noise sequences are supposed to be characterized by known distributions, whose moments are indicated by

$$\psi_\rho^{(i)} := \mathbb{E}\left[n_\rho^i(k)\right], \quad \psi_\theta^{(i)} := \mathbb{E}\left[n_\theta^i(k)\right], \quad i = 1, 2 \ldots. \tag{3.12}$$

Finally, the initial state x_0 forms, together with the zero mean sequences $\{a(k)\}$, $\{n_\rho(k)\}$, and $\{n_\theta(k)\}$, a family of independent variables. In the following the term white will be referred to a sequence of uncorrelated random vectors.

The design objective of the 2-D tracking problem (2-DTP) is to reconstruct the kinematic state $x(k)$ of the moving target from the nonlinear noisy measurements $\rho_m(k)$ and $\theta_m(k)$. Since the measurement process is nonlinear, the state estimation is, in general, an infinite dimensional problem which requires a nonlinear algorithm. Therefore only suboptimal solutions can be used for engineering applications, such as EKF, UKF, or PF. The novelty suggested in this work is the transformation of the above specified nonlinear problem in a form amenable to be treated by a linear processing of data, without introducing any linearization.

3.1.3 A Virtual Model for the 2-DTP

In order to get the new estimation algorithm, the concept of *virtual measurement map* is used through the definition 3.0.2.

In the following, a linear stochastic discrete-time dynamical system endowed with a virtual measurement map will be referred as *virtual linear stochastic model*.

Moreover, in order to formalize an optimality criterion for the filtering of virtual linear stochastic models the following definition is given.

Definition 3.1.1. *Consider a linear stochastic system endowed with a nonlinear noisy output transformation which admits a virtual measurement map y_v. Linear Virtual Measurement Kalman Filter (LVMKF) denotes the best minimum variance linear algorithm with respect to y_v. Moreover, ν-PVMKF denotes the best minimum variance polynomial algorithm, of order ν, with respect to y_v.*

Definition 3.0.2 introduces to the main idea of this work: to apply standard KF to the virtual linear stochastic model. The so obtained algorithm will constitute, as it is well known, the linear least square estimator given the defined virtual measurements. In order to exploit this idea for the 2-DTP, following theorem shows the existence of a virtual measurement map for this particular case.

Theorem 1. *The tracking system defined by (3.7), (3.10), and (3.11) admits a virtual measurement map with:*

$$y_v(k) = R^T(\theta_m(k)) \begin{bmatrix} \rho_m(k) \\ 0 \end{bmatrix}, \qquad C(k) = C = \bar{R}S, \tag{3.13}$$

$$n(k) = \mathring{R}(k)x(k) + R^T(\theta_m(k))Gn_\rho(k), \tag{3.14}$$

where

$$R(\alpha) := \begin{bmatrix} cos(\alpha) & sin(\alpha) \\ -sin(\alpha) & cos(\alpha) \end{bmatrix}, \qquad \bar{R} := \mathbb{E}\left[R(-n_\theta(k))\right] \tag{3.15}$$

$$\mathring{R}(k) := \left(R\left(-n_\theta(k)\right) - \bar{R}\right) S, \tag{3.16}$$

$$S = \begin{bmatrix} I_2 & 0_2 & 0_2 \end{bmatrix}, \qquad G = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{3.17}$$

The measurement error covariance matrix is given by

$$\Psi_n(k) = st^{-1}\left(\mathbb{E}\left[\mathring{R}(k)^{[2]}\right] st\left(\Psi_x(k)\right)\right) + \psi_\rho^{(2)} R^T(\theta_m(k)) G G^T R(\theta_m(k)), \tag{3.18}$$

with

$$\Psi_x(k) = A\Psi_x(k-1)A^T + F st^{-1}\left(\psi_a^{(2)}\right) F^T, \qquad \Psi_x(0) = \Psi_{x_0}, \tag{3.19}$$

where $st(\cdot)$ and $st^{-1}(\cdot)$ denote the stack operator and its inverse defined in Definition A.0.3.

Proof. Taking account that

$$\begin{aligned} x_1(k) cos(\theta(k)) &= \rho(k)\ cos^2(\theta(k)) \\ x_2(k) sin(\theta(k)) &= \rho(k)\ sin^2(\theta(k)) \end{aligned} \tag{3.20}$$

and the output equation (3.10), it readily follows that

$$\rho_m(k) - n_\rho(k) = \rho(k) = cos(\theta_m(k) - n_\theta(k)) x_1(k) + sin(\theta_m(k) - n_\theta(k)) x_2(k). \tag{3.21}$$

Moreover, from (3.11) it follows:

$$0 = -sin(\theta_m(k) - n_\theta(k)) x_1(k) + cos(\theta_m(k) - n_\theta(k)) x_2(k). \tag{3.22}$$

Now, from (3.21) and (3.22) it results that

$$\begin{bmatrix} \rho_m(k) \\ 0 \end{bmatrix} = R(\theta_m(k) - n_\theta(k)) S x(k) + G n_\rho(k), \tag{3.23}$$

where $R(\cdot)$, S, and G are the matrices defined in (3.15) and (3.17). Note that matrix $R(\cdot)$ satisfies the following properties for any $\alpha, \beta \in \mathbb{R}$:

$$R(\alpha + \beta) = R(\alpha) R(\beta), \tag{3.24}$$

$$R^{-1}(\alpha) = R^T(\alpha) = R(-\alpha). \tag{3.25}$$

By exploiting properties (3.24)-(3.25), (3.23) can be rewritten in the form

$$R^T\left(\theta_m\left(k\right)\right) \begin{bmatrix} \rho_m(k) \\ 0 \end{bmatrix} = R\left(-n_\theta(k)\right) S x(t) + R^T\left(\theta_m\left(k\right)\right) G n_\rho(k). \tag{3.26}$$

The left-hand side of last equation corresponds to the virtual output $y_v(k)$ given in (3.13). Finally, through a simple manipulation, it results:

$$y_v(k) = \bar{R} S x(t) + \left(R\left(-n_\theta(k)\right) - \bar{R}\right) S x(t) + R^T\left(\theta_m\left(k\right)\right) G n_\rho(k), \tag{3.27}$$

which corresponds to the virtual measurement map given in (3.13) and (3.14).

It remains to prove that the random sequence $\{n(k)\}$ is zero mean, white and uncorrelated with $\{x(k)\}$. This readily derives by observing that $\{n(k)\}$ is a combination of the three independent random sequences: $\{\mathring{R}(k)\}$, $\{n_\rho(k)\}$ and the state sequence $\{x(k)\}$, where the fist two are zero mean and white.

As far as (3.19) and (3.18) are concerned, by using (A.10) in Lemma A.0.1 and taking into account (3.9), one has:

$$\begin{aligned}\Psi_x(k) &:= \mathbb{E}\left[x(k)x^T(k)\right] \\ &= \mathbb{E}[(Ax(k-1)+Fa(k-1))(Ax(k-1)+Fa(k-1))^T] \\ &= A\Psi_x(k-1)A^T + Fst^{-1}\left(\psi_a^{(2)}\right)F^T,\end{aligned} \tag{3.28}$$

and, thanks to (A.9)-(A.10) in Lemma A.0.1 and the statistical properties of $\{\mathring{R}(k)\}$, $\{n_\rho(k)\}$ and $\{x(k)\}$:

$$\begin{aligned}\Psi_n(k) &:= \mathbb{E}\left[n(k)n^T(k)\right] \\ &= \mathbb{E}\left[\mathring{R}(k)x(k)x^T(k)\mathring{R}^T(k)\right] + \psi_\rho^{(2)}R^T(\theta_m(k))GG^TR(\theta_m(k)) \\ &= st^{-1}\left(\mathbb{E}\left[\mathring{R}(k)^{[2]}st\left(x(k)x^T(k)\right)\right]\right) + \psi_\rho^{(2)}R^T(\theta_m(k))GG^TR(\theta_m(k)) \\ &= st^{-1}\left(\mathbb{E}\left[\mathring{R}(k)^{[2]}\right]st\left(\Psi_x(k)\right)\right) + \psi_\rho^{(2)}R^T(\theta_m(k))GG^TR(\theta_m(k)).\end{aligned} \tag{3.29}$$

Now it is possible to state the first main result of the work that readily follows from Theorem 1 and Definition 3.1.1.

Proposition 1. *For the 2-DTP there exists the virtual model representation*

$$x(k+1) = Ax(k) + Fa(k) \tag{3.30}$$

$$y_v(k) = Cx(k) + n(k) \tag{3.31}$$

where $y_v(k)$, C and $n(k)$ are given by (3.13)-(3.17) *and, therefore, Kalman filter applied to this model constitutes the LVMKF for the 2-DTP.*

Of course, such proposal is not optimal with respect to all Borel transformations of the virtual measurements because of the nongaussianity of $n(k)$. By literature, this problem may be faced by several different approaches (e.g. UKF [44], PF [35], [11,12,59,78]). The idea followed in this work, in order to improve the performances of the linear algorithm, is looking for the best polynomial transformation of the virtual measurements to get a state estimation, obtaining, in this way, the ν-PVMKF defined in Definition 3.1.1. This solution can be considered as intermediate among Kalman filtering and the infinite dimensional optimal filtering.

3.1.4 Polynomial Extension

The first step for building the announced filter is to introduce the extended polynomial virtual measurement vector and the corresponding extended state.

Definition 3.1.2. *For the system model (3.7), (3.10) and (3.11), the extended output* $\mathcal{Y}_v(k)$ *and the extended state* $\mathcal{X}(k)$ *are defined as*

$$\mathcal{Y}_v(k) = \begin{bmatrix} \mathcal{Y}_v^{(1)}(k) \\ \mathcal{Y}_v^{(2)}(k) \\ \vdots \\ \mathcal{Y}_v^{(\nu)}(k) \end{bmatrix} \in \mathbb{R}^Q \ , \ \mathcal{X}(k) = \begin{bmatrix} x(k) \\ x^{[2]}(k) \\ \vdots \\ x^{[\nu]}(k) \end{bmatrix} \in \mathbb{R}^N, \tag{3.32}$$

where ν *is the filter order,* $Q = 2(2^\nu - 1)$, $N = 6(6^\nu - 1)/5$, *and*

$$\mathcal{Y}_v^{(i)}(k) = y_v^{[i]}(k) - R^{T^{[i]}}(\theta_m(k))G^{[i]}\psi_\rho^{(i)}. \tag{3.33}$$

Note that $\mathcal{Y}_v^{(i)}(k)$ contains information regarding the i-th power of $y_v(k)$. The idea of the polynomial extension is indeed based on the use of a standard linear estimation algorithm through the definition of a particular extended model. This last is required to have a completely linear structure and to be constituted by an output and a state vector composed by all Kronecker powers of the original output and state vectors up to a given order ν. In this way, such a model results to be amenable to the application of the standard linear Kalman filter, which constitute, in a natural way, the best minimum variance polynomial algorithm for the considered system.

In the following, Theorem 5, together with the statistical characterization given in Theorem 3, will formally show that this holds true for the extended output and state vectors defined in (3.32) for the 2-DTP.

Before to present the announced result, it is necessary to introduce some useful definitions and lemmas.

Definition 3.1.3. *Let* $\mathcal{M}$ *be the set of all finite dimensional matrices, then the map* Λ_ν, $\nu \geq 1$, *is defined as*

$$\Lambda_\nu(M) = \begin{bmatrix} M^{[1]} & 0 & \cdots & 0 \\ 0 & M^{[2]} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M^{[\nu]} \end{bmatrix}, \qquad M \in \mathcal{M}. \tag{3.34}$$

Lemma 3.1.1. *Let* $\Pi_{\nu,n}^{(i)} \in \mathbb{R}^{n^\nu \times r}$, $i, n, \nu \geq 1$, $r = n(n^\nu - 1)/(n-1)$ *be the matrix*

$$\Pi_{\nu,n}^{(i)} := \begin{bmatrix} 0_{n^i \times n} & \cdots & 0_{n^i \times n^{i-1}} & I_{n^i} & 0_{n^i \times n^{i+1}} & \cdots & 0_{n^i \times n^\nu} \end{bmatrix}, \tag{3.35}$$

where $0_{p\times q}$ *denotes the* $p \times q$ *null matrix. Given* $M \in \mathbb{R}^{r\times s}$, $L \in \mathbb{R}^{s\times q}$, $N \in \mathbb{R}^{n\times n}$, $v \in \mathbb{R}^n$, *and*

$$\mathcal{V} = \begin{bmatrix} v \\ v^{[2]} \\ \vdots \\ v^{[\nu]} \end{bmatrix}, \tag{3.36}$$

the following properties hold true:

$$\Lambda_{\nu}^{[i]}(ML) = \Lambda_{\nu}^{[i]}(M)\Lambda_{\nu}^{[i]}(L), \tag{3.37}$$

$$\Lambda_{\nu}(N) = \begin{bmatrix} N^{[1]}\Pi_{\nu,n}^{(1)} \\ N^{[2]}\Pi_{\nu,n}^{(2)} \\ \vdots \\ N^{[\nu]}\Pi_{\nu,n}^{(\nu)} \end{bmatrix}, \tag{3.38}$$

$$v^{[i]} = \Pi_{\nu,n}^{(i)}\mathcal{V}. \tag{3.39}$$

Proof. Property (3.37) can be readily proved by exploiting the diagonal structure of operator Λ_{ν} and noting that for any pair of suitable dimensioned matrices M and L it results that

$$(ML)^{[i]} = M^{[i]}L^{[i]}, \tag{3.40}$$

by iteratively applying property (A.7) in Lemma A.0.1. Properties (3.38) and (3.39) can be proved by direct computation.

Lemma 3.1.2. *Let $N_{\rho}(k) \in \mathbb{R}^{\nu}$ be defined as*

$$N_{\rho}(k) = \begin{bmatrix} n_{\rho}(k) & n_{\rho}^{2}(k) & \cdots & n_{\rho}^{\nu}(k) \end{bmatrix}^{T} \in \mathbb{R}^{\nu}. \tag{3.41}$$

Then:

$$\bar{N}_{\rho} := \mathbb{E}\left[N_{\rho}\left(k\right)\right] = \left[\psi_{\rho}^{(1)}, \psi_{\rho}^{(2)}, \ldots, \psi_{\rho}^{(\nu)}\right], \tag{3.42}$$

$$\bar{R}^{(i)} := \mathbb{E}\left[R^{[i]}\left(-n_{\theta}\left(k\right)\right)\right] = U^{[i]}L_{i}V^{[i]}, \tag{3.43}$$

$$\bar{\Lambda} := \mathbb{E}\left[\Lambda_{\nu}(R(-n_{\theta}(k))S)\right] = \begin{bmatrix} \bar{R}^{(1)}S & 0 & \cdots & 0 \\ 0 & \bar{R}^{(2)}S & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{R}^{(\nu)}S \end{bmatrix}, \tag{3.44}$$

with

$$V^{-1} = U = \begin{bmatrix} -j & j \\ 1 & 1 \end{bmatrix}, \qquad L_{i} = diag\left(\left[\mathbb{E}\left[e^{jl_{i}(q)n_{\theta}(k)}\right], \, q = 1, 2, \ldots, 2^{i}\right]\right), \tag{3.45}$$

where j denotes the imaginary unit, and $l_{i}(q) \in \mathbb{Z}$ satisfies

$$l_{i}(q) = 2\sum_{m=1}^{i-1}\left|\left[\frac{q-1}{2^{i-m}}\right]\right|_{2} + 2|q-1|_{2} - i, \tag{3.46}$$

in which $[\cdot]$ and $|\cdot|_{2}$ denote integer part and 2-modulo, respectively.

Proof. As far as $\bar{N}_\rho$ is concerned, it is enough to note that such a vector is composed by all the moments of the zero mean random variable $n_\rho(k)$ up to the order ν, indicated as defined in (3.9).

Regarding $\bar{R}^{(i)}$, it is easy to verify that the eigenvalues of matrix $R(\cdot)$ defined in (3.15), are

$$\lambda_1(k) = e^{-jn_\theta(k)},\ \lambda_2(k) = e^{jn_\theta(k)}, \tag{3.47}$$

and that the diagonalizing matrix U, constituted by the corresponding eigenvectors, is given by (3.45). Therefore, (3.43) holds true by defining L_i as

$$L_i := \mathbb{E}\left[diag\left(v^{[i]}(k)\right)\right], \text{with } v(k) = \left[\begin{array}{cc} \lambda_1(k) & \lambda_2(k) \end{array}\right]^T. \tag{3.48}$$

From Lemma A.0.6, it results that the q-th entry of $v^{[i]}(k)$, satisfy

$$\{v^{[i]}\}_q = v_{l_1} v_{l_2} \cdot \cdots \cdot v_{l_i}, \tag{3.49}$$

where coefficients $l_m \in \{1, 2\}$ $(m = 1, 2, \ldots, i)$ are given by (A.22) in Lemma A.0.6. By taking into account (3.49), it follows that

$$\{v^{[i]}\}_q = exp\left\{jn_\theta(k)\left[p_{i,2}(q) - p_{i,1}(q)\right]\right\}, \tag{3.50}$$

where $p_{i,1}(q)$ and $p_{i,2}(q)$ denote the cardinalities of the sets of the all coefficients l_m equal to 1 and 2, respectively. From (A.22), it results that

$$\begin{aligned} p_{i,2}(q) &= \sum_{m=1}^{i} (l_m - 1) = \sum_{m=1}^{i-1} \left| \left[\frac{q-1}{2^{i-m}} \right] \right|_2 + |q-1|_2, \\ p_{i,1}(q) &= i - p_2(q), \end{aligned} \tag{3.51}$$

from which one has:

$$l_i(q) := p_{i,1}(q) - p_{i,1}(q) = 2 \sum_{m=1}^{i-1} \left| \left[\frac{q-1}{2^{i-m}} \right] \right|_2 + 2|q-1|_2 - i. \tag{3.52}$$

Finally, from (3.50) and (3.52), it follows that

$$\left\{\mathbb{E}\left[v^{[i]}(k)\right]\right\}_q = \mathbb{E}\left[e^{jl_i(q)n_\theta(k)}\right]. \tag{3.53}$$

As far as $\bar{\Lambda}$ is concerned, (3.44) trivially follows from definitions in (3.34) and (3.43).

Theorem 2. *The extended output and extended state (3.32) defined for the system model (3.7), (3.10) and (3.11), satisfy the following model equations:*

$$\begin{aligned} \mathcal{X}(k+1) &= \mathcal{A}\mathcal{X}(k) + \mathcal{U} + \mathcal{F}(k), \\ \mathcal{Y}_v(k) &= \mathcal{C}(k)\mathcal{X}(k) + \mathcal{N}(k), \end{aligned} \tag{3.54}$$

where

$$\mathcal{A}=\begin{bmatrix} A & 0 & \cdots & 0 \\ \mathcal{O}_{2,1} & A^{[2]} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{O}_{\nu,1} & \mathcal{O}_{\nu,2} & \cdots & A^{[\nu]} \end{bmatrix},\ \mathcal{U}=\begin{bmatrix} 0 \\ F^{[2]}\psi_a^{(2)} \\ \vdots \\ F^{[\nu]}\psi_a^{(\nu)} \end{bmatrix},\ \mathcal{F}(k)=\begin{bmatrix} f_1(k) \\ f_2(k) \\ \vdots \\ f_\nu(k) \end{bmatrix} \tag{3.55}$$

$$\mathcal{C}(k)=\begin{bmatrix} \mathcal{C}^{(1)}(\theta_m(k)) \\ \mathcal{C}^{(2)}(\theta_m(k)) \\ \vdots \\ \mathcal{C}^{(\nu)}(\theta_m(k)) \end{bmatrix},\ \mathcal{N}(k)=\begin{bmatrix} \mathcal{N}^{(1)}(k) \\ \mathcal{N}^{(2)}(k) \\ \vdots \\ \mathcal{N}^{(\nu)}(k) \end{bmatrix}. \tag{3.56}$$

with

$$\mathcal{O}_{i,l}=M^i_{i-l,6}(F^{[i-l]}\otimes A^{[l]})(\psi_a^{(i-l)}\otimes I_n^{[l]}), \tag{3.57}$$

$$f_i(k)=\sum_{l=0}^{i-1} M^i_{i-l,6}(F^{[i-l]}\otimes A^{[l]})((a^{[i-l]}(k)-\psi_a^{(i-l)})\otimes I_n^{[l]})x^{[l]}(k), \tag{3.58}$$

$$\mathcal{C}^{(i)}(\theta_m(k))=\mathcal{M}^{(i)}_{\nu,2}\left(\bar{\Lambda}\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(I_N\otimes\bar{N}_\rho\right)+\bar{R}^{(i)}S^{[i]}\Pi^{(i)}_{\nu,6}, \tag{3.59}$$

$$\begin{aligned}\mathcal{N}^{(i)}(k)=&\,\mathcal{M}^{(i)}_{\nu,2}\left(\Lambda(k)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k)\otimes\mathring{N}_\rho(k)\right)\\&+\mathcal{M}^{(i)}_{\nu,2}\left(\mathring{\Lambda}\left(k\right)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k)\otimes\bar{N}_\rho\right)\\&+\mathring{R}^{(i)}(k)S^{[i]}\Pi^{(i)}_{\nu,6}\mathcal{X}(k)+R^{T[i]}(\theta_m(k))G^{[i]}\Pi^{(i)}_{\nu,1}\mathring{N}_\rho(k).\end{aligned} \tag{3.60}$$

in which $M^i_{k,n}$ and $\mathcal{M}^{(i)}_{\nu,n}$ are defined in Appendix by (A.14) and (A.27), respectively, and

$$\Lambda(k):=\Lambda_\nu\left(R\left(-n_\theta\left(k\right)\right)S\right), \tag{3.61}$$

$$\mathring{N}_\rho(k):=N_\rho(k)-\bar{N}_\rho\in\mathbb{R}^\nu, \tag{3.62}$$

$$\mathring{\Lambda}(k):=\Lambda(k)-\bar{\Lambda}\in\mathbb{R}^{Q\times N}, \tag{3.63}$$

$$\mathring{R}^{(i)}\left(k\right):=R^{[i]}\left(-n_\theta\left(k\right)\right)-\bar{R}^{(i)}\in\mathbb{R}^{2^i\times 2^i}, \tag{3.64}$$

$$\Gamma\left(\theta_m\left(k\right)\right):=\Lambda_\nu\left(R^T\left(\theta_m\left(k\right)\right)G\right)\in\mathbb{R}^{Q\times N}. \tag{3.65}$$

Moreover, $\{\mathcal{F}(k),k=1,2,3,\ldots\}$ and $\{\mathcal{N}(k),k=1,2,3,\ldots\}$ are uncorrelated zero mean white sequences and the second is also uncorrelated with $\{\mathcal{X}(k),k=1,2,3,\ldots\}$.

Proof. Concerning $\mathcal{A},\mathcal{U}$, and $\mathcal{F}(k)$ in (3.55), one can refer to [11]. As far as $\mathcal{C}(k)$ and $\mathcal{N}(k)$ in (3.56) are concerned, first recall (3.26) in the form:

$$y_v(k)=R(-n_\theta(k))Sx(k)+R^T(\theta_m(k))Gn_\rho(k). \tag{3.66}$$

By applying Lemma A.0.7 to (3.33), and taking into account (3.66) and (3.61), it results:

$$\begin{aligned}\mathcal{Y}_v^{(i)}(k) =& \mathcal{M}_{\nu,2}^{(i)}\left(\Lambda(k)\otimes\Lambda_\nu\left(R^T\left(\theta_m\left(k\right)\right)G\right)\right)\left(\mathcal{X}\left(k\right)\otimes N_\rho\left(k\right)\right)\\ &+ R^{[i]}(-n_\theta\left(k\right))S^{[i]}x^{[i]}\left(k\right) + R^{T[i]}(\theta_m\left(k\right))G^{[i]}n_\rho^i\left(k\right)\\ &- R^{T[i]}(\theta_m(k))G^{[i]}\psi_\rho^{(i)}(k).\end{aligned} \tag{3.67}$$

Moreover, using definitions (3.62)-(3.65) and properties in Lemma A.0.1 and in (3.39), one has:

$$\begin{aligned}\mathcal{Y}_v^{(i)}(k) =& \mathcal{M}_{\nu,2}^{(i)}\left(\bar{\Lambda}\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(I_N\otimes\bar{N}_\rho\right)\mathcal{X}\left(k\right)\\ &+ \mathcal{M}_{\nu,2}^{(i)}\left(\Lambda\left(k\right)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}\left(k\right)\otimes\mathring{N}_\rho\left(k\right)\right)\\ &+ \mathcal{M}_{\nu,2}^{(i)}\left(\mathring{\Lambda}\left(k\right)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}\left(k\right)\otimes\bar{N}_\rho\right)\\ &+ \mathring{R}^{(i)}\left(k\right)S^{[i]}\Pi_{\nu,6}^{(i)}\mathcal{X}\left(k\right) + \bar{R}^{(i)}S^{[i]}\Pi_{\nu,6}^{(i)}\mathcal{X}\left(k\right)\\ &+ R^{T[i]}(\theta_m\left(k\right))G^{[i]}\Pi_{\nu,1}^{(i)}\mathring{N}_\rho\left(k\right) + R^{T[i]}(\theta_m(k))G^{[i]}\Pi_{\nu,1}^{(i)}\bar{N}_\rho\\ &- R^{T[i]}(\theta_m(k))G^{[i]}\psi_\rho^{(i)}\\ =& \mathcal{C}^{(i)}(\theta_m\left(k\right))\mathcal{X}\left(k\right) + \mathcal{N}^{(i)}(k)\end{aligned} \tag{3.68}$$

which proves (3.54).

To finish the proof, the whiteness of $\{\mathcal{F}(k)\}$ is demonstrated in [11] meanwhile the whiteness of $\{\mathcal{N}(k)\}$ and its uncorrelation with $\{\mathcal{X}(k)\}$ is readily proved by observing that: $\{\Lambda(k)\}$, $\{X(k)\}$ and $\{\mathring{N}_\rho(k)\}$ are independent; $\{\mathring{R}^{(i)}(k)\}$ is independent of $\{X(k)\}$ and $\{\mathring{N}_\rho(k)\}$; and $\{\mathring{\Lambda}(k)\}$ and $\{\mathring{N}_\rho(k)\}$ are zero mean and white. Moreover, $\{\mathcal{F}(k)\}$ and $\{\mathcal{N}(k)\}$ are uncorrelated because of the independence of the state noise $a(k)$ and the measurement noises $n_\rho(k)$ and $n_\theta(k)$.

In order to build the optimal linear filter for the extended model (3.54) it is necessary to verify that the extended state and output noises admit finite covariances. Before to provide such matrices it is useful to characterize the statistics of the random sequences $\{\mathring{N}_\rho(k)\}$ and $\{\mathring{\Lambda}(k)\}$ through the following Lemma.

Lemma 3.1.3. *The second order statistical matrices $\Psi_\rho \in \mathbb{R}^{\nu\times\nu}$ and $\Psi_\theta \in \mathbb{R}^{N^2\times N^2}$ defined as*

$$\Psi_\rho = \mathbb{E}\left[\mathring{N}_\rho(k)\mathring{N}_\rho^T(k)\right] \tag{3.69}$$

$$\Psi_\theta = \mathbb{E}\left[\mathring{\Lambda}^{[2]}(k)\right], \tag{3.70}$$

satisfy:

$$\{\Psi_\rho\}_{r,s} = \psi_\rho^{(r+s)} - \psi_\rho^{(r)}\psi_\rho^{(s)} \tag{3.71}$$

$$\Psi_\theta = \Lambda_\nu^{[2]}(U)\bar{\Lambda}^{(2)}\Lambda_\nu^{[2]}(VS) - \bar{\Lambda}^{[2]} \tag{3.72}$$

where $\psi_\rho^{(i)}$, U, *and* V *are defined in (3.9) and (3.45) and* $\bar{\Lambda}^{(2)}$ *is given by:*

$$\bar{\Lambda}^{(2)} = diag\left[\sum_{\substack{h_1,\dots,h_\nu \geq 0 \\ h_1+\dots+h_\nu=2}} M^2_{h_1,\dots,h_\nu}\left[\left(\bigotimes_{i=1}^{\nu}\left(\Pi^{(i)T}_{\nu,2}\right)^{[h_i]}\right) diag\left(L_{h_1+2h_2+\dots+\nu h_\nu}\right)\right]\right] \tag{3.73}$$

with L_i *defined in (3.45).*

Proof. As far as Ψ_ρ is concerned, the (r,s) entry is given by

$$\begin{aligned}\{\Psi_\rho\}_{r,s} &= \mathbb{E}\left[\{\mathring{N}_\rho(k)\}_r\{\mathring{N}_\rho(k)\}_s\right] \\ &= \mathbb{E}\left[\left(n^r_\rho(k)-\psi^{(r)}_\rho\right)\left(n^s_\rho(k)-\psi^{(s)}_\rho\right)\right] = \psi^{(r+s)}_\rho - \psi^{(r)}_\rho\psi^{(s)}_\rho.\end{aligned} \tag{3.74}$$

Recalling that U and V diagonalize matrix $R(-n_\theta(k))$ and by exploiting property (3.37) , it results

$$\mathbb{E}\left[\Lambda^{[2]}_\nu\left(R(-n_\theta(k))S\right)\right] = \mathbb{E}\left[\Lambda^{[2]}_\nu\left(U diag(v(k))VS\right)\right] \tag{3.75}$$

$$= \Lambda^{[2]}_\nu(U)\,\mathbb{E}\left[\Lambda^{[2]}_\nu\left(diag(v(k))\right)\right]\Lambda^{[2]}_\nu(VS)\,, \tag{3.76}$$

where $v(k)$ is the same vector defined in (3.48); moreover, by observing that

$$\Lambda^{[2]}_\nu(diag(v(k))) = diag(\mathcal{V}^{[2]}(k)),\ \text{with}\ \mathcal{V}(k) = \begin{bmatrix} v(k) \\ v^{[2]}(k) \\ \vdots \\ v^{[\nu]}(k)\end{bmatrix}, \tag{3.77}$$

Lemma A.0.5 implies that

$$\bar{\Lambda}^{(2)} := \mathbb{E}\left[\Lambda^{[2]}_\nu(diag(v(k)))\right] = \mathbb{E}\left[diag(\mathcal{V}^{[2]}(k))\right] \tag{3.78}$$

satisfies (3.73) since $L_i = \mathbb{E}[diag(v^{[i]}(k))]$, as defined in (3.48). Finally, by recalling definitions (3.44), (3.61), and (3.63), one has:

$$\mathbb{E}\left[\mathring{\Lambda}^{[2]}(k)\right] = \mathbb{E}\left[\Lambda^{[2]}(R(-n_\theta(k))S)\right] - \bar{\Lambda}^{[2]} \tag{3.79}$$

that can be readily proved corresponding to (3.70) by substituting (3.76) to the first term and taking into account (3.78).

Theorem 3. *The covariances of* $\{\mathcal{F}(k)\}$ *and* $\{\mathcal{N}(k)\}$ *are given by:*

$$\mathcal{Q}(k) := \mathbb{E}\left[\mathcal{F}(k)\mathcal{F}^T(k)\right],\ \mathcal{R}(k) := \mathbb{E}\left[\mathcal{N}(k)\mathcal{N}^T(k)\right] \tag{3.80}$$

where

$$\mathcal{Q}_{r,s}(k) = \sum_{l=0}^{r-1}\sum_{m=0}^{s-1} M^r_{r-l,6}(F^{[r-l]}\otimes A^{[l]})P^{r,s}_{l,m}(k)(F^{[s-m]}\otimes A^{[m]})(M^s_{s-m,6})^T \tag{3.81}$$

with

$$\begin{aligned} P_{l,m}^{r,s}(k) =& st^{-1}\Big[(I_2^{[s-m]} \otimes C_{2^{r-l},6^m}^T \otimes I_6^{[l]}) \\ & \cdot \left((\psi_a^{(s+r-m-l)} - \psi_a^{(s+m)} \otimes \psi_a^{(r-l)}) \otimes C_{1,6^m} \otimes I_6^{[l]}\right) \Pi_{\nu,6}^{(l+m)} \mu_{\mathcal{X}}(k)\Big], \end{aligned} \tag{3.82}$$

and where

$$\mathcal{R}(k) = \sum_{\ell=1}^{4}\sum_{m=\ell+1}^{4}\left(\Phi_{\ell,m}(k) + \Phi_{\ell,m}^T(k)\right) + \sum_{\ell=1}^{4}\Phi_{\ell,\ell}(k), \tag{3.83}$$

with the non-zero terms satisfying:

$$\begin{aligned} \Phi_{1,1}(k) &= \mathcal{M}_{\nu,2}\left(st^{-1}\left((\Psi_\theta + \bar{\Lambda}^{[2]})st(\Psi_{\mathcal{X}}(k))\right) \otimes \Xi(k)\right)\mathcal{M}_{\nu,2}^T, \\ \Phi_{1,4}(k) &= \mathcal{M}_{\nu,2}\left(\left(\bar{\Lambda}\mu_{\mathcal{X}}(k)\right) \otimes \Xi(k)\right), \\ \Phi_{2,2}(k) &= \mathcal{M}_{\nu,2}\left(\Theta(k) \otimes \left(\Gamma\left(\theta_m\left(k\right)\right)\bar{N}_\rho\bar{N}_\rho^T\Gamma^T\left(\theta_m\left(k\right)\right)\right)\right)\mathcal{M}_{\nu,2}^T, \\ \Phi_{2,3}(k) &= \mathcal{M}_{\nu,2}\left(\Theta(k) \otimes \left(\Gamma\left(\theta_m\left(k\right)\right)\bar{N}_\rho\right)\right), \\ \Phi_{3,3}(k) &= \Theta(k), \\ \Phi_{4,4}(k) &= \Xi(k), \end{aligned} \tag{3.84}$$

in which

$$\begin{aligned} \Theta(k) =& st^{-1}\left(\Psi_\theta st\left(\Psi_{\mathcal{X}}(k)\right)\right), \\ \Xi(k) =& \Gamma\left(\theta_m\left(k\right)\right)\Psi_\rho\Gamma^T\left(\theta_m\left(k\right)\right), \\ \mathcal{M}_{\nu,2} =& \left[\begin{array}{cccc} \mathcal{M}_{\nu,2}^{(1)^T} & \mathcal{M}_{\nu,2}^{(2)^T} & \cdots & \mathcal{M}_{\nu,2}^{(\nu)^T} \end{array}\right]^T \in \mathbb{R}^{Q\times Q^2}, \end{aligned} \tag{3.85}$$

$$\mu_{\mathcal{X}}(k) := \mathbb{E}[\mathcal{X}(k)] = \mathcal{A}\mu_{\mathcal{X}}(k-1) + \mathcal{U}, \tag{3.86}$$

$$\Psi_{\mathcal{X}}(k) := \mathbb{E}[\mathcal{X}(k)\mathcal{X}^T(k)] \tag{3.87}$$

$$= \mathcal{A}\Psi_{\mathcal{X}}(k-1)\mathcal{A}^T + \mathcal{Q}(k-1) \tag{3.88}$$

$$+\mu_{\mathcal{X}}(k-1)\mathcal{U}^T + \mathcal{U}\mu_{\mathcal{X}}^T(k-1). \tag{3.89}$$

Proof. Concerning $\mathcal{Q}(k)$ one can refer to [11].

From the expression of the components of $\mathcal{N}(k)$, given in (3.60), and using property (3.38), it can be noted that the whole $\mathcal{N}(k)$ consists of the sum of the following terms:

$$\begin{aligned} \mathcal{L}_1(k) &= \mathcal{M}_{\nu,2}\left(\Lambda\left(k\right) \otimes \Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k) \otimes \mathring{N}_\rho(k)\right), \\ \mathcal{L}_2(k) &= \mathcal{M}_{\nu,2}\left(\mathring{\Lambda}\left(k\right) \otimes \Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k) \otimes \bar{N}_\rho\right), \\ \mathcal{L}_3(k) &= \mathring{\Lambda}(k)\mathcal{X}(k), \\ \mathcal{L}_4(k) &= \Gamma\left(\theta_m\left(k\right)\right)\mathring{N}_\rho(k). \end{aligned} \tag{3.90}$$

Equation (3.83) holds true by introducing the notation

$$\Phi_{\ell,m}(k) = \mathbb{E}\left[\mathcal{L}_\ell(k)\mathcal{L}_m^T(k)\right], \quad \ell, m = 1, \ldots, 4. \tag{3.91}$$

In order to prove expression (3.84), considering $\Phi_{1,1}(k)$, one has:

$$\Phi_{1,1}(k) = \mathcal{M}_{\nu,2}\tilde{\Phi}_{1,1}(k)\mathcal{M}_{\nu,2}^T, \tag{3.92}$$

with

$$\begin{aligned}
\tilde{\Phi}_{1,1}(k) =&\mathbb{E}\left[\left(\Lambda\left(k\right)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k)\otimes\mathring{N}_\rho(k)\right)\left(\left(\Lambda\left(k\right)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k)\otimes\mathring{N}_\rho(k)\right)\right)^T\right]\\
=&\mathbb{E}\left[\left(\Lambda\left(k\right)\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(\mathcal{X}(k)\otimes\mathring{N}_\rho(k)\right)\left(\mathcal{X}^T(k)\otimes\mathring{N}_\rho^T(k)\right)\left(\Lambda^T\left(k\right)\otimes\Gamma^T\left(\theta_m\left(k\right)\right)\right)\right]\\
=&\mathbb{E}\left[\Lambda(k)\mathcal{X}(k)\mathcal{X}^T(k)\Lambda^T(k)\right]\otimes\left(\Gamma(\theta_m(k))\mathbb{E}\left[\mathring{N}_\rho(k)\mathring{N}_\rho^T(k)\right]\Gamma^T(\theta_m(k))\right)\\
=&\mathbb{E}\left[st^{-1}\left(\Lambda^{[2]}(k)st\left(\mathcal{X}(k)\mathcal{X}^T(k)\right)\right)\right]\otimes\left(\Gamma(\theta_m(k))\Psi_\rho\Gamma^T(\theta_m(k))\right)\\
=&st^{-1}\left(\mathbb{E}\left[\Lambda^{[2]}(k)\right]\right)st\left(\mathbb{E}\left[\mathcal{X}(k)\mathcal{X}^T(k)\right]\right)\otimes\Xi(k)\\
=&\mathcal{M}_{\nu,2}\left(st^{-1}\left((\Psi_\theta+\bar{\Lambda}^{[2]})st(\Psi_\mathcal{X}(k))\right)\otimes\Xi(k)\right)\mathcal{M}_{\nu,2}^T,
\end{aligned} \tag{3.93}$$

where the independence of $\{\Lambda(k)\}$, $\{\mathcal{X}(k)\}$, and $\{\mathring{N}_\rho(k)\}$, and definitions (3.69) and (3.70) have been taken into account. All the other terms of (3.83) can be obtained following the same line. It only will result useful to rewrite in advance $\mathcal{L}_3(k)$ and $\mathcal{L}_4(k)$ in the form

$$\begin{aligned}
\mathcal{L}_3(k) &= \left(\mathring{\Lambda}(k)\otimes 1\right)\left(\mathcal{X}(k)\otimes 1\right),\\
\mathcal{L}_4(k) &= \left(1\otimes\Gamma\left(\theta_m\left(k\right)\right)\right)\left(1\otimes\mathring{N}_\rho(k)\right).
\end{aligned} \tag{3.94}$$

As far as (3.86) and (3.89) are concerned, one has:

$$\begin{aligned}
\mu_\mathcal{X}(k) =&\mathbb{E}[\mathcal{A}\mathcal{X}(k-1)+\mathcal{U}+\mathcal{F}(k)] = \mathcal{A}\mu_\mathcal{X}(k-1)+\mathcal{U},\\
\Psi_\mathcal{X}(k) =&\mathbb{E}\left[\left(\mathcal{A}\mathcal{X}(k-1)+\mathcal{U}+\mathcal{F}(k)\right)\left(\mathcal{A}\mathcal{X}(k-1)+\mathcal{U}+\mathcal{F}(k)\right)^T\right]\\
=&\mathcal{A}\Psi_\mathcal{X}(k-1)\mathcal{A}^T+\mathcal{Q}(k-1)+\mu_\mathcal{X}(k-1)\mathcal{U}^T+\mathcal{U}\mu_\mathcal{X}^T(k-1).
\end{aligned} \tag{3.95}$$

Finally, it is possible to state the second main result of the work that readily follows from theorems 5 and 3, and Definition 3.1.1.

Proposition 2. *Kalman filter applied to the extended state model* (3.54) *constitutes the* ν*-PVMKF for the 2-DTP.*

3.1.5 Simulation Results

In order to analyze the performance of the new approach the results of a few simulation are provided in the following.

The considered scenario is the typical maneuvering target setting [56]. All simulations were realized by checking the behaviour of the versions of the new filters LVMKF and 2-PVMKF compared to the classical EKF, UKF, and standard PF. The target trajectory

(ground truth) is a sample path of a Zero-Mean First-Order Markov Process [75] characterized by two parameters indicated by τ and σ_a. Parameter τ is the maneuver time constant that depends on how long the target maneuver lasts. Parameter σ_a is the instantaneous standard deviation of the acceleration. For a given target trajectory, a set of 100 realizations of Gaussian measurement noises with standard deviations indicated with σ_ρ and σ_θ were generated with a sampling period $\Delta = 1s$, according to the model (3.10)-(3.11). Moreover, the angle error standard deviation σ_θ was considered within a set of 5 different values.

The evaluation metric of interest is the standard relative position error ($RPE_i(k)$) which is defined for each sample measurement noise realization i at time k as

$$RPE_i(k) = \sqrt{\left(\left(\hat{e}_1^{(i)}(k)\right)^2 + \left(\hat{e}_2^{(i)}(k)\right)^2\right)\left((p_1(k))^2 + (p_2(k))^2\right)^{-1}} \cdot 100\%, \qquad (3.96)$$

where $\hat{e}_l^{(i)}(k) = p_l(k) - \hat{p}_l^{(i)}(k)$, $l = 1, 2$, with $\left(\hat{p}_1^{(i)}(k), \hat{p}_2^{(i)}(k)\right)$ being the estimated positions at time k. The average value of $RPE_i(k)$ with respect to the given set of samples will be denoted by $RPE(k)$. The average value of $RPE(k)$ with respect to time will be indicated with RPE. Moreover, a run is considered to be convergent only if the $RPE_i(k) < 15\%$ at the end of the simulation time.

It is well known that the filtering performances strictly depend on the accuracy of the state model. In the particular case of target tracking, this represents a crucial point since the dynamical features of all possible targets are in general unknown and hardly identifiable. The most common state models used in literature are listed below, assuming matrices A and F in (3.7) in the forms

$$A = \begin{bmatrix} \bar{A} & 0 \\ 0 & \bar{A} \end{bmatrix}, \qquad F = \begin{bmatrix} \bar{F} & 0 \\ 0 & \bar{F} \end{bmatrix}. \qquad (3.97)$$

1. Zero-Mean First-Order Markov model (MM) [75]:

$$\bar{A}_{MM} = \begin{bmatrix} 1 & \Delta & \tau\left(\Delta - \tau + \tau e^{-\Delta/\tau}\right) \\ 0 & 1 & \tau\left(1 - e^{-\Delta/\tau}\right) \\ 0 & 0 & e^{-\Delta/\tau} \end{bmatrix}, \qquad (3.98)$$

   $\bar{F}_{MM}$ as given in [75];

2. Wiener Sequence acceleration model (WS) [56]:

$$\bar{A}_{WS} = \begin{bmatrix} 1 & \Delta & \Delta^2/2 \\ 0 & 1 & \Delta \\ 0 & 0 & 1 \end{bmatrix}, \quad \bar{F}_{WS} = \sigma_a \begin{bmatrix} \Delta^2/2 \\ \Delta \\ 1 \end{bmatrix}; \qquad (3.99)$$

3. White-Noise Jerk model (WJ) [56]:

$$\bar{A}_{WJ} = \bar{A}_{WS}, \quad \bar{F}_{WS} = \sigma_a\sqrt{2/\tau} \begin{bmatrix} \Delta^5/20 & \Delta^4/8 & \Delta^3/6 \\ \Delta^4/8 & \Delta^3/3 & \Delta^2/2 \\ \Delta^3/6 & \Delta^2/2 & \Delta \end{bmatrix}^{\frac{1}{2}}; \qquad (3.100)$$

Algorithms	σ_θ (rad)					ACT
	0.87e-2	1.31e-2	1.75e-2	2.18e-2	2.62e-2	(sec)
LVMKF	0.34	0.46	0.63	0.69	0.89	0.21
2-PVMKF	0.34	0.46	0.62	0.67	0.88	1.97
UKF	0.33	0.45	0.59	0.67	0.88	0.04
EKF	0.33	0.49	0.92	1.96	4.02	0.02
PF	0.52	0.78	0.90	2.15	2.44	11.76

Table 3.1: *Simulation 1: RPE and Average Computational Time (ACT).*

4. Constant Velocity model (CV) [56]:

$$\bar{A}_{CV} = \begin{bmatrix} 1 & \Delta & \Delta^2/2 \\ 0 & 1 & \Delta \\ 0 & 0 & 0 \end{bmatrix}, \quad \bar{F}_{CV} \begin{bmatrix} 0 \\ 0 \\ \sigma_a \end{bmatrix}. \tag{3.101}$$

For all models $\{a(k)\}$ is assumed to be a standard white Gaussian random sequence and, as a consequence, the state noise covariance matrix results to be equal to $\Psi_a = FF^T$. Notice that model MM is the discrete-time counter-part of the model used to generate the ground truth.

An usual strategy adopted to avoid the filtering divergences consists in increasing the amount of system state noise [27], that corresponds to use a suitable scaled covariance matrix $\tilde{\Psi}_a = \alpha\Psi_a$, where $\alpha > 1$ is a tuning parameter . However, in order to get a consistent comparison, in all simulations, no tuning technique was applied for each of the mentioned filters (i.e. $\tilde{\Psi}_a = \Psi_a$) with the exception of the PF, which required such a tuning procedure in order to obtain acceptable results in some of the considered settings.

Regarding the measurement noises, EKF, UKF and PF were implemented by using the right covariance matrix $\Psi_n = diag([\sigma_\rho^2\ \sigma_\theta^2])$. In the cases of the new algorithms, the required parameters in (3.18) and (3.15) for the LVMKF and in (3.42) and (3.45) for the 2-PVMKF were computed basing on the Gaussian assumption, as it follows:

$$\psi_\rho^{(i)} = \begin{cases} (i-1)!!\sigma_\rho^i & i\ odd \\ 0 & i\ even \end{cases}, \tag{3.102}$$

$$\bar{R} = e^{-\sigma_\theta^2/2} I_2, \qquad \mathbb{E}\left[e^{jl_i(q)n_\theta(k)}\right] = e^{-(l_i(q)\sigma_\theta)^2/2}, \qquad i = 1,2,3,4. \tag{3.103}$$

The described simulation were computed for several different trajectories (i.e. different realizations of the state noise), obtaining strictly similar results. The implementations of the filtering procedures were realized on a MATLAB© platform and computed with an Intel(R) Core(TM) i7 CPU 2.70 GHz, 8GB RAM.

Simulation 1: accuracy and computational time

Table 3.1 shows the numerical results in terms of RPE, together with the average computational time (ACT) per measurement noise realization, of a typical simulation whose param-

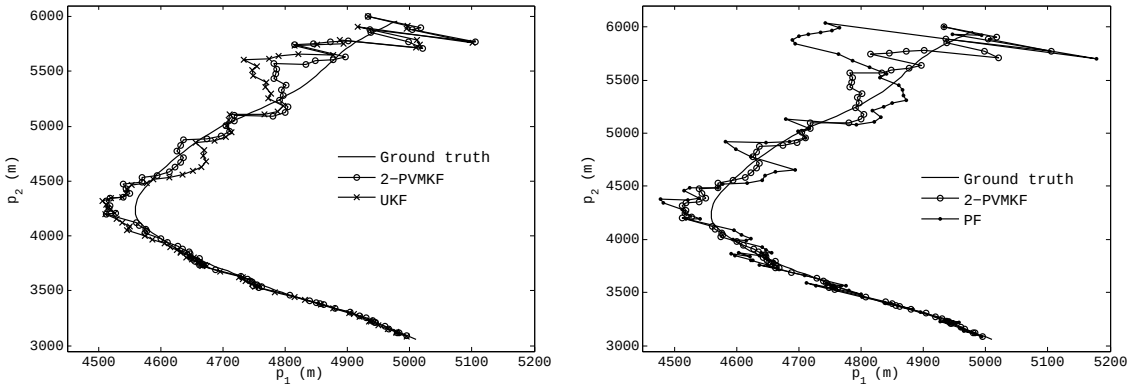

Figure 3.1: *Graphical comparison of 2-PVMK with UKF (left) and PF (right) [$\sigma_\theta = 1.75e-2rad$, $\sigma_\rho = 0.35m$]*

$\sigma_\theta\ (rad)$	0.87e-2	1.31e-2	1.75e-2	2.18e-2	2.62e-2
LVMKF, 2-PVMKF	0	0	0	0	0
UKF	0	0	0	0	0
EKF	0	0	0	0	5
PF	13	27	46	49	76

Table 3.2: *Simulation 1: failure percentage.*

eters are $\tau = 10s$, $\sigma_a = 0.32$, and $\sigma_\rho = 0.35m$, with initial position $(5km, 6km)$, initial velocity $(-10m/s, -20m/s)$, zero initial acceleration and simulation time $Ts = 100s$. Results are averaged over the set of runs resulting convergent for all the considered algorithms. Fig. 3.1 graphically represents an example of the estimation results for one of the measurement realizations. All filters were implemented by using the MM reference model. The PF algorithm was implemented in the generic form, with Gaussian importance density and systematic low-variance resampling [3]. The resampling threshold was set to the usual value $N_{th} = 2N_p/3$ [38], where N_p is the number of particles. In the presented results the tuning parameter was set to $\alpha = 10$ and the number of particle, chosen as a reference value, is $N_p = 3000$.

It is evident that LVMKF, 2-PVMKF and UKF have a similar behaviour and the inadequacy of the standard EKF, especially for significant values of angle errors. It is here important to focus on the computational complexity. For n_x denoting the dimension of the state vector, it is clear that LVMKF is $\mathcal{O}(n_x^3)$, as well as KF and UKF. Otherwise, ν-PVMKF has complexity $\mathcal{O}(n_x^{3\nu})$ that increases with the approximation order ν. On the other hand, from theory [11, 12] follows that a larger order may return more accurate estimation performances. This property is shared with the PF which is $\mathcal{O}(N_p n_x^2)$ and is proved to improve the estimation results with a larger number of particles. However, as showed in Table 3.1, the execution time of the PF in this reference case ($N_p = 3000$) is significantly higher than those of the other filters, although the RPE values are similar to those of the

EKF. Moreover, Table 3.2 reports the percentage of not convergent runs for the considered algorithms. It can be noted that the PF failures generally increase with the measurement error. In contrast, both UKF and the new filters have no failures.

Simulation 2: robustness analysis

In order to evaluate the robustness of the new filtering approach, two different checking methods were applied. The first one consists in simulating all filtering process using each one of the state models (3.98)-(3.101). In the second one, the performances of the considered filtering procedures were checked by varying the model parameters from the correct ones, using the most accurate reference model MM.

Table 3.3 shows the numerical results in terms of RPE of a typical simulation whose parameters are $\tau = 10s$, $\sigma_a = 0.32$, and $\sigma_\rho = 0.35m$, with initial position $(12km, 8km)$, initial velocity $(-10m/s, -20m/s)$, zero initial acceleration and simulation time $Ts = 100s$. In this case, results are averaged over all the 100 runs and PF was omitted tacking into account the previous discussion. It is clear that LVMKF, 2-PVMKF and UKF have similar behaviour for low values of angular noise variance with reference to models MM, WS and WJ, while, with the increasing of the noise angle variance both LVMKF and 2-PVMKF present growing performances with respect to UKF. As far as the CV model is concerned, the performances of the new filters are definitely superior in all conditions. Also in these simulations, the EKF behaviour is decidedly worse than that of the other filters. This is confirmed by Table 3.4 where the percentage of not convergent runs are presented, tacking into account that both UKF and the new filters have no failures.

In the second checking method, parameter σ_a is assumed to be a random variable endowed with a log-uniform distribution, having the real standard deviation σ_a^* as mean value. The above described simulations were repeated on the same data by setting parameter σ_a at each of the values in the set $\{\sigma_a^* 2^{i/2} : i = -8, -7, \ldots, 7, 8\}$. Finally, the robustness index is defined as the average of the RPEs so obtained. Table **??** reports such an index referred to the scenario of simulation 2 showing the high quality in terms of state estimation and robustness of the proposed algorithms.

Simulation 3: comparison of LVMKF and 2-PVMKF

In all the reported simulation results, 2-PVMKF has better estimation performances with respect to the LVMKF, as it is expected from the theoretical argumentation. However, in most cases, they exhibit very similar RPE values. In order to stress the effective difference between the two algorithms a nongaussian setting was considered. In particular, the above described scenario was reproduced by generating uniform distributed measurement noises instead of the conventional Gaussian ones. Recalling that σ_ρ and σ_θ denote the standard deviations of such distributions, the radius and angle error sequences were uniformly generated in the intervals $[-\sqrt{3}\sigma_\rho, \sqrt{3}\sigma_\rho]$ and $[-\sqrt{3}\sigma_\theta, \sqrt{3}\sigma_\theta]$, respectively. Moreover, parameters in (3.102) and (3.103) were substituted by

$$\psi_\rho^{(1)} = \psi_\rho^{(3)} = 0, \quad \psi_\rho^{(2)} = \sigma_\rho^2, \quad \psi_\rho^{(4)} = \frac{9}{5}\sigma_\rho^4, \tag{3.104}$$

σ_θ (rad)	0.87e-2	1.31e-2	1.75e-2	2.18e-2	2.62e-2
			MM		
LVMKF	0.30	0.43	0.56	0.67	0.78
2-PVMKF	0.30	0.42	0.55	0.66	0.77
UKF	0.28	0.39	0.54	0.63	0.82
EKF	0.33	1.16	3.09	6.72	13.97
			WS		
LVMKF	0.34	0.49	0.64	0.78	0.90
2-PVMKF	0.33	0.48	0.63	0.77	0.90
UKF	0.32	0.49	0.67	0.81	1.00
EKF	2.77	11.23	21.42	33.60	48.60
			WJ		
LVMKF	0.33	0.47	0.61	0.74	0.87
2-PVMKF	0.32	0.47	0.60	0.73	0.86
UKF	0.33	0.52	0.81	1.11	1.45
EKF	1.61	8.60	17.30	29.70	45.47
			CV		
LVMKF	0.51	0.68	0.79	0.90	0.99
2-PVMKF	0.50	0.67	0.77	0.89	0.96
UKF	0.60	0.86	1.30	1.74	2.56
EKF	0.66	1.25	2.12	4.66	10.61

Table 3.3: *Simulation 2: numerical RPE results*

$$\bar{R} = \frac{\sin(\sqrt{3}\sigma_\theta)}{\sqrt{3}\sigma_\theta} I_2 \qquad \mathbb{E}\left[e^{jl_i(q)n_\theta(k)}\right] = \frac{\sin(l_i(q)\sqrt{3}\sigma_\theta)}{l_i(q)\sqrt{3}\sigma_\theta}, \qquad i = 1,2,3,4. \tag{3.105}$$

Table 3.6 reports the numerical results in terms of RPE of a simulation whose parameters are $\tau = 10s$, $\sigma_a = 0.32$, and $\sigma_\rho = 3.5m$, with initial position $(4km, 4km)$, initial velocity $(-5m/s, -10m/s)$, zero initial acceleration and simulation time $Ts = 100s$. In this case the only new algorithms are considered. Since no failure occurred for both algorithms, RPEs are averaged over all the 100 runs. As one can note, in this nongaussian setting the polynomial algorithm clearly overperforms the linear one especially when filtering uses less accurate state models such as WJ and CV. Fig. 3.2 graphically represents an example of the estimation results for one of the measurement realizations where it can observed the improved performance of the 2-PVMKF algorithm with respect to the LVMKF.

3.1.6 Conclusions

In this paper, the 2-D tracking problem is considered exploiting the idea of "virtual measurement map". This consists in transforming the nonlinear measurement output function into a linear time-varying one corrupted by an additive nongaussian noise. The resulting model is amenable to the application of Kalman filter as linear least square estimator with respect to the defined virtual output map, without any linearization. Moreover, in order

σ_θ (rad)	0.87e-2	1.31e-2	1.75e-2	2.18e-2	2.62e-2
MM	0	0	1	1	19
WS	7	24	49	67	82
WJ	3	24	42	67	76
CV	0	0	0	1	12

Table 3.4: *Simulation 2: EKF failure percentage.*

σ_θ (rad)	0.87e-2	1.31e-2	1.75e-2	2.18e-2	2.62e-2
LVMKF	0.35	0.48	0.62	0.74	0.85
2-PVMKF	0.35	0.48	0.62	0.74	0.84
UKF	0.37	0.61	1.00	1.30	1.74
EKF	0.59	2.01	5.18	11.10	22.48

Table 3.5: *Simulation 2: Robustness analysis*

to manage the nongaussian noise, a polynomial extension is designed, which could improve the performance of the linear filter. The comparison with the state of the art tracking algorithms (EKF, UKF and particle filter) confirms the effectiveness of this proposal. In particular the present algorithms result to be more robust with respect to model inaccuracies, especially for high measurement noise levels.

σ_θ (rad)	0.87e-2	1.31e-2	1.75e-2	2.18e-2	2.62e-2
			MM		
LVMKF	0.39	0.57	0.72	0.81	0.94
2-PVMKF	0.35	0.51	0.65	0.75	0.88
			WS		
LVMKF	0.37	0.54	0.64	0.76	0.86
2-PVMKF	0.35	0.49	0.59	0.72	0.83
			WJ		
LVMKF	0.36	0.51	0.63	0.72	0.83
2-PVMKF	0.33	0.47	0.57	0.67	0.79
			CV		
LVMKF	1.85	2.32	2.70	2.86	3.14
2-PVMKF	1.60	2.06	2.46	2.68	2.99

Table 3.6: *Simulation 3: numerical RPE results for nongaussian noise*

3.2 Perspective Vision

3.2.1 Introduction

The estimation of the time-varying distance of an object from a camera along its optical axis is a classical problem in machine vision [33]. It has received a lot of attention because of its importance in several practical applications, such as autonomous vehicle navigation, aerial tracking, path planning, surveillance, etc. All these applications usually require to have at disposition the 3-D Euclidean coordinates of moving features or the position of a static object to be recovered from a 2-D image sequence provided by a charge-coupled device (CCD) camera. The overall class of practical problems can be generally represented by considering the relative motion between a perspective camera and an observed object.

In literature, there are many range identification techniques for perspective vision systems. Some of them utilize the extended Kalman filter (EKF) ([17, 50, 60, 76]). However, EKF involves linearization of the nonlinear vision model and requires *a priori* knowledge of the noise distribution. In order to overcome the shortcomings of the linear model, nonlinear system analysis and estimation tools are used to develop nonlinear observers able to identify the range when the motion parameters are known. Such an approach is introduced in [41]. The basic idea is developed in [61] where the perspective problem is treated as a particular case of implicit output system. Following works [15]and [23] suggest to build an observer by expressing the perspective system in terms of the nonlinear feature dynamic. A different approach is adopted in [1] and [20] where the estimation is carried out using methods from linear control theory. A reduced-order observer is suggested in [51] as an application of the more general observer technique introduced in [52]. Finally, interesting recent solutions are proposed in [57, 62] and [21].

In this work a new approach for the estimation of the 3-D position of a moving target is suggested. The main idea is the use of the measurements process as time-varying param-

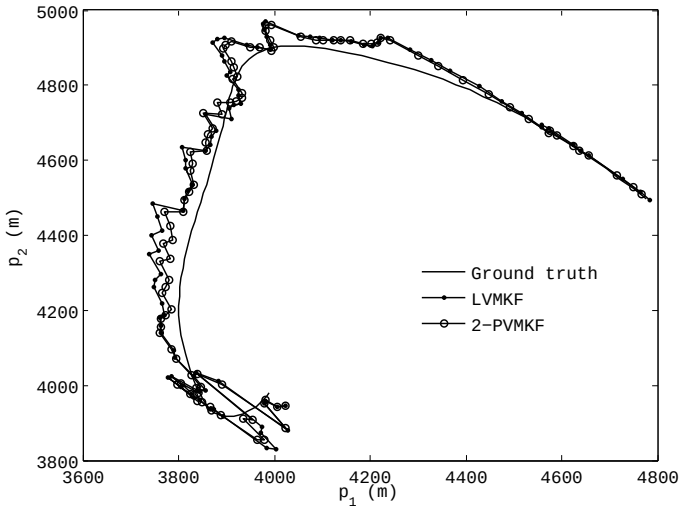

Figure 3.2: *Graphical comparison of LVMKF with 2-PVMK* $[\sigma_\theta = 2.18e-2rad\,,\ \sigma_\rho = 3.5m]$

eters affecting a suitable defined output matrix, derived by the introduction of the virtual output presented at the beginning of this chapter. This simple manipulation transforms the nonlinear stationary measurement map into a linear time-varying one. This allows to solve the problem avoiding any linearization procedure, via a linear time-varying observer whose gain matrix is computed through a delay Riccati equation (DRE).

3.2.2 Problem Formulation

The motion of feature point on a rigid object relative to a calibrated pinhole camera can be described by the affine system [15, 23, 51, 79]

$$\dot{x}(t) = A(t)x(t) + b(t), \tag{3.106}$$

where $A(t) \in \mathbb{R}^{3\times3}$, $b(t) \in \mathbb{R}^3$, and

$$x(t) = [x_1(t), x_2(t), x_3(t)]^T \ \in \ \mathbb{R}^3$$

contains the unmeasurable coordinates of the feature point in an inertial reference frame with x_3 being perpendicular with the camera image plane.

The dynamic matrix $A(t)$ and the drift term $b(t)$ are composed by motion parameters $a_{i,j}(t)$ and $b_i(t)$ which are possible time-varying and are assumed known. The measurable image-space coordinates, denoted by $y(t) \in \mathbb{R}^2$, are given as

$$y(t) = \left[\ y_1(t)\ y_2(t)\ \right]^T = \eta \left[\ \frac{x_1(t)}{x_3(t)} \quad \frac{x_2(t)}{x_3(t)}\ \right]^T, \tag{3.107}$$

where η is the focal length of the camera. Without loss of generality, it can be assumed that $\eta = 1$. For the perspective system in (3.106) and (3.107), following assumptions are made [23].

Assumption 1. *The motion parameters $a_{i,j}(t)$ and $b_i(t)$, $i, j = 1, 2, 3$ are bounded functions of time, i.e. $a_{i,j},\ b_i \in \mathcal{L}_\infty$.*

Assumption 2. *The image-space coordinates $y_1(t)$ and $y_2(t)$ are bounded functions of time, i.e. $y_1,\ y_2 \in \mathcal{L}_\infty$.*

Assumption 3. *The object feature motion avoids the degenerate case where the point feature intersect the image plane. That is $x_3(t) > \varepsilon_0$, where $\varepsilon_0 \in \mathbb{R}$ is an arbitrarily small positive constant. Moreover, $x_3 \in \mathcal{L}_\infty$.*

Remark 1. *Assumptions 2 and 3 are standard hypothesis (see also [15, 23, 51, 79]) that are practically properties of the physical system rather than assumptions.*

The problem objective is the reconstruction of the coordinates $x_1(t),\ x_2(t),\ x_3(t)$ from the measurements of the image-space coordinates $y_1(t)$ and $y_2(t)$.

3.2.3 The New Observer

The Virtual Measurement Map

Taking into account the first definition in (3.107), it follows that, for any t

$$x_1(t) - y_1(t)x_3(t) = 0, \tag{3.108}$$
$$x_2(t) - y_2(t)x_3(t) = 0. \tag{3.109}$$

It can be now defined the virtual output function which is identically equal to a zero vector, i.e.

$$y_v(t) := \begin{bmatrix} 0 & 0 \end{bmatrix}^T, \tag{3.110}$$

for which the following measurement map holds true:

$$y_v(t) = C(t)x(t), \tag{3.111}$$

with

$$C(t) = \begin{bmatrix} I_2 & -y(t) \end{bmatrix}, \tag{3.112}$$

where I_n denotes the identity matrix in $\mathbb{R}^n$. The expression of the time-varying output matrix $C(t)$ in (3.112) directly follows from (3.108) and (3.109).

It is remarkable that the above manipulation transforms the nonlinear stationary measurement map (3.107) into a linear time-varying one (3.111). It can be said that *the definition of a virtual output converts nonlinearity into non-stationarity*. This naturally leads to a simpler and efficient solution to the proposed problem since no approximation (e.g. linearization) has to be performed.

The Observer Equation

The reconstruction of coordinates $x(t)$ from the measurements of $y(t)$ can be obtained by the design of a linear time-varying observer for the affine system in (3.106) and (3.111). The standard approach to linear state observation, motivated partly on grounds of hindsight,

is to generate an asymptotic estimate of the state by using another linear state equation that accepts as inputs the output $y(t)$ and input $b(t)$ signals. For the special case of this paper such a system state equation has the standard structure of a Luemberger time-varying observer [72]:

$$\begin{aligned}\dot{\hat{x}}(t) =& A(t)\hat{x}(t) + b(t) + K^*(t)\left(y_v(t) - C(t)\hat{x}(t)\right),\\ \hat{x}(t_0) =& \hat{x}_0,\end{aligned} \tag{3.113}$$

where $\hat{x}(t) \in \mathbb{R}^3$ is the estimated state, i.e. the estimated coordinates of the feature point, $\hat{x}_0 \in \mathbb{R}^3$ is the initial estimate, and $K^*(t) \in \mathbb{R}^{3\times 2}$ is the *observer gain matrix*. This matrix has the standard form

$$K^*(t) = P(t)C^T(t), \tag{3.114}$$

where $P(t) \in \mathbb{R}^{3\times 3}$ is a symmetric matrix that has been chosen to satisfy the following matrix dynamical delay system which has a form that resembles the classical Riccati equation:

$$\begin{aligned}\dot{P}(t) =\ & A(t)P(t) + P(t)A^T(t) + 4\alpha^2 P(t)\\ & - 2P(t)C^T(t)C(t)P(t)\\ & + 2e^{-4\alpha^2\delta}P(t)\Psi^T(t)C^T(t-\delta)C(t-\delta)\Psi(t)P(t)\end{aligned} \tag{3.115}$$

with $P(t_0) > 0$ and where $\alpha, \delta \in \mathbb{R}$, $\delta > 0$, and $\Psi(t) \in \mathbb{R}^{3\times 3}$ is the solution of

$$\begin{aligned}\dot{\Psi}(t) &= A(t-\delta)\Psi(t) - \Psi(t)A(t),\\ \Psi(t_0) &= \Phi(t_0 - \delta, t_0),\end{aligned} \tag{3.116}$$

$\Phi(t,\tau)$ denoting the state-transition matrix associated to $A(t)$. The role of parameters α and δ will be clarified in the sequel.

Taking into account (3.110) and (3.114), (3.113) can be rewritten as follows:

$$\dot{\hat{x}}(t) = \left(A(t) - P(t)C^T(t)C(t)\right)\hat{x}(t) + b(t). \tag{3.117}$$

Note that the delay Riccati equation (DRE) (3.115) is forced though the output matrix $C(\cdot)$ both by the actual value of measurements $y(t)$ and their delayed values $y(t-\delta)$. It is now known that using past values of the output signal helps observability [10].

In order to clarify the way the observer works, the proposed solution is summarized in Algorithm 1. Note that the values of measurements $y(t)$ before the initial time t_0 are zero valued. This is obviously one of the possible choices for real frameworks. A second possibility is to start the estimation at the delayed time $t_0 + \delta$ and use the output signal $y(t)$ which is actually available within the interval $[t_0, t_0 + \delta]$.

Proof of Convergence

Once introduced the observer equation, the relevant convergence is now proved. In order to quantify the performance of any observer, an estimation error, denoted by $e(t) \in \mathbb{R}^3$, is usually defined as

$$e(t) := x(t) - \hat{x}(t).$$

Algorithm 1 DRE Observer

given: $A|_{[t_0-\delta,T]}, b|_{[t_0,T]}, y|_{[t_0,T]}, \alpha, \delta > 0$;

initial conditions: $\hat{x}(t_0) = \hat{x}_0,\ P(t_0) > 0,$

$$\Psi(t_0) = \Phi(t_0 - \delta, t_0),\ y|_{[t_0-\delta,t_0)} = 0;$$

instant time $t = t_0$, **start integrating:**

$$\begin{aligned}
&\dot{\hat{x}}(t) = \left(A(t) - P(t)C^T(t)C(t)\right)\hat{x}(t) + b(t);\\
&\dot{P}(t) = A(t)P(t) + P(t)A^T(t) + 4\alpha^2 P(t)\\
&\quad - 2P(t)C^T(t)C(t)P(t)\\
&\quad + 2e^{-4\alpha^2\delta}P(t)\Psi^T(t)C^T(t-\delta)C(t-\delta)\Psi(t)P(t);\\
&\dot{\Psi}(t) = A(t-\delta)\Psi(t) - \Psi(t)A(t);\\
&C(t) = \begin{bmatrix} I_2 & -y(t) \end{bmatrix}.
\end{aligned}$$

For observer forms such as (3.113) with a generic gain $K(t)$, the error time derivative is given by

$$\dot{e}(t) = \left(A(t) - K(t)C(t)\right)e(t), \tag{3.118}$$

where $K(t)$ is the observer gain matrix.

The goal of any observer design is to make the observer-error system (3.118) asymptotically stable. In the sequel a sufficient condition for the uniform exponential stability of such a system is given in the particular case of observer (3.117).

In order to analyze the asymptotic behaviour of the observer-error system, a key result form [72] is recalled.

Let the reconstructibility Gramian for the plant in (3.106) and (3.111) be

$$N(t_0, t) = \int_{t_0}^{t} \Phi^T(\tau, t)C^T(\tau)C(\tau)\Phi(\tau, t)d\tau.$$

Note that $N(t_0, t)$ is a symmetric and positive semi-definite matrix. Furthermore, let

$$M(t_0, t) = \int_{t_0}^{t} 2e^{-4\alpha^2(\tau-t)}\Phi^T(\tau, t_0)C^T(\tau)C(\tau)\Phi(\tau, t_0)d\tau.$$

Theorem 4. *Suppose for the linear state equations (3.106) and (3.111) there exist positive constants δ, ε_1, and ε_2 such that*

$$\varepsilon_1 I_3 \leq N(t-\delta, t) \leq \varepsilon_2 I_3 \tag{3.119}$$

for all t. Then, given a constant $\alpha \in \mathbb{R}$, the observer gain matrix

$$K(t) = \left[\Phi^T(t-\delta, t)M(t-\delta, t)\Phi(t-\delta, t)\right]^{-1}C^T(t) \tag{3.120}$$

is such that the resulting observer-error state equation (3.118) is uniformly exponentially stable with rate α^2.

For the proof the reader is referred to [72].

Now, the main result of the paper is summarized in the following theorem.

Theorem 5. *Suppose for the linear state equations (3.106) and (3.111) there exist positive constants* δ *and* ε_1 *such that*

$$N(t-\delta,t) \geq \varepsilon_1 I_3 \tag{3.121}$$

for all t. *Then, given a constant* $\alpha \in \mathbb{R}$, *the observer gain matrix* $K^*(t)$ *defined in (3.114), and satisfying (3.115) and (3.116), is such that the resulting observer-error state equation (3.118) is uniformly exponentially stable with rate* α^2.

Proof. This proof consists of two parts. In the first part it will be proved that condition (3.121) implies condition (3.119) of Theorem 4. Then it will be demonstrate that the special form of the observer gain $K^*(t)$ (3.114) is identically equal to the generic form given by Theorem 4 in (3.120).

Condition (3.121) corresponds to the left-side inequality of condition (3.119). Since $N(t-\delta,t)$ is a symmetric matrix the following inequality holds true (Rayleigh-Ritz inequality):

$$N(t-\delta,t) \leq \lambda_{max}(N(t-\delta,t))I_3,$$

for every t and δ, and where $\lambda_{max}(M)$ is the maximum eigenvalue of a given symmetric matrix M. As a consequence, the right-hand inequality of condition (3.119) is satisfied if there exists a positive finite constant ε_2 such that $\lambda_{max}(N(t-\delta,t)) \leq \varepsilon_2$. Recall that for a symmetric positive semi-definite matrix $Q \in \mathbb{R}^{n\times n}$ the following norm can be defined:

$$\parallel Q \parallel = \lambda_{max}(Q).$$

Therefore, (3.119) is proved if ε_2 is such that

$$\parallel N(t-\delta,t) \parallel \leq \varepsilon_2. \tag{3.122}$$

In order to demonstrate condition (3.122), note that from Assumptions 1 and 2 it follows that there exist two finite positive constants γ and σ such that, for any t:

$$\parallel A(t) \parallel \leq \gamma \qquad \parallel C^T(t)C(t) \parallel \leq \sigma. \tag{3.123}$$

Taking into account (3.123) and Lemma B.0.10 it can be said that there exist a positive constant β such that, for every t and $\delta > 0$:

$$\begin{aligned}
\|N(t-\delta,t) \| &\leq \int_{t-\delta}^{t} \parallel \Phi^T(\tau,t)C^T(\tau)C(\tau)\Phi(\tau,t) \parallel d\tau \\
&\leq \int_{t-\delta}^{t} \parallel \Phi^T(\tau,t) \parallel\parallel C^T(\tau)C(\tau) \parallel\parallel \Phi(\tau,t) \parallel d\tau \\
&\leq \beta^2\sigma\delta < +\infty,
\end{aligned}$$

which proves condition (3.122) with $\varepsilon_2 = \beta^2\sigma\delta$.

Since (3.119) is satisfied, from Theorem 4 follows that the observer gain matrix in (3.120) guarantees the uniform exponential stability of the observer-error system (3.118). Therefore, it remains to prove that the solution of the DRE (3.115) satisfies, for any t,

$$P^{-1}(t) = \Phi^T(t-\delta,t)M(t-\delta,t)\Phi(t-\delta,t), \tag{3.124}$$

so that $K^*(t) = K(t)$, for every t.

In order to prove (3.124) the time derivative of $M(t-\delta,t)$ is needed. This can be computed utilizing properties in Lemma B.0.11:

$$\begin{aligned}\dot{M}(t-\delta,t) = &-4\alpha^2 M(t-\delta,t)\\ &+2\Phi^T(t,t-\delta)C^T(t)C(t)\Phi(t,t-\delta)\\ &-2e^{-4\alpha^2\delta}C^T(t-\delta)C(t-\delta)\\ &-A^T(t-\delta)M(t-\delta,t)\\ &-M(t-\delta,t)A(t-\delta).\end{aligned} \tag{3.125}$$

The time derivative of both sides of (3.124) can be now obtained taking into account (3.125) and Lemma B.0.11:

$$\begin{aligned}\dot{P}^{-1}(t) = &\dot{\Phi}^T(t-\delta,t)M(t-\delta,t)\Phi(t-\delta,t)\\ &+\Phi^T(t-\delta,t)\dot{M}(t-\delta,t)\Phi(t-\delta,t)\\ &+\Phi^T(t-\delta,t)M(t-\delta,t)\dot{\Phi}(t-\delta,t)\\ =&-A(t)^TP^{-1}(t)-P^{-1}(t)A(t)\\ &-4\alpha^2P^{-1}(t)+2C^T(t)C(t)\\ &-2e^{-4\alpha^2\delta}\Phi^T(t-\delta,t)C^T(t-\delta)C(t-\delta)\Phi(t-\delta,t).\end{aligned} \tag{3.126}$$

By defining

$$\Psi(t) := \Phi(t-\delta,t),$$

which satisfies (3.116), as proved in Lemma B.0.11, and using Lemma B.0.8, the DRE (3.115) readily follows from (3.126). □

3.2.4 Simulation Results

In this section, a detailed simulation study is presented to evaluate the performance of the proposed estimation technique. Since the estimation is carried out through the the delayed measurement $y(t-\delta)$, these are set to zero for $t < \delta$.

Consider the example given in [15] and [21] of the perspective system:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-0.2 & 0.4 & -0.6\\ 0.1 & -0.2 & 0.3\\ 0.3 & -0.4 & 0.4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} + \begin{bmatrix}0.5\\ 0.25\\ 0.3\end{bmatrix},$$

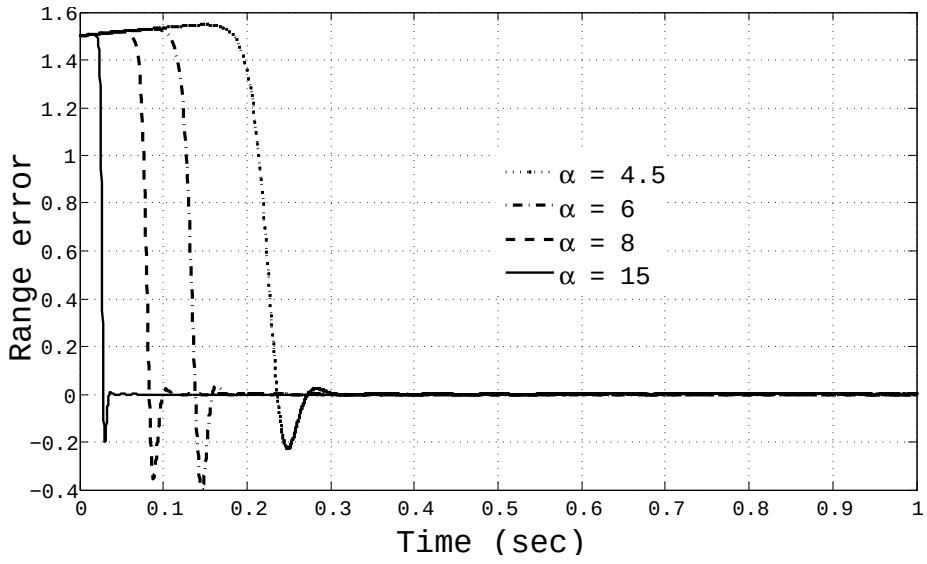

Figure 3.3: *Time history of the observation error for different values of* α.

with initial conditions

$$x(0) = [1\ 1.5\ 2.5]^T,$$
$$\hat{x}_0 = [0.4\ 0.6\ 1]^T.$$

The observer performance for different values of the rating parameter α, namely $\alpha = 4.5,\ 6,\ 8,\ 15$, is shown in Fig. 3.3, with $\delta = 0.01$. Note that the convergence rate can be arbitrarily chosen simply by setting parameter α.

In order to compare this observer with an existing approach, the observer in [51] is considered and the results are shown in Fig. 3.4 with $\alpha = 5$ and the constant design parameter of [51] $\lambda = 30$.

To demonstrate the robustness of the suggested tool, the system in [21] is taken into account:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Three set of initial condition (ICs) have been considered in the format $[x^T(0)\ \hat{x}_0^T]^T$, with $\delta = 0.1$:

$$\begin{aligned} IC1: &\quad [-1, 2, 2, 1/6, 1/3, 1/3]^T, \\ IC2: &\quad [-1, 2, 1, -0.03, 0.12, 0.30]^T, \\ IC3: &\quad [-2, 3, 4, -0.4, 2.4, 0.4]^T. \end{aligned} \tag{3.127}$$

In Fig. 3.5 the norm of error $\| x(t) - \hat{x}(t) \|$ is graphically represented for the different ICs in (3.127).

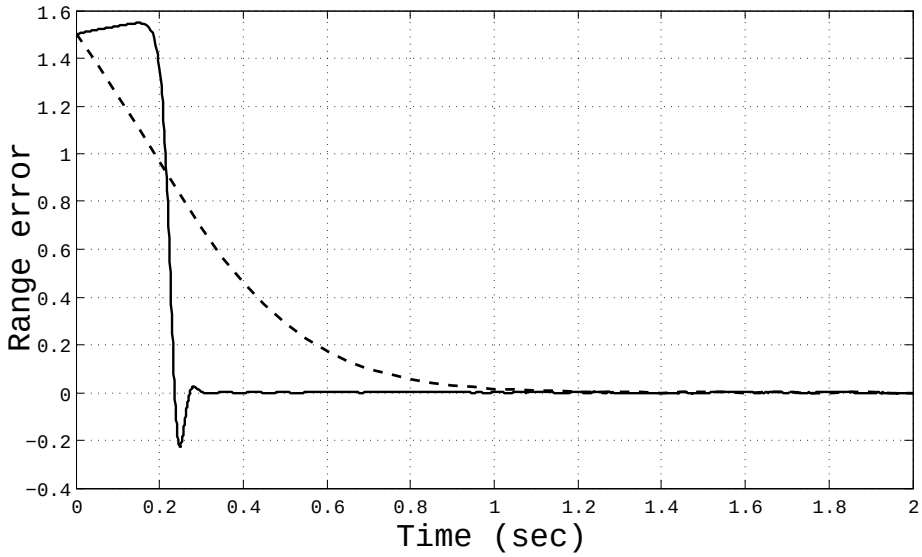

Figure 3.4: *State estimate error in x_3-coordinate. Solid curve is the new observer instead the dashed curve is the observer presented in [51].*

3.3 Conclusions

As a concluding remark, it deserves to point out the novelty of the concept denoted as "virtual output measurement process", that could be very promising for the treatment of important classes of output measurement functions. Moreover the gain matrix of the proposed observer, computed trough a special kind of Riccati equation, guarantees global convergence and high performances of this method.

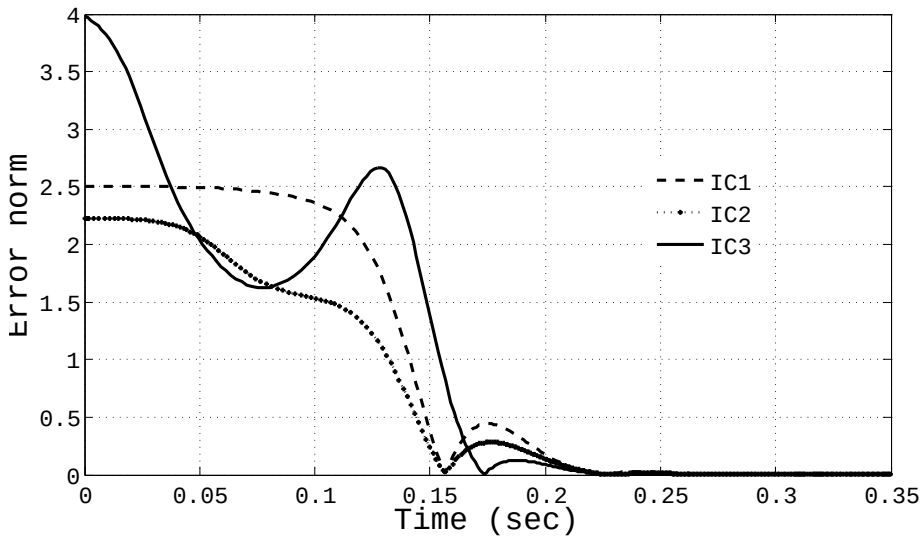

Figure 3.5: *Norm of the state estimate error.*

APPENDIX $\mathcal{A}$

The Kronecker Algebra

Throughout this thesis, Kronecker algebra is widely used. Here, for the sake of completeness, some definitions and properties are recalled and some new results on this subject are given.

Definition A.0.1. *Let M and N be real matrices of dimensions $r \times s$ and $p \times q$ respectively. Then, the Kronecker product $M \otimes N$ is defined as the $(r \cdot p) \times (s \cdot q)$ matrix*

$$M \otimes N = \begin{bmatrix} m_{1,1}N & \cdots & m_{1,s}N \\ \vdots & \ddots & \vdots \\ m_{r,1}N & \cdots & m_{r,s}N \end{bmatrix}, \tag{A.1}$$

where $m_{i,j}$ are the entries of M.

Of course, this kind of product is not commutative.

Definition A.0.2. *Let M be a $r \times s$ real matrix, the Kronecker power of M is defined as*

$$\begin{aligned} M^{[0]} &= 1 \in \mathbb{R}, \\ M^{[l]} &= M \otimes M^{[l-1]} \quad l \geq 1. \end{aligned} \tag{A.2}$$

Definition A.0.3. *Let M be the $r \times s$ real matrix*

$$M = \begin{bmatrix} m_1 & m_2 & \cdots & m_s \end{bmatrix}, \tag{A.3}$$

where m_i denotes the i-th column of M. Then, the stack of M is the $r \cdot s$ vector

$$st(M) = \begin{bmatrix} m_1^T & m_2^T & \cdots & m_s^T \end{bmatrix}^T. \tag{A.4}$$

Moreover, the inverse operation is defined as

$$st^{-1}\left(\begin{bmatrix} m_1^T & m_2^T & \cdots & m_s^T \end{bmatrix}^T\right) = M. \tag{A.5}$$

It can be easily proved that the stack operator is linear. Some useful properties of the Kronecker product and power are given in the following.

Lemma A.0.1. *For any matrices A, B, C, and D, and any vectors u and v, it is:*

$$(A+B)\otimes(C+D) = (A\otimes C)+(A\otimes D)+(B\otimes C)+(B\otimes D), \tag{A.6}$$
$$(A\otimes B)\cdot(C\otimes D) = (A\cdot C)\otimes(B\cdot D), \tag{A.7}$$
$$(A\otimes B)^T = A^T\otimes B^T, \tag{A.8}$$
$$st(A\cdot B\cdot C) = \left(C^T\otimes A\right)\cdot st(B), \tag{A.9}$$
$$st(uv^T) = v\otimes u, \tag{A.10}$$

the $\cdot$ denoting the standard matrix product.

For the proof see [6].

Lemma A.0.2. *For any given pair of matrices $A\in\mathbb{R}^{r\times s}$, $B\in\mathbb{R}^{n\times m}$ it results:*

$$A\otimes B = C_{r,n}^T\,(A\otimes B)\,C_{s,m}, \tag{A.11}$$

where the (h,l) entry of the commutation matrix $C_{u,v}\in\mathbb{R}^{(uv)\times(uv)}$ is given by

$$\{C_{u,v}\}_{h,l} = \begin{cases} 1, & \text{if } l = (|h-1|_\nu)\,u + \left(\left[\frac{h-1}{\nu}+1\right]\right); \\ 0, & \text{otherwise}. \end{cases} \tag{A.12}$$

For the proof see [12].

Lemma A.0.3. *Let $x, y\in\mathbb{R}^n$, for any integer h, it is:*

$$(x+y)^{[h]} = \sum_{k=0}^{h} M_{k,n}^h (x^{[k]}\otimes y^{[h-k]}), \tag{A.13}$$

where $M_{k,n}^h$ satisfies

$$\begin{aligned} M_{h,n}^h &= M_{0,n}^h = I_n^{[h]}, \\ M_{k,n}^h &= (M_{k,n}^{h-1}\otimes I_n) + (M_{k-1,n}^{h-1}\otimes I_n)(I_n^{[k-1]}\otimes G_{h-k}),\ 1\le k\le h, \end{aligned} \tag{A.14}$$

with G_l given by

$$\begin{aligned} G_1 &= C_{n,n}^T, \\ G_l &= (I_n\otimes G_{l-1})(G_1\otimes I_{l-1}), l>1, \end{aligned} \tag{A.15}$$

where $C_{n,n}^T$ is given by (A.12).

For the proof see [12].

Lemma A.0.4. *Let $a_i \in \mathbb{R}^n$, $1 \leq i \leq p$, $p \in \mathbb{N}$. Then:*

$$(a_1 + a_2 + \ldots + a_p)^{[h]} = \sum_{\substack{h_1,\ldots,h_p \geq 0 \\ h_1+\cdots+h_p=h}} M^h_{h_1,\ldots,h_p} \left(a_1^{[h_1]} \otimes a_2^{[h_2]} \otimes \cdots \otimes a_p^{[h_p]} \right), \tag{A.16}$$

where $M^h_{h_1,\ldots,h_p} \in \mathbb{R}^{n^h \times n^h}$ satisfies

$$\begin{aligned} &M^h_{h_1,\ldots,h_p} = 0_{n^h} \text{ if at least one of } h_i \text{ is negative}, \\ &M^h_{h_1,\ldots,h_p} = I_n \text{ for } h = 1, \\ &M^h_{h_1,\ldots,h_p} = \sum_{1 \leq i \leq p-1} \left(M^h_{h_1,\ldots,h_i-1,\ldots,h_p} \otimes I_n \right) \left(I_n^{[h_1,\ldots,h_i-1]} \otimes G_{h_{i+1}\ldots,h_p} \right) \\ &\qquad + \left(M^{h-1}_{h_1,\ldots,h_p-1} \otimes I_n \right) \text{ for } h > 1. \end{aligned} \tag{A.17}$$

with G_l given by (A.15).

For the proof see [12].

Lemma A.0.5. *Let $\mathcal{V} \in \mathbb{R}^r$, $r = n(n^\nu - 1)/(n-1)$, $\nu, n \in \mathbb{N}$, be defined as in (3.36). Then:*

$$\mathcal{V}^{[h]} = \sum_{\substack{h_1,\ldots,h_\nu \geq 0 \\ h_1+\cdots+h_\nu=h}} M^h_{h_1,\ldots,h_\nu} \left[\left(\bigotimes_{i=1}^{\nu} (\Pi^{(i)T}_{\nu,n})^{[h_i]} \right) v^{[h_1+2h_2+\ldots+\nu h_\nu]} \right]. \tag{A.18}$$

Proof. Vector $\mathcal{V}$ can be represented as

$$\mathcal{V} = \sum_{i=1}^{\nu} \Pi^{(i)T}_{\nu,n} v^{[i]}. \tag{A.19}$$

Therefore, from Lemma A.0.4 and iteratively using property (A.7) it follows that

$$\begin{aligned} \mathcal{V}^{[h]} = \left(\sum_{i=1}^{\nu} \Pi^{(i)T}_{\nu,n} v^{[i]} \right)^{[h]} &= \sum_{\substack{h_1,\ldots,h_\nu \geq 0 \\ h_1+\cdots+h_\nu=h}} M^h_{h_1,\ldots,h_\nu} \left[\bigotimes_{i=1}^{\nu} \left(\Pi^{(i)T}_{\nu,n} v^{[i]} \right)^{[h_i]} \right] \\ &= \sum_{\substack{h_1,\ldots,h_\nu \geq 0 \\ h_1+\cdots+h_\nu=h}} M^h_{h_1,\ldots,h_\nu} \left[\left(\bigotimes_{i=1}^{\nu} (\Pi^{(i)T}_{\nu,n})^{[h_i]} \right) v^{[h_1+2h_2+\ldots+\nu h_\nu]} \right]. \end{aligned} \tag{A.20}$$

Lemma A.0.6. *Let $v \in \mathbb{R}^n$. Then, for any $i \geq 0$, the q-th entry if $v^{[i]}$ is*

$$\{v^{[i]}\}_q = v_{l_1} v_{l_2} \cdot \cdots \cdot v_{l_i}, \tag{A.21}$$

where

$$\begin{aligned} l_m &= \left| \left[\frac{q-1}{n^{i-m}} \right] \right|_n + 1, \ m = 1, 2, \ldots, i-1, \\ l_q &= |q-1|_n + 1. \end{aligned} \tag{A.22}$$

For the proof see [11].

Lemma A.0.7. *Let* $x, y \in \mathbb{R}^n$, $A, B \in \mathbb{R}^{n \times n}$, *and*

$$\mathcal{X} = \begin{bmatrix} x^{[1]} \\ \vdots \\ x^{[\nu]} \end{bmatrix}, \qquad \mathcal{Y} = \begin{bmatrix} y^{[1]} \\ \vdots \\ y^{[\nu]} \end{bmatrix}, \qquad [Ax]_e = \begin{bmatrix} A^{[1]}x^{[1]} \\ \cdots \\ A^{[\nu]}x^{[\nu]} \end{bmatrix}. \tag{A.23}$$

For any $h = 1, 2, \ldots, \nu$, $\nu \geq 1$, *it is:*

$$[Ax]_e = \Lambda_\nu(A)\mathcal{X}, \tag{A.24}$$

$$(x+y)^{[h]} = \mathcal{M}_{\nu,n}^{(h)} (\mathcal{X} \otimes \mathcal{Y}) + x^{[h]} + y^{[h]}, \tag{A.25}$$

$$(Ax+By)^{[h]} = \mathcal{M}_{\nu,n}^{(h)} (\Lambda_\nu(A) \otimes \Lambda_\nu(B)) (\mathcal{X} \otimes \mathcal{Y}) + (Ax)^{[h]} + (By)^{[h]}, \tag{A.26}$$

where matrices $\mathcal{M}_{\nu,n}^{(h)} \in \mathbb{R}^{n^h \times r^2}$, $r = n(n^\nu - 1)/(n-1)$ *are such that*

$$\begin{aligned} \mathcal{M}_{\nu,n}^{(1)} &= 0, \\ \mathcal{M}_{\nu,n}^{(h)} &= \sum_{k=1}^{h-1} M_{k,n}^h \left(\Pi_{\nu,n}^{(k)} \otimes \Pi_{\nu,n}^{(h-k)} \right), \ h > 1. \end{aligned} \tag{A.27}$$

Proof. Equation(A.24) is trivially satisfied. For the case $h = 1$, (A.25) holds true. For $h > 1$, by consecutively applying Lemma A.0.3, property (3.39) and property (A.7), it results:

$$\begin{aligned} (x+y)^{[h]} &= \textstyle\sum_{k=1}^{h-1} M_{k,n}^h \left(\left(\Pi_{\nu,n}^{(k)} \mathcal{X} \right) \otimes \left(\Pi_{\nu,n}^{(h-k)} \mathcal{Y} \right) \right) + x^{[h]} + y^{[h]} \\ &= \textstyle\sum_{k=1}^{h-1} M_{k,n}^h \left(\Pi_{\nu,n}^{(k)} \otimes \Pi_{\nu,n}^{(h-k)} \right) (\mathcal{X} \otimes \mathcal{Y}) + x^{[h]} + y^{[h]} \\ &= \mathcal{M}_{\nu,n}^h (\mathcal{X} \otimes \mathcal{Y}) + x^{[h]} + y^{[h]}, \end{aligned} \tag{A.28}$$

which proves (A.25). By applying last property to vectors Ax and By, one has:

$$(x+y)^{[h]} = \mathcal{M}_{\nu,n}^{(h)} ([Ax]_e \otimes [By]_e) + (Ax)^{[h]} + (By)^{[h]}, \tag{A.29}$$

that can be rewritten exploiting properties (A.24) and (A.7) as

$$\begin{aligned} (x+y)^{[h]} &= \mathcal{M}_{\nu,n}^{(h)} ((\Lambda_\nu(A)\mathcal{X}) \otimes (\Lambda_\nu(B)\mathcal{Y})) + (Ax)^{[h]} + (By)^{[h]} \\ &= \mathcal{M}_{\nu,n}^{(h)} (\Lambda_\nu(A) \otimes \Lambda_\nu(B)) (\mathcal{X} \otimes \mathcal{Y}) + (Ax)^{[h]} + (By)^{[h]}, \end{aligned} \tag{A.30}$$

which corresponds to (A.26).

APPENDIX $\mathcal{B}$

Mathematical Tools

Appendix

Lemma B.0.8. *Let $M : \mathbb{R} \to \mathbb{R}^{n\times n}$ be a continuous function. Suppose $M(t)$ to be invertible for all t, then*

$$\dot{M}^{-1}(t) = -M^{-1}(t)\dot{M}(t)M^{-1}(t), \tag{B.1}$$

for every t.

Proof. Since $M(t)$ is invertible for all t, then

$$M(t)M^{-1}(t) = I_n \qquad \forall t.$$

From the time derivative of both sides of last equation, it follows that

$$\dot{M}(t)M^{-1}(t) + M(t)\dot{M}^{-1}(t) = 0$$

from which (B.1) can be trivially obtained. □

Let $\Phi(t,\tau) \in \mathbb{R}^{n\times n}$ be the state-transition matrix associated with the linear system

$$\dot{\zeta}(t) = A(t)\zeta(t)$$

where $\zeta(t) \in \mathbb{R}^n$ and $A(t) \in \mathbb{R}^{n\times n}$. Following Lemma give a few key properties of this matrix.

Lemma B.0.9. *The transition matrix for $A(t)$ is invertible for any $t, \tau \in \mathbb{R}$ and*

$$\Phi^{-1}(t,\tau) = \Phi(\tau,t)$$

Lemma B.0.10. *If there exists a finite positive constant γ such that $\| A(t) \| \leq \gamma$, for all t, then given a finite $\delta > 0$ there exists a finite $\beta > 0$ such that $\| \Phi(t, \tau) \| \leq \beta$ for all $t, \tau \in \mathbb{R}$ such that $|t - \tau| < \delta$.*

For the proof of Lemma B.0.9 and B.0.10 the reader is referred to [72].

Lemma B.0.11. *For every $t, \tau, \delta \in \mathbb{R}$, $\delta > 0$ it is:*

$$\dot{\Phi}(t, \tau) = A(t)\Phi(t, \tau); \tag{B.2}$$

$$\frac{\partial \Phi(t, \tau)}{\partial \tau} = -\Phi(t, \tau)A(\tau); \tag{B.3}$$

$$\dot{\Phi}(t - \delta, t) = A(t - \delta)\Phi(t - \delta, t) - \Phi(t - \delta, t)A(t). \tag{B.4}$$

Proof. Equation (B.2) is a general property of the state-transition matrix.

From Lemma B.0.8 and B.0.9, and (B.2) the partial derivative of the $\Psi(t, \tau)$ with respect to the second argument τ can be computed as follows:

$$\begin{aligned} \frac{\partial \Phi(t, \tau)}{\partial \tau} &= \frac{\partial \Phi^{-1}(\tau, t)}{\partial \tau} \\ &= -\Phi^{-1}(\tau, t)\frac{\partial \Phi(\tau, t)}{\partial \tau}\Phi^{-1}(\tau, t) \\ &= -\Phi^{-1}(\tau, t)A(\tau)\Phi(\tau, t)\Phi^{-1}(\tau, t) \\ &= -\Phi^{-1}(\tau, t)A(\tau) = -\Phi(t, \tau)A(\tau), \end{aligned} \tag{B.5}$$

which proves (B.3).

Finally, (B.4) is trivially derived by applying properties (B.2) and (B.3). □

Bibliography

[1] R. Abdursul, H. Inaba, and B. K. Ghosh. Nonlinear observers for perspective time-invariant linear systems. *Automatica*, 40(3):481–490, 2004.

[2] T Alarcon, HM Byrne, and PK Maini. A multiple scale model for tumor growth. *Multiscale Modeling and Simulation*, 3(2):440, 2005.

[3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Trans. Signal Processing*, 50(2), February 2002.

[4] A.V. Balakrishnan. *Kalman filtering theory*. Optimization Software, Inc., New York, 1984.

[5] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Application to Tracking and Navigation: Theory, Algorithm, and Software*. Wiley, New York, 2001.

[6] R. Bellman. *Introduction to Matrix Analysis*. McGraw-Hill, New York, 1970.

[7] I. Bilink and J. Tabrikian. Maneuvering target tracking in the presence of glint using the nonlinear gaussian mixture kalman filter. *IEEE Transaction on Aerospace and Electronic Systems*, 46(1):246–262, 2010.

[8] W. D. Blair, G. A. Watson, and T. R. Rice. Tracking manuevering targets with an interacting multiple model filter containing exponentially correlated acceleration models. In *Southeastern Symposium on System Theory*, Columbia, SC, March 1991.

[9] Z Bozdech, M Llinas, BL Pulliam, ED Wong, J Zhu, and JL DeRisi. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol*, I:E5, 2003.

[10] F. Cacace, A. Germani, and C. Manes. Observability through delayed measurements: a new approach to state observers design. *International Journal of Control*, 83(11):2395–2410, 2010.

[11] F. Carravetta, A. Germani, and M. Raimondi. Polynomial filtering for linear discrete tiem non-gaussian systems. *SIAM J. Control and Optimization*, 34(5):1666–1690, September 1996.

[12] F. Carravetta, A. Germani, and M. Raimondi. Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise. *IEEE Transaction on Automatic Control*, 42(8):1106–1126, August 1997.

[13] J. A. Castellanos, J. Neira, and J. D. Tardos. Limits to consistency of the EKF-based SLAM. In *Intelligent Autonomous Vehicle (IAV 2004)*, Lisbon, Portugal, July 2004.

[14] P Castorina and D Zappalą. Tumor gompertzian growth by cellular energetic balance. *Phys Stat Mech Appl*, 365(2):473, 2006.

[15] X. Chen and H. Kano. A new state observer for perspective systems. *IEEE Trans. Autom. Control*, 47(4):658–663, April 2002.

[16] D Cheneval, T Kastelic, P Fuerst, and CN Parker. A review of methods to monitor the modulation of mrna stability: a novel approach to drug discovery and therapeutic intervention. *J Biomol Screen*, 15:609–622, 2010.

Bibliography

[17] A. Chiuso, P. Favaro, H. Jin, and S Soatto. Structure from motion casually integrated over time. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):523–535, 2002.

[18] G Ciccarella, M Dalla Mora, and A Germani. A luenberger-like observer for nonlinear systems. *Int J Contr*, 57(3):537, 1993.

[19] G Ciccarella, M Dalla Mora, and A Germani. Observers for discrete-time nonlinear systems. *System & Control Letters*, 20:373, 1993.

[20] O. Dahl, F. Nyberg, J. Olst, and A. Heyden. Linear design of a nonlinear observer for perspective systems. In *IEEE Conference on Robotics and Automation (ICRA05)*, pages 266–268, Barcellona, Spain, 2005.

[21] O. Dahl, Y. Wang, A. F. Lynch, and A. Heyden. Observer forms for persective systems. *Automatica*, 46:1829–1834, 2010.

[22] LG de Pillis, W Gu, KR Fister, T Head, K Maples, A Murugan, T Neal, and K Yoshida. Chemotherapy for tumors: An analysis of the dynamics and a study of quadratic and linear optimal controls. *Math Biosci*, 209(1):292, 2007.

[23] W. Dixon, Y. Fang, D. Dawson, and T. Flynn. Range identification for perspective vision systems. *IEEE Trans. Autom. Control*, 48(12):2232–2238, December 2003.

[24] E Eden, R Navon, I Steinfeld, D Lipson, and Z Yakhini. Gorilla: A tool for discovery and visualization of enriched go terms in ranked gene lists. *BMC Bioinformatics*, pages 10–48, 2009.

[25] R Elkon, E Zlotorynski, KI Zeller, and R Agami. Major role for mrna stability in shaping the kinetics of gene induction. *BMC Genomics*, 11:259, 2010.

[26] L Farina, A De Santis, S Salvucci, G Morelli, and I Ruberti. Embedding mrna stability in correlation analysis of time-series gene expression data. *PLoS Comput Biol*, 4:e1000141, 2008.

[27] P. Fearhead. *Sequential Monte Carlo methods in filter theory*. PhD thesis, University of Oxford, 1998.

[28] J. M. Fitts. Aided tracking as applied to high accuracy pointing systems. *IEEE Transaction on Aerospace and Electronic Systems*, AES-9, May 1973.

[29] P.H. Foo and G.W. Ng. Combining the interacting multiple model method with particle filters for manoeuvering target tracking. *IET Radar, Sonar and Navigation*, 5(3):234–255, 2011.

[30] J Garcia-Martinez, A Aranda, and J Perez-Ortin. Genomic run-on evaluates transcription rates for all yeasts genes and identifies gene regulatory mechanisms. *Mol Cell*, 15:303–313, 2004.

[31] NL Garneau, J Wilusz, and CJ Wilusz. The highways and byways of mrna decay. *Nat Rev Mol Cell Bio*, 8:113–126, 2007.

[32] AP Gerber, D Herschlag, and PO Brown. Extensive association of functionally and cytotopically related mrnas with puf family rna-binding proteins in yeast. *PLoS Biol*, 2:342–354, 2004.

[33] B. K. Ghosh and E. P. Loucks. A perspective theory for motion and shape estimation in machine vision. *SIAM Journal of Control and Optimization*, 33(5):1530–1559, September 1995.

[34] B Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Phil Trans*, 115:513, 1825.

[35] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *Proc. Inst. Elect. Eng. F*, 140(2), 1993.

[36] J Grigull, S Mnaimneh, J Pootoolal, MD Robinson, and TR Hughes. Genome-wide analysis of mrna stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol Cell Biol*, 25:5534–5547, 2004.

[37] F. Guo, Z. Sun, and K. Huangfu. A modified covariance extended Kalman filtering algorithm in passive location. In *Proceedings of IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*, pages 307–311, October 2003.

[38] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P. J. Nordlund. Particle filters for positioning, navigation and tracking. *IEEE Trans. on Signal Processing*, 50(2):425 – 437, 2002.

[39] J Houseley and D Tollervay. The many pathways of rna degradation. *Cell*, 136:763–776, 2009.

[40] N. Ikoma, N. Ichimurat, T. Higuchit, and H. Maeda. Maneuvering target tracking by using particle filter. In *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th*, 2001.

[41] M. Jankovic and B. K. Ghosh. Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear observer. *System & Control Letters*, 25(1):63–73, 1995.

[42] A Jemal, R Siegel, E Ward, Y Hao, and J Xu. Cancer statistics, 2009. *CA: a cancer journal for clinicians*, 59:225, Jan 2009.

[43] L. Jing, H. ChongZhao, and H. Yu. Adaptive MCMC particle filter for tracking maneuvering target. In *IEEE Chinese Control Conference*, 2011.

[44] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defence Sensing, Simulation and Controls*, Orlando, FL, 1997.

[45] S. J. Julier and J. K. Uhlmann. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transaction on Automatic Control*, 45(3):477–482, March 2000.

[46] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. In *Proceedings of the IEEE*, volume 92, March 2004.

[47] T Kailath. Linear systems. *Prentice-Hall*, 29:436–442, 1980.

[48] RE Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82:35–45, 1960.

[49] Eyung W. Kang. *Radar system analysis, design, and simulation*. Artech house, Boston, 2008.

[50] H. Kano, B. K. Ghosh, and H. Kanai. Single camera based motion and shape estimation using extended kalman filtering. *Mathematical and Computer Modelling*, 34(5):511–525, 2001.

[51] D. Karagiannis and A. Astolfi. A new solution to the problem of range identification in perspective vision systems. *IEEE Trans. Autom. Control*, 50(12):2074–2077, December 2005.

[52] D. Karagiannis, D. Carnevale, and A. Astolfi. Invariant manifold based reduced-order observer design for nonlinear systems. *IEEE Trans. Autom. Control*, 53(11), December 2008.

[53] JD Keene. The global dynamics of rna stability orchestrates responses to cellular activation. *BMC Biol*, 8:95, 2010.

[54] Krzysztof Kowalski and Willi-Hans Steeb. *Nonlinear Dynamical Systems and Carleman Linearization*. World Scientific Publishing Company, 1991.

[55] M. Landau. Radar tracking of airborne targets. In *National Aerospace and Electronics Conference (NAECON)*, Dayton, OH, 1976.

[56] X. R. Li and V. P. Jilkov. Survey of manuevering target tracking. Part I: Dynamical models. *IEEE Transaction on Aerospace and Electronic Systems*, 39(4), October 2003.

[57] A. De Luca, G. Oriolo, and P. R. Giordano. Feature depth observation for image-based visual servoing: theory and experiments. *International Journal of Robotics Research*, 27(10):1093–1116, 2008.

[58] S. Mandzuka. Ship tracking control: Optimal estimation of navigation parameters. In *42th International Symposium ELMAR*, Zadra, 2000.

[59] J. L. Maryak, J. C. Spall, and B. D. Heydon. Use of the kalman filter for inference in state-space models with unknown noise distributions. *IEEE Transactions on Automatic Control*, 49:87–90, 2004.

[60] L. Matthies, T. Kanade, and R. Szeliski. Kalman filter-based algorithms for estimating depth from image sequences. *International Journal of Computer Vision*, 3:209–236, 1989.

[61] A. Matveev, X. Hu, R. Frezza, and H. Rehbinder. Observers for systems with implicit output. *IEEE Trans. Autom. Control*, 45(1), 2000.

[62] F. Morbidi and D. Prattichizzo. Range estimation from a moving camera: an immersion and invariance approach. In *IEEE Conference on Robotics and Automation (ICRA09)*, pages 2747–2752, Kobe, Japan, 2009.

[63] M. R. Morelande and S. Challa. Manoeuvring target tracking in clutter using particle filters. *IEEE Trans. on Aerospace and Electronic Systems*, 41:252–270, 2005.

Bibliography

[64] SE Munchel, RK Shultzaberger, N Takizawa, and K Weis. Dynamic profiling of mrna turnover reveals gene-specific and system-wide regulation of mrna decay. *Mol Biol Cell*, 22:2787–2795, 2011.

[65] L Norton. A gompertzian model of human breast cancer growth. *Canc Res*, 48(24 Part1):7067, 1988.

[66] DA Orlando, CY Lin, A Bernard, JY Wang, JES Socolar, ES Iversen, AJ Hartemink, and SB Haase. Global control of cell-cycle transcription by coupled cdk and network oscillators. *Nature*, 453:944–947, 2008.

[67] J. B. Pearson. *Basic Studies in Airborne Radar Tracking Systems*. PhD thesis, University of California at Los Angeles, 1970.

[68] J. B. Pearson and E. B. Stear. Kalman filter applications in airborne radar tracking. *IEEE Transaction on Aerospace and Electronic Systems*, AES-10:319–329, 1974.

[69] T Pramila, W Wu, S Miles, WS Noble, and LL Breeden. The forkhead transcription factor hcm1 regulates chromosome segregation genes and fills the s-phase gap in the transcriptional circuitry of the cell cycle. *Genes Dev*, 20:2266–2278, 2006.

[70] T Roose, S Chapman, and P Maini. Mathematical models of avascular tumor growth. *SIAM Rev*, 49(2):179, 2007.

[71] J Ross. mrna stability in mammalian cells. *Microbiol Rev*, 59:423–450, 1995.

[72] W. J. Rugh. *Linear System Theory*. Prentice Hall, Inc., Upper Saddle River , New Jersey, second edition, 1996.

[73] O Shalem, O Dahan, MR Martinez, I Furman, E Segal, and Y Pilpel. Transient transcriptional responses to stress are generated by opposing effects of mrna production and degradation. *Mol Syst Biol*, 4:223, 2008.

[74] JL Shock, KF Fischer, and JL DeRisi. Whole-genome analysis of mrna decay in *Plasmodium falciparum* reveals a global lengthening of mrna half-life during the intra-erythrocytic developmental cycle. *Genome Biol*, 8:R134, 2007.

[75] R. A. Singer. Estimating optimal tracking filter performance for manned maneuvering targets. *IEEE Transaction on Aerospace and Electronic Systems*, AES-6(4), 1970.

[76] S. Soatto, R. Frezza, and P. Perona. Motion estimation via dynamic vision. *IEEE Trans. Autom. Control*, 41(3):393–413, 1996.

[77] T. L. Song and J. Speyer. A stochastic analysis of a modified gain extended Kalman filter with applications to estimation with bearings only measurement. *IEEE Transaction on Automatic Control*, AC-30(10):940–949, 1985.

[78] J. C. Spall. Estimation via markov chain Monte Carlo. *IEEE Control Systems Magazine*, 23(2):34–45, 2003.

[79] R. Y. Tsai and T. S. Huang. Estimating three-dimensional motion parameters of a rigid planar patch. *IEEE Trans. on Acustics, Speech, and Signal Processing*, ASSP-29(6):1147–152, December 1981.

[80] BP Tu, A Kudlicki, M Rowicka, and SL McNight. Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. *Science*, 310:1152–1158, 2005.

[81] R. van der Merwe, E. A. Wan, and S. J. Julier. Sigma-point Kalman filters for nonlinear estimation and sensor-fusion: Applications to integrated navigation. In *Proceedings of AIAA Guidance Navigation and Controls Conference*, Providence, RI, 2004.

[82] E. A. Wan and R. van der Merwe. The unscented Kalman filter for nonlinear estimation. In *Proceedings of IEEE Symposium (AS-SPCC)*, pages 153–158, Lake Louise, Alberta, Canada, October 2000.

[83] Y Wang, CL Liu, JD Storey, RJ Tibshirani, D Hershlag, and PO Brown. Precision and functional specificity in mrna decay. *PNAS*, 99:5860–5865, 2002.

[84] R Weissleder. Molecular imaging in cancer. *Science2006*, 312(5777):1168, May 2006.

[85] R. Zhan and J. Wan. Iterated unscented Kalman filter for passive target tracking. *IEEE Transaction on Aerospace and Electronic Systems*, 43(3):1155–1162, July 2007.