Random Sampling of a Continuous-time Stochastic Dynamical System: Analysis, State Estimation, and Applications

by Mario Micheli

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences,
University of California at Berkeley, in partial satisfaction
of the requirements for the degree of

Master of Science, Plan II.

Approval for the Report and Comprehensive Examination:

Committee:

Prof. Shankar S. Sastry
Research Advisor

Date

Prof. Michael I. Jordan
Second Reader

Date

Acknowledgements

One realizes how much he loves a place, how much he owes to it, and how much he will miss it, only when he has to leave.

In my case, Berkeley is the place. Even though I was actually the one who decided to leave, I have to say that Berkeley has been one of the most amazing experiences of my life (well, so far!), for many reasons. First of all, for the unique academic environment: the fame UCB has all over the world is fully deserved. But also, by all means just as important, for the human experience I have had: in Berkeley I have come in contact with people from all over the world, and the cultural exchange I have experienced has been amazing; the acceptance and welcoming of diversity that is so typical of Berkeley has had a remarkable influence on my view of the world and society; in Berkeley I have made my best friends, whom I shall stay in contact with for the rest of my life.

Therefore, my first thanks go Berkeley and its community. My very special thanks go to Prof. Shankar Sastry, my advisor, who made all this possible; in particular, I am grateful to him for his precious advice and human understanding. Also, I wish to thank Prof. Michael Jordan, who encouraged me in my research on stochastic systems and supported me while I was writing this thesis.

The support I have received from my parents has been decisive, and I thank them for it. "Follow your dreams", my mother told me many times. Well, I have tried, and I think I came very close to that.

This Master's thesis is dedicated to my girlfriend Myriam, who never stopped showing me her intense support, even though we've been an ocean and a couple of continents apart for so long. Her strength has been unbelievable, even in the hardest times. I can't think of anyone who could have stood all that she has, and I'll be forever grateful to her.

Last, but not least... Go Bears!

Mario Micheli

Berkeley, California
December 2001

Contents

List of Figures

Notation

We made an effort to make the notation we used in the text as standard as possible, so that anyone with some background in Theory of Probability [6] should be able to read it with little difficulty. In any case, here is a brief list of the symbols that are most frequently used:

$\mathbb{N}$	set of natural numbers: $1, 2, 3, \ldots$
$\mathbb{Z}$	set of integer numbers: $\ldots, -2, -1, 0, 1, 2, \ldots$
$\mathbb{Z}^+$	set of nonnegative integer numbers: $0, 1, 2, 3, \ldots$
$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
$\Re[z]$, $\Im[z]$	real and imaginary parts of a complex number z
$\mathcal{S}^n$	set on n-tuples of elements belonging to set $\mathcal{S}$
$\mathcal{S}^{m\times n}$	set on $m \times n$ arrays with entries in set $\mathcal{S}$
δ_{ij}	Kronecker's delta
$\delta(\cdot)$	Dirac delta distribution (generalized function)
$1(\cdot)$	indicator function
$\log(\cdot)$	natural logarithm
$F_X(\cdot)$	prob. distribution function of random variable (or random vector) X
$f_X(\cdot)$	prob. density function of random variable (or random vector) X
$\mathbb{P}[\mathcal{A}]$	probability of an event $\mathcal{A}$
$\mathbb{E}[X]$	expectation of random variable (or vector) X
$\mathrm{Var}[X]$	variance (or covariance matrix) of random variable (or vector) X
$\mathcal{U}[a, b]$	family of uniformly distributed r.v.'s in interval $[a, b]$
$\mathcal{E}(\lambda)$	family of unilaterally exponential r.v.'s, with parameter λ
$\mathcal{P}(\lambda)$	family of Poisson r.v.'s, with parameter λ
$\mathcal{N}(m, K)$	family of (multivariate) Gaussian distributions, with $\mathbb{E}[X] = m \in \mathbb{R}^n$ and $\mathrm{Var}[X] = K \in \mathbb{R}^{n\times n}$

Within each chapter Theorems, Lemmas, Propositions and Corollaries are numbered in the order in which they appear. Their text is in *emphasized* font. Proofs in the text end with symbol $\Box$ to demarcate the proof from the following text.

Chapter 1

Introduction

In this first chapter we will present the state estimation problem that will constitute the main topic of this thesis: we will justify its study by proving that it models the physical problem of estimating a dynamical variable that is measured by a network of sensors. In the last section we will describe how the following chapters are organized, emphasizing the main goals of our investigation.

1.1 Problem overview

In this project we study the problem of state estimation of a linear, stochastic dynamical system whose state is a *continuous* function of time, and whose output is measured only in *discrete* time instants; such instants are not evenly spaced in time but are separated by random time intervals.

In other terms, we are considering the following dynamical model:[1]

$$\begin{cases} \dot{x}(t) &= Fx(t) + Gv(t) \\ y(t_k) &= Cx(t_k) + z(t_k) \end{cases} \qquad t \in \mathbb{R}, \quad k \in \mathbb{N} \tag{1.1}$$

where $x : \mathbb{R} \to \mathbb{R}^n$, $y : \mathbb{R} \to \mathbb{R}^p$, are stochastic processes, and $F \in \mathbb{R}^{n \times n}$, $G \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ are time-invariant real matrices. In linear model (1.1) two different white, zero-mean Gaussian stationary noise inputs appear: continuous-time noise $v(t)$, $t \in \mathbb{R}$, and discrete time noise $z(t_k)$, indexed by parameter $k \in \mathbb{N}$, with

$$\mathbb{E}[v(t)\, v^T(\tau)] = S\, \delta(t - \tau), \qquad \text{and} \qquad \mathbb{E}[z(t_i)\, z^T(t_j)] = R\, \delta_{ij},$$

[1] A formally correct way to write the first of equations (1.1) would be the following:

$$dx = F\, x\, dt + G\, dw$$

which is the standard notation for stochastic differential equations (SDEs), whose solutions are called Itô processes, or diffusions [6]. However, we preferred to write it as in (1.1) since it is closer to the standard notation for linear dynamical system used by the Electrical Engineering community.

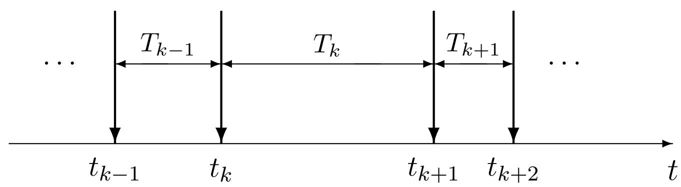


Figure 1.1: Random sampling process.

where $\delta(\cdot)$ is the Dirac distribution while δ_{ij} is Kronecker's delta. $S \in \mathbb{R}^{m\times m}$ and $R \in \mathbb{R}^{p\times p}$ are constant positive definite matrices; we also assume that $v(\cdot)$ and $z(\cdot)$ are independent of each other. We will call *state equation* the first of (1.1), whereas the second one will be named *measurement equation*.

Time instants $\{t_k\}_{k=1}^{\infty}$ are positive, ordered (i.e. $t_{k+1} > t_k$, $\forall k \in \mathbb{N}$) and are such that time intervals

$$T_0 \stackrel{\Delta}{=} t_1, \qquad T_k \stackrel{\Delta}{=} t_{k+1} - t_k \quad \text{for } k \geq 1$$

are i.i.d. exponential random variables[2] with parameter λ, $T_k \sim \mathcal{E}(\lambda)$; we shall also assume that T_k and $v(t)$ are independent, for all $k \in \mathbb{N}$ and $t \in \mathbb{R}$; Figure 1.1 illustrates the random sampling. In the language of stochastic processes, time instants $\{t_k\}_{k=1}^{\infty}$ are the arrivals of a *Poisson process* [6], and λ represents its *intensity*. Indicating with $N(s,t]$ the number of arrivals in time interval $(s,t]$ we have that $N(s,t] \sim \mathcal{P}\big(\lambda(t-s)\big)$, i.e. it is a Poisson r.v. of parameter $\lambda(t-s)$:

$$P\big[N(s,t] = k\big] = e^{-\lambda(t-s)}\,\frac{[\lambda(t-s)]^k}{k!}\,, \qquad k \geq 0$$

and $\mathbb{E}\big[N(s,t]\big] = \mathrm{Var}\big[N(s,t]\big] = \lambda(t-s)$. Being the mean of the number of arrivals in time interval $(s,t]$ *proportional* to λ, the name *intensity* is justified for it.

Given such a model, we wish to estimate state $x(t)$ in an *on-line* manner at any time instant $t \in \mathbb{R}$, using measurements set $\mathcal{Y}(t) = \{y(t_k) : t_k \leq t\}$ and $\mathcal{T}(t) = \{T_{k-1} : t_k \leq t\}$; in other words at time t we only know the realization of the Poisson process *up to* time t, and the corresponding measurements. We also wish to have, at every time instant, a measure of the estimation error, i.e. an error variance.

We shall see in the following chapters that the relation between parameter λ and the eigenvalues of matrix F (which represent the dynamics of state x) plays a fundamental role in determining the properties, and the effectiveness, of the estimation

[2]An exponential random variable $X \sim \mathcal{E}(\lambda)$ has probability density $f_X(x) = \lambda\, e^{-\lambda x}\ 1(x \geq 0)$; its mean and variance are respectively given by $\mathrm{E}[X] = 1/\lambda$ and $\mathrm{Var}[X] = 1/\lambda^2$. The exponential r.v. has some very distinctive properties; for example, it is a well known fact that it is *memoryless*; also, among all random variables whose mean is $1/\lambda$ and whose support is $\mathbb{R}^+ = [0,\infty)$, it is the one with maximum *differential entropy* [4].

equations: intuitively, one can imagine that a system with positive eigenvalues is somehow hard to track when the sampling intensity is too low.

Remark: sampling of Brownian motion. Note that when $F = 0$ in (1.1) then we are dealing with random sampling of *Brownian motion*, which turns out to be a particular case of our problem. In the next Chapters we will give some special attention to this subproblem.

1.2 Motivation: Sensor Networks

Our study finds its motivations in the analysis of sensor networks. Assume that a physical process $x(t)$ is describable by the first of equations (1.1), i.e. by the state equation. A certain number N of sensors measure such process in the following way: each one of them periodically yields a noisy measurement of process $x(t)$, every T seconds; the measurement equation for the i-th sensor is:

$$y_i(nT + \xi_i) = Cx(nT + \xi_i) + z(nT + \xi_i), \qquad n \in \mathbb{Z},\ i \in \{1, 2, \dots, N\}$$

where C and $z(\cdot)$ are defined as in the previous section, $\xi_i \in [0, T)$ is the "phase" of each sensor (i.e. how long after time $t = 0$ the i-th sensor yields its first measurement) and sampling period T is the same for all sensors.

If the sensors are *not synchronized* (i.e. the ξ_i's are all different) then the process that is obtained by summing the arrivals of *all* sensors may be approximated by a Poisson process with intensity $\lambda = N/T$. In other words, the waiting time between two successive arrivals may be approximated by an exponential random variable $\mathcal{E}(\lambda)$. This is rigorously justified by the following proposition.

Proposition 1.1 *Let $\lambda \in (0, \infty)$, $N \in \mathbb{N}$ and let $\{\xi_i\}_{i=1}^N$ be N i.i.d., uniformly distributed random variables in interval $[0, N/\lambda]$. Let $Y_N = \min_i\{\xi_i\}$, that is the time before the "first arrival" occurs in $[0, N/\lambda]$. Then*

$$Y_N \xrightarrow{D} \mathcal{E}(\lambda) \text{ as } N \to \infty,$$

where $\xrightarrow{D}$ denotes convergence in distribution.

Proof. For $y \in [0, N/\lambda]$ the probability distribution function of r.v. Y_N is given by

$$\begin{aligned} F_{Y_N}(y) &= \mathbb{P}[Y_N \le y] = \mathbb{P}[\forall i, \xi_i > y] = \prod_{i=1}^{N} \mathbb{P}[\xi_i > y] = \\ &= \prod_{i=1}^{N} \left(1 - \frac{\lambda y}{N}\right) = \left(1 - \frac{\lambda y}{N}\right)^N \longrightarrow e^{-\lambda y}, \text{ as } N \to \infty; \end{aligned}$$

therefore, Y_N tends, in distribution, to the exponential r.v. $\mathcal{E}(\lambda)$ as $N \to \infty$. □

Hence, if T is known and N is sufficiently large, defining $\lambda = N/T$ it is reasonable to consider the waiting time for the next arrival as an exponential r.v. of parameter λ; in other terms, we may think of the arrivals as if they were generated by a Poisson process. Thus the characterization of the waiting time we made in the previous section in the definition of our problem is justified; such choice of random variable will be of great usefulness in the development of our study.

1.3 Thesis structure

In the next Chapter we shall propose a method for estimating state $x(t)$; it will consist of sampling the first of equations (1.1) in correspondence of the sampling instants $\{t_k\}$, and then apply Kalman Filtering to the discrete-time system so obtained; we shall then statistically characterize the parameters of such system in the one-dimensional case ($m = n = p = 1$). In particular, we shall see that the estimation error is a Markov process.

In Chapter 3 we shall compute a complete statistical description of the estimation error process, always in the one-dimensional case; in particular, we shall discuss the strong, and sometimes subtle, dependence of such description on the relationship between the continuous-time dynamics of the first of equations (1.1) and the sampling intensity λ.

Assuming we have control on parameter λ (which is proportional to the number of sensors N), in Chapter 4 we shall show how to bound the error estimation variance (with arbitrary probability) by an appropriate choice of the sampling intensity (i.e. by choosing an appropriate number of sensors).

Finally, Chapter 5 is dedicated to conclusions and a brief discussion on possible directions of future research.

Chapter 2

Sampling the State Equation

Having formulated our problem in the previous Chapter and related it to the physical situation of sensor networks, we are now ready to start investigating ways to perform state estimation. In this Chapter we illustrate a Kalman Filter-based method to do it in correspondence of the Poisson arrivals that characterize the measurement process. Such method implies sampling the state equation (i.e. the first of (2.1)) thus yielding a discrete-time system, as we shall see, with *random* parameters.

We compute a complete statistical description of these parameters it the one-dimensional case, emphasizing its dependence on the original system's dynamics and on the sampling intensity, i.e. on the number of sensors in the network we are modelling. The study of the effectiveness of the proposed state estimation algorithm will be the topic of the next Chapter.

2.1 Sampled System

In order to estimate state x at time instants t_k, $k \in \mathbb{N}$, we consider the *sampled* version (see, for example, [5] or [11]) of the *state equation*, i.e. the first of (1.1), where the samples are taken in correspondence of the Poisson arrivals.

The discrete-time, stochastic system that is obtained this way is the following:

$$\begin{cases} x(t_{k+1}) & = & A_k x(t_k) + w(t_k) \\ y(t_k) & = & C x(t_k) + z(t_k) \end{cases} \qquad k \in \mathbb{N}, \tag{2.1}$$

where matrix A_k and input noise $w(t_k)$ are given, respectively, by exponential matrix

$$A_k = e^{F(t_{k+1}-t_k)} = e^{FT_k}, \tag{2.2}$$

and vector

$$w(t_k) \;=\; \int_{t_k}^{t_{k+1}} e^{F(t_{k+1}-\tau)}\, G\, v(\tau)\, d\tau \;=\; \int_0^{T_k} e^{F\tau}\, G\, v(t_{k+1}-\tau)\, d\tau\,. \tag{2.3}$$

Note that A_k depends on random variable T_k, therefore it is a *random matrix*. The randomness of noise $w(t_k)$ derives from its dependence on both continuous-time noise $v(t)$ and random variable T_k, which appears as second extreme of integration in (2.3).

For on-line estimation purposes we are interested in calculating the mean and covariance matrix of $w(t_k)$, *given* time interval T_k (in fact when estimating state $x(t_{k+1})$ time interval T_k is known; further details will be given in the next section). It is simple to verify that $\mathbb{E}\big[w(t_k)\,\big|\,T_k\big] = 0$. We will define matrix Q_k as follows:

$$\begin{aligned}
Q_k &\triangleq \mathbb{E}\big[w(t_k)w^T(t_k)\,\big|\,T_k\big] \\
&= \mathbb{E}\left[\int_0^{T_k} d\tau \int_0^{T_k} d\tau'\, e^{F\tau} G\, v(t_{k+1}-\tau)\, v^T(t_{k+1}-\tau')\, G^T e^{F^T\tau'} \,\middle|\, T_k\right] \\
&= \int_0^{T_k} d\tau \int_0^{T_k} d\tau'\, e^{F\tau} G\, \mathbb{E}\big[v(t_{k+1}-\tau)\, v^T(t_{k+1}-\tau')\,\big|\,T_k\big]\; G^T e^{F^T\tau'} \\
&\overset{(*)}{=} \int_0^{T_k} d\tau \int_0^{T_k} d\tau'\, e^{F\tau} G\, S\, \delta(\tau-\tau')\, G^T e^{F^T\tau'} \\
&= \int_0^{T_k} d\tau\, e^{F\tau} G S G^T e^{F^T\tau}\,,
\end{aligned} \tag{2.4}$$

where step $(*)$ is justified by the independence between T_k and $v(t)$.

One can prove in a similar way that $\mathbb{E}\big[w(t_j)w^T(t_k)\,\big|\,T_j, T_k\big] = 0$ for $j \neq k$:

$$\begin{aligned}
&\mathbb{E}\big[w(t_k)w^T(t_j)\,\big|\,T_k, T_j\big] = \\
&\quad= \mathbb{E}\left[\int_0^{T_k} d\tau \int_0^{T_j} d\tau'\, e^{F\tau} G\, v(t_{k+1}-\tau)\, v^T(t_{j+1}-\tau')\, G^T e^{F^T\tau'} \,\middle|\, T_k, T_j\right] \\
&\quad= \int_0^{T_k} d\tau \int_0^{T_j} d\tau'\, e^{F\tau} G\, \mathbb{E}\big[v(t_{k+1}-\tau)\, v^T(t_{j+1}-\tau')\,\big|\,T_k, T_j\big]\; G^T e^{F^T\tau'} \\
&\quad= \int_0^{T_k} d\tau \int_0^{T_j} d\tau'\, e^{F\tau} G\, S\, \delta\big((t_{k+1}-\tau)-(t_{j+1}-\tau')\big)\, G^T e^{F^T\tau'}\,,
\end{aligned}$$

where, again, the last step is justified by the independence between process $\{T_k\}_{k=1}^{\infty}$ and $v(t)$; since intervals $(t_{k+1}-T_k, t_{k+1}) = (t_k, t_{k+1})$ and $(t_{j+1}-T_j, t_{j+1}) = (t_j, t_{j+1})$ are disjoint for $j \neq k$ (see Figure 1.1), we have that the above integral must be equal to zero. Therefore, random process $w(t_k)$, conditioned on $\{T_j\}_{j=0}^{\infty}$, is white Gaussian[1] noise. In particular, $w(t_k)|T_k \sim \mathcal{N}(0, Q_k)$.

Matrix Q_k, expressed by (2.4), may be obtained as the solution of linear matrix equation:

$$\dot{Q}(t) = FQ(t) + Q(t)F^T + GSG^T, \tag{2.5}$$

[1] If a linear transformation, like the integral operator in (2.3), is applied to a Gaussian stochastic process then the resulting process is Gaussian as well (see [3, p. 37]).

with initial condition $Q(0) = 0$, calculated in T_k, i.e. $Q_k = Q(T_k)$ (see Appendix A and [2, p. 58]). In general, such problem can be solved *numerically* (again, see Appendix A); an analytical solution may be found in only a few cases, depending on the structure of matrices F, S and G. For the purposes of state estimation, which will be the topic of the next section, Q_k has to be evaluated on-line: in fact one only knows the value random variable T_k assumes only at time t_{k+1}, i.e. in correspondence of the $(k+1)$-th Poisson arrival.

Remark. Note that matrix Q_k, as a function of T_k, is a *random matrix* (as well as A_k). Therefore we can say that in our model we have two levels of stochasticity, since Q_k is the random covariance of Gaussian noise $w(t_k)$: the realization of r.v. T_k determines the value of Q_k by (2.4), then $w(t_k)$ is still random with covariance Q_k.

2.2 State Estimation algorithm

At this point we perform state estimation on discrete-time dynamical system (2.1) using the Kalman Filter equations (see [9], [11], or [7]). Note that system (2.1) is *time-variant*[2] as both (random) state matrix A_k and (random) covariance matrix Q_k depend on T_k, i.e. on parameter k.

2.2.1 Estimation at arrival times

Traditional Kalman Filtering is a state estimation algorithm that applies to discrete-time state-space models like the one represented in (2.1), in the case where all parameters (matrices and noise variances) are known *a priori*. When noise is Gaussian it yields the "best" state estimate $\hat{x}(t_k)$, meaning that it minimizes quantity $\mathbb{E}\big[||x(t_k) - \hat{x}(t_k)||^2 \big| \{y(t_j)\}_{j\le k}\big]$, the mean of the square distance between $x(t_k)$ and $\hat{x}(t_k)$, for all k.

While in the traditional case sampling period T is fixed, in our case it is random: this implies that matrices $\{A_k\}_{k=1}^{\infty}$ and $\{Q_k\}_{k=1}^{\infty}$ are a priori unknown. We shall have to modify the definitions and the equations in the way that is described below. In any case the modified algorithm will still yield the optimal estimate, in the sense that it will minimize $\mathbb{E}\big[||x(t_k) - \hat{x}(t_k)||^2 \big| \{y(t_j), T_{j-1}\}_{j\le k}\big]$.

Define the following quantities (refer to Figure 1.1):

$$\hat{x}_{k|k} \triangleq \mathbb{E}\big[x(t_k)\big|\{y(t_j), T_{j-1}\}_{j\le k}\big]\,, \quad P_{k|k} \triangleq \mathrm{Var}\big[x(t_k)\big|\{y(t_j), T_{j-1}\}_{j\le k}\big]\,,$$
$$\hat{x}_{k+1|k} \triangleq \mathbb{E}\big[x(t_{k+1})\big|\{y(t_j), T_j\}_{j\le k}\big]\,, \quad P_{k+1|k} \triangleq \mathrm{Var}\big[x(t_{k+1})\big|\{y(t_j), T_j\}_{j\le k}\big]\,;$$

following the same lines of the proof of Kalman filter's equations, one gets the

[2] Note that the notion of *time* for this system is different from time parameter t in system (1.1), as the parameter we are now referring to is k.

following ones:

$$\begin{aligned}
\hat{x}_{k+1|k} &= A_k \hat{x}_{k|k} && (2.6)\\
P_{k+1|k} &= A_k P_{k|k} A_k^T + Q_k && (2.7)\\
\hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} + P_{k+1|k} C^T (C P_{k+1|k} C^T + R)^{-1} (y_{k+1} - C \hat{x}_{k+1|k}) && (2.8)\\
P_{k+1|k+1} &= P_{k+1|k} - P_{k+1|k} C^T (C P_{k+1|k} C^T + R)^{-1} C P_{k+1|k}\,. && (2.9)
\end{aligned}$$

As for the ordinary Kalman filter, we will name the first two of the above formulas *time update equations*, while the least two will be called *measurement update equations*. Note that, as in the traditional case, the evolution of $P_{k|k}$ does not directly depend on the measured quantity y_k.

At this point we should take a short break in order to fully understand the main differences between these equations and the ordinary Kalman Filter. First of all A_k and Q_k are functions of T_k, therefore they are not deterministic but random matrices; hence the sequence of error covariance matrices $\{P_{k|k}\}_{k=1}^{\infty}$, which in the ordinary case is (for all k) completely deterministic and can be computed off-line before measurements start, in our case is itself a *random process*.

Secondly, while in the ordinary case time update $k \to k+1$ (i.e. equation (2.6)) can be performed at time t_k, in the case of random sampling one has to wait time t_{k+1} as matrices A_k and Q_k are needed in equations (2.6) and (2.7): both of them depend on $T_k = t_{k+1} - t_k$, and at time t_k one does not know when arrival t_{k+1} will occur, i.e. what value T_k will take. Therefore the time update and measurement update steps will *both* be performed at arrival time t_{k+1}.

2.2.2 $\{P_{k|k}\}_{k=0}^{\infty}$ is a Markov process

Plugging equation (2.7) into (2.9) yields the following:

$$\begin{aligned}
P_{k+1|k+1} = {}& (A_k P_{k|k} A_k^T + Q_k) + \\
& -(A_k P_{k|k} A_k^T + Q_k) C^T (C P_{k+1|k} C^T + R)^{-1} C (A_k P_{k|k} A_k^T + Q_k);
\end{aligned}$$

we now apply the Matrix Inversion Lemma[3] to the right hand side of the above equation and get the following, simpler form:

$$P_{k+1|k+1} = \left((A_k P_{k|k} A_k^T + Q_k)^{-1} + C^T R^{-1} C \right)^{-1}. \qquad (2.10)$$

Define function $h(A_k, Q_k, P_{k|k})$ to be equal to the right hand side of (2.10); then let $P_{0|0} = \mathrm{Var}[x(0)]$, i.e. the *a priori* covariance matrix of process x; assume $x(0)$ is

[3]The Matrix Inversion Lemma states that the following equality holds:

$$(S + UZV)^{-1} = S^{-1} - S^{-1} U (V S^{-1} U + Z^{-1})^{-1} V S^{-1},$$

whenever the matrices are of appropriate size and the required inverses exist [9]. In our case, pick $S = (A_k P_{k|k} A_k^T + Q_k)^{-1}$, $Z = R^{-1}$, $U = C^T$ and $V = C$: the left hand side of the equation above is given by the right hand side of (2.10).

a Gaussian r.v., independent of all other random variables. We will have:[4]

$$\begin{aligned}
&\mathbb{P}\big[\, P_{k+1|k+1} \le p_{k+1} \,\big|\, P_{k|k} = p_k, P_{k-1|k-1} = p_{k-1}, \ldots, P_{0|0} = p_0 \,\big] \\
&\quad = \quad \mathbb{P}\big[\, h(A_k, Q_k, P_{k|k}) \le p_{k+1} \,\big|\, P_{k|k} = p_k, P_{k-1|k-1} = p_{k-1}, \ldots, P_{0|0} = p_0 \,\big] \\
&\quad = \quad \mathbb{P}\big[\, h(A_k, Q_k, p_k) \le p_{k+1} \,\big|\, P_{k|k} = p_k, P_{k-1|k-1} = p_{k-1}, \ldots, P_{0|0} = p_0 \,\big] \\
&\quad = \quad \mathbb{P}\big[\, h(A_k, Q_k, p_k) \le p_{k+1} \,\big|\, P_{k|k} = p_k \,\big] \,,
\end{aligned}$$

where the last step is so justified: A_k and Q_k are functions of T_k and noise $v(t)$, $t \in (t_k, t_{k+1})$ (see (2.2) and (2.4)), whereas $\{P_{j|j}\}_{j=0}^{k-1}$ are functions of $\{T_j\}_{j=1}^{k-2}$, $v(t)$, $t \in (t_k, t_{k+1})$ and $x(0)$, which, by hypothesis, are independent of T_k and $v(t)$, $t \in (t_k, t_{k+1})$; hence A_k and Q_k are independent of $\{P_{j|j}\}_{j=0}^{k-1}$.

Therefore, process $\{P_{k|k}\}_{k=0}^{\infty}$ is Markov; in fact it is also *homogeneous* since the T_k's are i.i.d. and noise $v(t)$ is white and Gaussian. We shall investigate important consequences of this fact in the next Chapter, where we will compute a complete statistical description of the process in question.

2.2.3 Estimation between Poisson arrivals.

Being $x(t)$ a continuous-time process one may be interested in estimating its value for any time $t \in \mathbb{R}$, which may not necessarily coincide with a Poisson arrival. For this purpose, define:

$$\hat{x}_t \triangleq \mathbb{E}\big[x(t) \,\big|\, \{y(t_j), T_{j-1}\}_{t_j \le t}\big] \,, \qquad P_t \triangleq \mathrm{Var}\big[x(t) \,\big|\, \{y(t_j), T_{j-1}\}_{t_j \le t}\big] \,.$$

Then one can easily show that for $t \in (t_k, t_{k+1})$ the above quantities may be expresses as follows:

$$\begin{aligned}
\hat{x}_t &= e^{F(t-t_k)} \hat{x}_{k|k} \\
P_t &= e^{F(t-t_k)} P_{k|k}\, e^{F^T(t-t_k)} + \int_0^{t-t_k} d\tau\; e^{F\tau} G S G^T e^{F^T \tau}.
\end{aligned}$$

Random process $\hat{x}_t$ (for a given realization of the Poisson process) is a *piecewise continuous* function of time; discontinuities occur in correspondence of the Poisson arrivals. In the case of *Brownian motion* ($F = 0$) the above equations take, for $t \in (t_k, t_{k+1})$, the simpler form:

$$\begin{aligned}
\hat{x}_t &= \hat{x}_{k|k} \\
P_t &= P_{k|k} + G S G^T (t - t_k);
\end{aligned}$$

note in particular that $\hat{x}_t$ becomes *piecewise constant* whereas P_t becomes *piecewise linear*; in both cases discontinuities occur in correspondence of the Poisson arrivals $\{t_k\}_{k=1}^{\infty}$.

[4]With the expression $P_{k+1|k+1} \le p_{k+1}$ we intend that every element of matrix $P_{k+1|k+1}$ is less than or equal to the corresponding element in matrix p_{k+1}.

2.3 Characterization of Sampled System's parameters

As we remarked more than once in the previous sections quantities A_k and Q_k are random matrices that depend on time interval T_k. An interesting problem is finding a statistical description of such quantities, which have to be computed on-line at every Poisson arrival in order to perform state estimation.

In the one-dimensional case ($m = n = p = 1$) it is possible to find a *complete statistical description* (given by a probability density) of A_k and Q_k in function of the continuous system's parameters and the Poisson process intensity, as we will show in the following subsections. Before proceeding, let us introduce some notation for one-dimensional continuous systems: $F = \phi$, $G = g$, $S = \sigma^2$. In the 1-D case ϕ represents the only eigenvalue of the continuous-time system matrix; such system is therefore asymptotically stable for $\phi < 0$, simply stable for $\phi = 0$, and unstable for $\phi > 0$.

2.3.1 Statistical description of A_k

For 1-D systems parameter A_k, given by (2.2), is simply expressed as

$$A_k = e^{\phi T_k}, \quad \phi \in \mathbb{R}. \tag{2.11}$$

We wish to find its complete statistical description, knowing that $T_k \sim \mathcal{E}(\lambda)$.

We note that in the case $\phi = 0$ we have $A_k = 1$ w.p. 1 (with probability one), which is something that one could foresee ($\dot{x}(t) = v(t)$ implies $x(t_{k+1}) = w(t_k)$).

In order calculate the probability density of A_k in the other two cases (asymptotic stability and instability of the continuous time-system), we shall use the following theorem, provided by the Theory of Probability (see [6], [10]).

Theorem 2.1 *Let X be a continuous r.v. with probability density $f_X(x)$, and $g(\cdot)$ a real function, continuous and differentiable in $\mathbb{R}$, such that its derivative $g'(\cdot)$ is equal to zero only for a finite or countably infinite set of points. Fix any $y \in \mathbb{R}$ and define $\mathcal{H}(y) = \{x_1, x_2, \dots\} = g^{-1}(y)$, i.e. $\mathcal{H}(y)$ is the set of points that are mapped in y by function $g(\cdot)$, and suppose that $g'(x) \neq 0$ for any $x \in \mathcal{H}(y)$. Then the probability density of random variable $Y \stackrel{\Delta}{=} g(X)$ is equal to:*

$$f_Y(y) = \sum_{x \in \mathcal{H}(y)} \frac{f_X(x)}{|g'(x)|},$$

intending that $f_Y(y) = 0$ if $\mathcal{H}(y) = \emptyset$.

Exponential function $g(t) = e^{\phi t}$ satisfies the hypotheses of the above theorem, which may therefore be used to calculate the probability density of random variable $A_k = g(T_k)$. Performing the necessary calculations yields different results, depending on the stability of the original continuous-time system. In fact, the following proposition holds.

Proposition 2.2 *In the one-dimensional case, the probability density of A_k has the following form, depending on the stability of the original continuous-time system (1.1):*

- *for* **asymptotically stable systems** *($\phi < 0$):*

$$f_{A_k}(a) = -\frac{\lambda}{\phi\, a^{\frac{\lambda}{\phi}+1}} \cdot 1(0 < a \leq 1) \tag{2.12}$$

- *for* **unstable systems** *($\phi > 0$):*

$$f_{A_k}(a) = \frac{\lambda}{\phi\, a^{\frac{\lambda}{\phi}+1}} \cdot 1(a \geq 1)\,,$$

- *for* **Brownian motion** *($\phi = 0$): $f_{A_k}(a) = \delta(a-1)$,*

where $1(\cdot)$ and $\delta(\cdot)$ are, respectively, the indicator and Dirac's delta functions.

Proof. The case $\phi = 0$ is trivial. Now, let $\phi \neq 0$ and define $g(t) = e^{\phi t}$. If $\phi < 0$ then $\mathcal{H}(a) = g^{-1}(a) = \{\frac{1}{\phi}\log a\}$ for $a \in (0,1]$, $\mathcal{H}(a) = \emptyset$ otherwise. If $\phi > 0$ then $\mathcal{H}(a) = \{\frac{1}{\phi}\log a\}$ for $a \geq 1$, and $\mathcal{H}(a) = \emptyset$ otherwise. At this point, just apply Theorem 2.1. $\square$

These curves, for both the asymptotically stable and the unstable cases, are plotted in Figure 2.1. First of all note that for $\phi < 0$ probability density $f_{A_k}(a)$ has *bounded support*: in particular we have that $\mathbb{P}\big[A_k \in (0,1]\big] = 1$ for all $k \in \mathbb{N}$, i.e. sampled system (2.1) is asymptotically stable w.p. 1 (the presence of noise $w(t_k)$ will actually avoid convergence of state x to zero). This result may not be too surprising.

More interesting things occur in the case of instability of the original continuous-time system. First of all, for $\phi > 0$ probability density $f_{A_k}(a)$ has *unbounded support*, and $\mathbb{P}\big[A_k \in (1,+\infty)\big] = 1$ for all $k \in \mathbb{N}$, i.e. sampled system (2.1) is *unstable* w.p. 1. Secondly, probability density $f_{A_k}(a)$ *does not have all moments*.

In fact, one can see that integral

$$\mathbb{E}\big[A_k^n\big] = \int_{-\infty}^{\infty} \big(e^{\phi t}\big)^n f_{T_k}(t)\,dt = \lambda \int_0^{\infty} e^{(n\phi-\lambda)t}\,dt$$

only converges for $\lambda > n\phi$, and in that case it yields:

$$\mathbb{E}\big[A_k^n\big] = \frac{\lambda}{\lambda - n\phi}\,. \tag{2.13}$$

In particular, the mean and variance of A_k are given by:

$$\mathbb{E}\big[A_k\big] = \frac{\lambda}{\lambda-\phi}\,, \qquad \mathrm{Var}\big[A_k\big] = \mathbb{E}\big[A_k^2\big] - \mathbb{E}\big[A_k\big]^2 = \frac{\phi^2\lambda}{(\lambda-2\phi)(\lambda-\phi)^2}$$

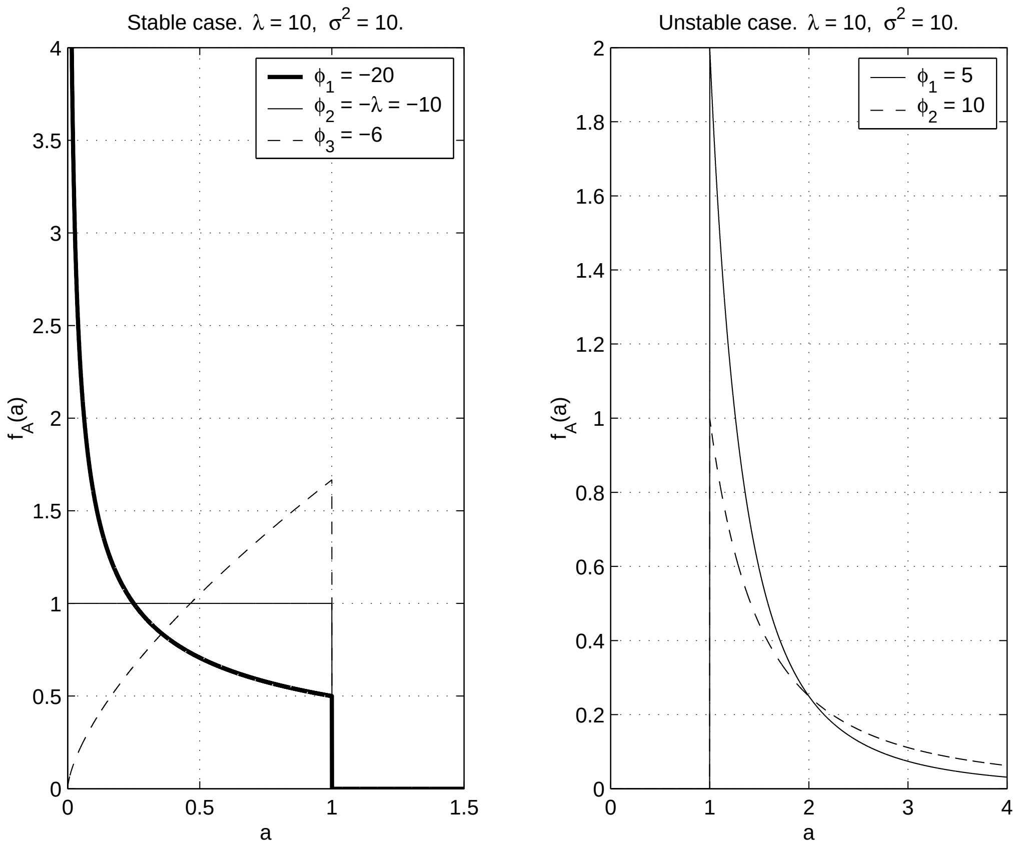


Figure 2.1: Probability density $f_{A_k}(a)$, plotted for different values of eigenvalue ϕ (there is a scale difference between the two figures). Note from equation (2.12) that if $\lambda = -\phi$ then A_k has a *uniform* probability density.

and they exist if, respectively, $\lambda > \phi$ and $\lambda > 2\phi$. Note that, when $\phi > 0$, $\text{Var}[A_k]$ is a monotone increasing function of ϕ (provided that $\lambda > 2\phi$), i.e. it increases with the "instability" of the original continuous-time dynamical system; furthermore it is a monotone decreasing function of λ, i.e. it gets smaller for more "intense" measurement arrivals.

Condition $\lambda > n\phi$ is verified for any $n \in \mathbb{N}$ in the case of asymptotically stable systems ($\phi < 0$), because intensity λ is always positive. On the other hand if the system is unstable ($\phi > 0$) there exists an $N^* \in \mathbb{N}$ such that $\lambda < n\phi$ for all $n > N^*$; for these values of n, random variable A_k does not have finite n-th order moments.

In particular if $\lambda < \phi$ then A_k does not even have a finite first-order moment, i.e. *its mean diverges to infinity*. This can be interpreted as follows, remembering the meaning of parameters λ (which is the intensity of the Poisson process, i.e. the average number of measurements in a time interval of one second) and ϕ (the greater is $\phi > 0$, the more "unstable" is continuous dynamical system (1.1)):

if the measurements are not "dense" enough in time,[5] or if the degree of instability of the continuous-time dynamical system is too high, then the mean of A_k tends to infinity; therefore A_k will tend to assume very high values, that is, the degree of instability of the sampled system will tend to be high as well. All these results are summarized by the following Corollary.

Corollary 2.3 *When $\phi < 0$ random variable A_k has all moments, given by (2.13). When $\phi = 0$, r.v. A_k obviously has all moments since it is equal to one w.p. 1. Finally, when $\phi > 0$, let $n_A = \max\{n : \lambda > n\phi\}$: r.v. A_k has the n-th order moment if and only if $n \leq n_A$; when this occurs, $\mathbb{E}[A_k^n]$ is given by (2.13).*

2.3.2 Statistical description of Q_k

Analogous, but maybe even more interesting results hold for Q_k, the random variance of noise $w(t_k)$. In the case of 1-D systems if $\phi \neq 0$ integral (2.4) may be expressed in analytical form as follows:

$$Q_k = \begin{cases} g^2\sigma^2 T_k & \text{for } \phi = 0, \\ \frac{g^2\sigma^2}{2\phi}\left(e^{2\phi T_k} - 1\right) & \text{for } \phi \in \mathbb{R} \setminus \{0\}. \end{cases} \tag{2.14}$$

In order to compute its probability density we apply once again Theorem 2.1 using real function $g(t) = \frac{g^2\sigma^2}{2\phi}\left(e^{2\phi t} - 1\right)$ for $\phi \neq 0$, which satisfies the prerequisites that are demanded by the hypotheses. In the case $\phi = 0$ we simply have that Q_k is proportional to T_k, which has an exponential distribution, so Q_k has an exponential distribution as well, with mean $g^2\sigma^2/\lambda$. In general, the following proposition holds.

Proposition 2.4 *In the one-dimensional case, the probability density of Q_k has the following form, depending on the stability of the original continuous-time system (1.1):*

- *for* **asymptotically stable systems** *($\phi < 0$):*

$$f_{Q_k}(q) = \frac{\lambda}{g^2\sigma^2\left(\frac{2\phi}{g^2\sigma^2}q + 1\right)^{1+\frac{\lambda}{2\phi}}} \cdot 1\left(0 \leq q < \frac{g^2\sigma^2}{2|\phi|}\right);$$

- *for* **unstable systems** *($\phi > 0$):*

$$f_{Q_k}(q) = \frac{\lambda}{g^2\sigma^2\left(\frac{2\phi}{g^2\sigma^2}q + 1\right)^{1+\frac{\lambda}{2\phi}}} \cdot 1(q \geq 0).$$

[5]In terms of sensor networks, the number of sensors is too low.

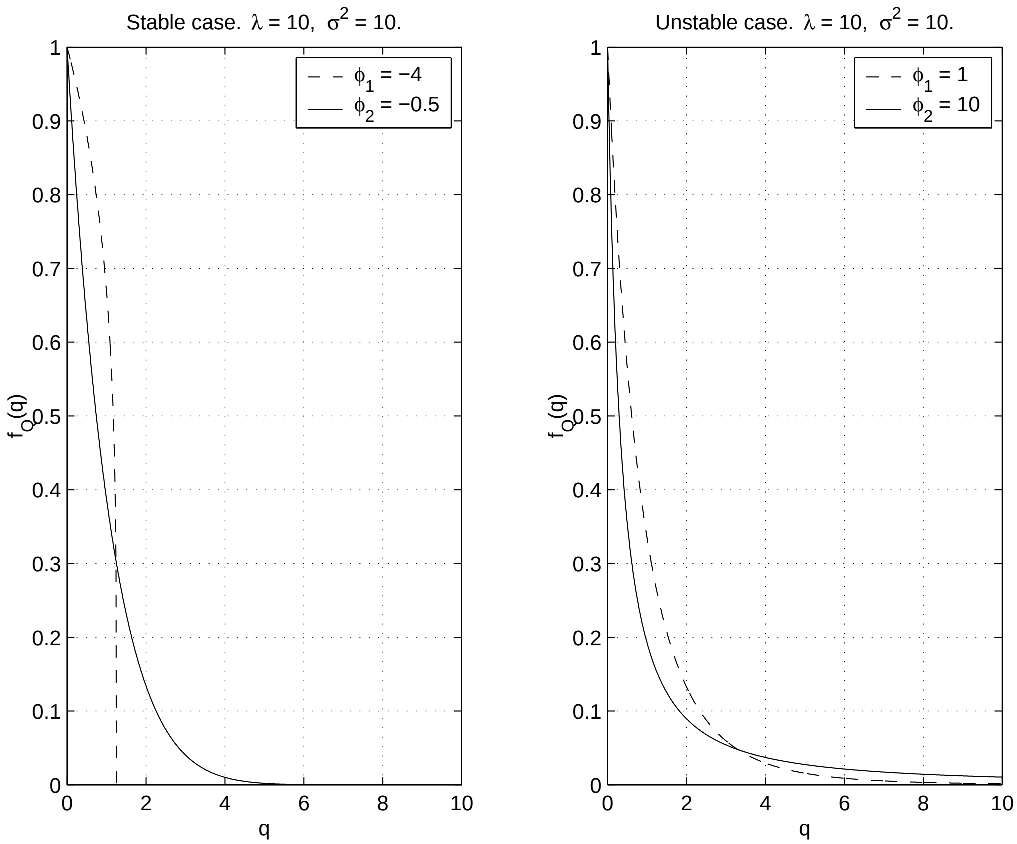


Figure 2.2: Probability density $f_{Q_k}(q)$, plotted for different values of eigenvalue ϕ. Note that in the case of asymptotic stability of the continuous-time system ($\phi < 0$) the support of f_{Q_k} is bounded, being instead *unbounded* for $\phi > 0$.

- *for* **Brownian motion** *($\phi = 0$):*

$$f_{Q_k}(q) = \frac{\lambda}{g^2\sigma^2} \exp\left(-\frac{\lambda}{g^2\sigma^2} q\right) \cdot 1(q \geq 0),$$

  i.e. $Q_k \sim \mathcal{E}\left(\frac{\lambda}{g^2\sigma^2}\right)$.

Proof. We have discussed the case $\phi = 0$ above. Now, let $\phi \neq 0$ and define $g(t) = \frac{g^2\sigma^2}{2\phi}\left(e^{2\phi t} - 1\right)$. Some computation is required to show that if $\phi < 0$ then $\mathcal{H}(q) = g^{-1}(q) = \{\frac{1}{2\phi} \log\left(\frac{2\phi}{g^2\sigma^2} q + 1\right)\}$ for $q \in \left[0, \frac{g^2\sigma^2}{2|\phi|}\right)$, $\mathcal{H}(q) = \emptyset$ otherwise. If $\phi > 0$ then $\mathcal{H}(q) = g^{-1}(q) = \{\frac{1}{2\phi} \log\left(\frac{2\phi}{g^2\sigma^2} q + 1\right)\}$ for $q \geq 0$, $\mathcal{H}(q) = \emptyset$ otherwise. Now just apply Theorem 2.1. □

These graphs of the above functions are plotted in Figure 2.2, for different (asymptotically stable and unstable) values of eigenvalue ϕ. We note again that for $\phi < 0$ probability density $f_{Q_k}(q)$ of random variance Q_k has bounded support. Such support now *depends on parameters* σ^2 *and* ϕ; note that $\mathbb{P}\left[Q_k < \frac{g^2\sigma^2}{2|\phi|}\right] = 1$. In the case of unstable systems ($\phi > 0$) we have instead that the support of $f_{Q_k}(q)$ is *unbounded*; in fact, we have that $f_{Q_k}(q) > 0$ for all $q > 0$.

Again, interesting things happen as far as the moments of Q_k are concerned. Consider the following integral:

$$\mathbb{E}\left[Q_k^n\right] = \frac{\lambda g^2\sigma^2}{2\phi}\int_0^\infty \left(e^{2\phi t} - 1\right)^n e^{-\lambda t}\, dt = \frac{\lambda g^2\sigma^2}{2\phi}\int_0^\infty \sum_{m=0}^{n} \binom{n}{m} e^{(2m\phi-\lambda)t} dt\,;$$

it only converges if $\lambda > 2n\phi$, which is the condition for the existence (i.e. the finiteness) of the n-th moment of random variable Q_k. When $\lambda > 2n\phi$ we have that

$$\mathbb{E}\left[Q_k^n\right] = \frac{g^2\sigma^2}{2\phi}\sum_{m=0}^{n}\binom{n}{m}\frac{\lambda}{\lambda - 2m\phi}\,; \tag{2.15}$$

in particular, the mean and variance of Q_k are given by:

$$\mathbb{E}\left[Q_k\right] = \frac{g^2\sigma^2}{\lambda - 2\phi}\,, \qquad \mathrm{Var}\left[Q_k\right] = \frac{g^4\sigma^4}{(\lambda-2\phi)^2}\frac{\lambda}{\lambda - 4\phi}$$

and they exist if, respectively, $\lambda > 2\phi$ and $\lambda > 4\phi$. Note that both the mean and the covariance of Q_k (when they exist) are monotone increasing functions of continuous noise variance σ^2, monotone increasing functions of eigenvalue ϕ and monotone *decreasing* functions of measurement intensity λ.

Condition $\lambda > 2n\phi$ is verified for any $n \in \mathbb{N}$ in the case of asymptotically stable systems ($\phi < 0$), since $\lambda > 0$. If the continuous-time system is unstable ($\phi > 0$) then it will not have n-th order moments from a certain index n onwards. In particular if $\lambda < \phi$ then A_k does not even have a finite first-order moment, i.e. *its mean diverges to infinity*: this again happens when the intensity of the Poisson process (that is, the average number of measurements per second) is too small with respect to the degree of instability of the continuous-time system. Note that "infinite" variance Q_k means that no reliable estimation can be made on the amplitude of noise $w(t_k)$, which could easily take very high values; this would clearly have negative effects on the effectiveness of state estimation. We shall investigate these effects in the next Chapter.

The above results regarding the moments of r.v. Q_k are summarized by the following Corollary.

Corollary 2.5 *When $\phi < 0$ random variable Q_k has all moments, given by (2.15). When $\phi = 0$ we have that $Q_k \sim \mathcal{E}\left(\frac{\lambda}{g^2\sigma^2}\right)$ hence it obviously has all moments. Finally, when $\phi > 0$, let $n_Q = \max\{n : \lambda > 2n\phi\}$: r.v. Q_k has the n-th order moment if and only if $n \leq n_Q$; when this occurs, $\mathbb{E}[Q_k^n]$ is given by (2.15).*

Remark: Brownian motion. It is important to summarize the results we obtained in the case of random sampling of Brownian motion ($\phi = 0$), since it is such a significant subcase. As far as A_k is concerned, it is equal to 1 with probability 1 (the corresponding density is a Dirac delta function); on the other hand, Q_k is exponentially distributed with mean $g^2\sigma^2/\lambda$.

2.4 Conclusions

In this Chapter we have performed an analysis of system (2.1), obtained by sampling at the Poisson arrivals the original continuous-time system (1.1) we introduced in the previous Chapter.

After a discussion on the randomness the matrices that characterize this system, we formulated a Kalman Filter-based algorithm to estimate state x from the noisy measurements that are given by the second of equations (2.1) in correspondence of the Poisson arrivals. We also extended the estimation algorithm to instants that lie between two consecutive Poisson arrivals. We also proved a property of process $\{P_{k|k}\}_{k=0}^{\infty}$ (the sequence of estimation error covariance matrices) that will be of great use in the next chapter: namely, we proved that it is a *Markov* process.

Finally, we computed a full statistical description of random matrices A_k and Q_k for the one-dimensional case. We saw how such description dramatically depends on the sign of eigenvalue ϕ and on the ratio between such eigenvalue and the Poisson sampling intensity λ, which is proportional to the number of sensors in a sensor network in case this were the physical situation that system (1.1) represents (as we discussed in Chapter 1). In particular, when system (1.1) is unstable A_k and Q_k may tend to have very high values, especially in the case of low sampling intensity.

The next Chapter will be dedicated to calculating a statistical description of process $\{P_{k|k}\}_{k=0}^{\infty}$, which measures the *effectiveness* of the estimation algorithm we formulated in section 2.2. We shall again limit ourselves to the one-dimensional case, computing, however, its *complete* statistical description. We will perform calculations on such process in a similar way as we we did on A_k and Q_k, uncovering its dependency on eigenvalue ϕ and on sampling intensity λ.

Remark: prediction vs estimation. As we noted in section 2.2.1 it is necessary to wait until arrival t_{k+1} occurs before using the time update equations (2.6) and (2.7); this means that no *prediction* can really be performed at time t_k about time t_{k+1} as in the case of ordinary Kalman filtering. In fact at time t_{k+1}, i.e. when A_k and Q_k are available, one uses both the time update and the measurement update equations obtaining $\hat{x}_{k|k}$ and $P_{k|k}$, i.e. performing directly state *estimation*, whose effectiveness will be studied thoroughly in the next Chapter.

Chapter 3

Statistical Description of the State Estimator

We have seen how the estimation error variance process $\{P_{k|k}\}_{k=1}^{\infty}$ is *stochastic*. The main goal of this Chapter is providing a complete statistical description of this process, in function of the original continuous-time system's dynamics (represented by eigenvalue ϕ) and the Poisson sampling intensity λ. We will immediately see how the Markov nature of such process plays a key role.

The estimation error variance describes the effectiveness of the estimation algorithm we formulated in the previous Chapter. Therefore it will be useful, and interesting, to analyze in depth how different choices of the sampling intensity (which is proportional to the number of sensors in a network) influences this effectiveness. We will see how in some cases, namely for asymptotically stable systems ($\phi < 0$), such influence may be an unexpected one; however, we will provide a thorough explanation for this apparently counterintuitive behavior.

3.1 Introduction: exploiting the Markov property

In the previous Chapter we formulated Kalman Filter-based algorithm for estimating state x in correspondence of the Poisson arrivals of the measurement process. The corresponding equations are the following:

$$\begin{aligned}
\hat{x}_{k+1|k} &= A_k \hat{x}_{k|k} \\
P_{k+1|k} &= A_k P_{k|k} A_k^T + Q_k && (3.1) \\
\hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} + P_{k+1|k} C^T (C P_{k+1|k} C^T + R)^{-1} (y_{k+1} - C \hat{x}_{k+1|k}) \\
P_{k+1|k+1} &= P_{k+1|k} - P_{k+1|k} C^T (C P_{k+1|k} C^T + R)^{-1} C P_{k+1|k} \, . && (3.2)
\end{aligned}$$

we remind the reader that the first two of the above formulas are called *time update equations*, while the least two are named *measurement update equations*.

We already discussed the fact that the sequence of estimation error covariance matrices, i.e. $\{P_{k|k}\}_{k=1}^{\infty}$, is (contrarily to what happens for the ordinary Kalman Filter) a *stochastic process*; in fact it has the nice property of being a homogeneous Markov process, as we proved in section 2.2.2.

This has important consequences regarding the complete statistical description of the random process in question. Assume that the probability density (or distribution) of $x(0)$ is known and let $P_0 = \mathrm{Var}[x(0)]$ be the corresponding covariance matrix; define $P_{0|0} \triangleq P_0$. Consider probability distribution:[1]

$$\begin{aligned} &F^{(k+1)}(p_{k+1}, p_k, \dots, p_1, p_0) = \\ &\qquad = \mathbb{P}\big[\, P_{k+1|k+1} \le p_{k+1},\ P_{k|k} \le p_k,\ P_{k-1|k-1} \le p_{k-1}, \dots, P_{0|0} \le p_0 \,\big] \end{aligned}$$

and the corresponding probability density:[2]

$$\begin{aligned} &f^{(k+1)}(p_{k+1}, p_k, \dots, p_1, p_0) = \\ &\qquad = \frac{\partial^{k+2}}{\partial p_{k+1} \partial p_k \dots \partial p_1 \partial p_0}\ F^{(k+1)}(p_{k+1}, p_k, \dots, p_1, p_0)\,; \end{aligned}$$

thanks to the Markov property we may rewrite the above density as follows:

$$f^{(k+1)}(p_{k+1}, p_k, \dots, p_1, p_0) = f_{k+1|k}(p_{k+1}|p_k) \cdot f_{k|k-1}(p_k|p_{k-1}) \cdot \ldots \cdot f_{1|0}(p_1|p_0) \cdot f_0(p_0),$$

where $f_0(\cdot)$ is the probability density of covariance matrix p_0, which in general may be random as well (in case it were deterministic, $f_0(\cdot)$ would just take the form of a $n \times n$ dimensional delta function). Since $\{P_{k|k}\}_{k=1}^{\infty}$ is a *homogeneous* Markov process each of the above conditional densities $f_{j+1|j}(\,\cdot\,|\,\cdot\,)$, $j \in \{0, \dots, k\}$, does *not* depend on index j but only on the value of its arguments. So the joint density of random matrices $\{P_{j|j}\}_{j=0}^{k+1}$ takes the simpler form:

$$f^{(k+1)}(p_{k+1}, p_k, \dots, p_1, p_0) = f(p_{k+1}|p_k) \cdot f(p_k|p_{k-1}) \cdot \ldots \cdot f(p_1|p_0) \cdot f_0(p_0), \tag{3.3}$$

where

$$f(p|q) \triangleq \frac{\partial}{\partial p}\ \mathbb{P}\big[\, P_{j+1|j+1} \le p \,\big|\, P_{j|j} = q \,\big] \tag{3.4}$$

does not depend on index j but only on the value assumed by matrices p and q.

The rest of this Chapter is dedicated computing an analytical expression for (3.4), which is sufficient to yield complete statistical description of process $\{P_{k|k}\}_{k=1}^{\infty}$, i.e. the sequence of (random) error estimation covariance matrices. We shall limit ourselves to the one-dimensional case, for which such calculations can be carried

[1] As in the previous Chapter, with the expression $P_{k+1|k+1} \le p_{k+1}$ we intend that every element of matrix $P_{k+1|k+1}$ is less than or equal to the corresponding element of matrix p_{k+1}.

[2] Here symbol $\partial/\partial p_{k+1}$ indicates differentiation with respect to each single element of matrix p_{k+1}. In the one-dimensional case ($m = n = p = 1$), which we will analyze in detail later in the Chapter, it just indicates ordinary partial differentiation with respect to variable p_{k+1}.

out in a relatively simple manner. We will see how the expression of conditional density (3.4) is very closely related on the stability of the original continuous time system (1.1) and on the intensity of the Poisson sampling process, sometimes in quite a surprising manner.

3.2 Computation of conditional probability density

From now on we shall consider the one-dimensional case ($m = n = p = 1$), as we did in the previous Chapter. For this purpose, we will use the usual scalar quantities: $F = \phi$, $G = g$, $S = \sigma^2$. In the 1-D case ϕ obviously represents the only eigenvalue of the continuous-time system matrix F. In fact R and C are scalar too; however, we will not change symbols for these.

In the scalar case, equation (3.2) may be rewritten as follows:

$$P_{k+1|k+1} = P_{k+1|k} - \frac{C^2 P_{k+1|k}^2}{C^2 P_{k+1|k} + R} = \frac{R\, P_{k+1|k}}{C^2 P_{k+1|k} + R} = \frac{R}{C^2} \frac{P_{k+1|k}}{P_{k+1|k} + \frac{R}{C^2}} \;;$$

note, in particular, that $\mathbb{P}\big[P_{k+1|k+1} \le R/C^2\big] = 1$, for all k; in other words the estimation error variance is bounded by R/C^2 (which is basically determined by R, i.e. the variance of noise z in the measurement equation in (1.1)), independently of the system's dynamics and the Poisson sampling process. Substituting (3.1) into the above equation yields:

$$P_{k+1|k+1} = \frac{R}{C^2} \frac{A_k^2 P_{k|k} + Q_k}{A_k^2 P_{k|k} + Q_k + \frac{R}{C^2}} \;; \tag{3.5}$$

which is equation (2.10) written in scalar form.

We remember from Chapter 2 that both A_k and Q_k depend on random waiting time T_k, as in (2.11) and in (2.14):

$$A_k = e^{\phi T_k}, \;\; \forall \phi \in \mathbb{R} \quad \text{and} \quad Q_k = \begin{cases} g^2\sigma^2 T_k & \text{for } \phi = 0\,, \\ \frac{g^2\sigma^2}{2\phi}\big(e^{2\phi T_k} - 1\big) & \text{for } \phi \neq 0\,. \end{cases}$$

Substituting such expressions into (3.5) yields:

$$P_{k+1|k+1} = \begin{cases} \dfrac{R}{C^2} \dfrac{P_{k|k} + g^2\sigma^2 T_k}{P_{k|k} + g^2\sigma^2 T_k + \frac{R}{C^2}} & \text{for } \phi = 0\,, \\[2ex] \dfrac{R}{C^2} \dfrac{\Big(P_{k|k} + \frac{g^2\sigma^2}{2\phi}\Big) e^{2\phi T_k} - \frac{g^2\sigma^2}{2\phi}}{\Big(P_{k|k} + \frac{g^2\sigma^2}{2\phi}\Big) e^{2\phi T_k} - \frac{g^2\sigma^2}{2\phi} + \frac{R}{C^2}} & \text{for } \phi \neq 0\,, \end{cases} \tag{3.6}$$

making the dependency on T_k explicit. Note that the expression that is valid for $\phi = 0$ may be obtained from the other one just by continuity in ϕ, i.e. by taking the

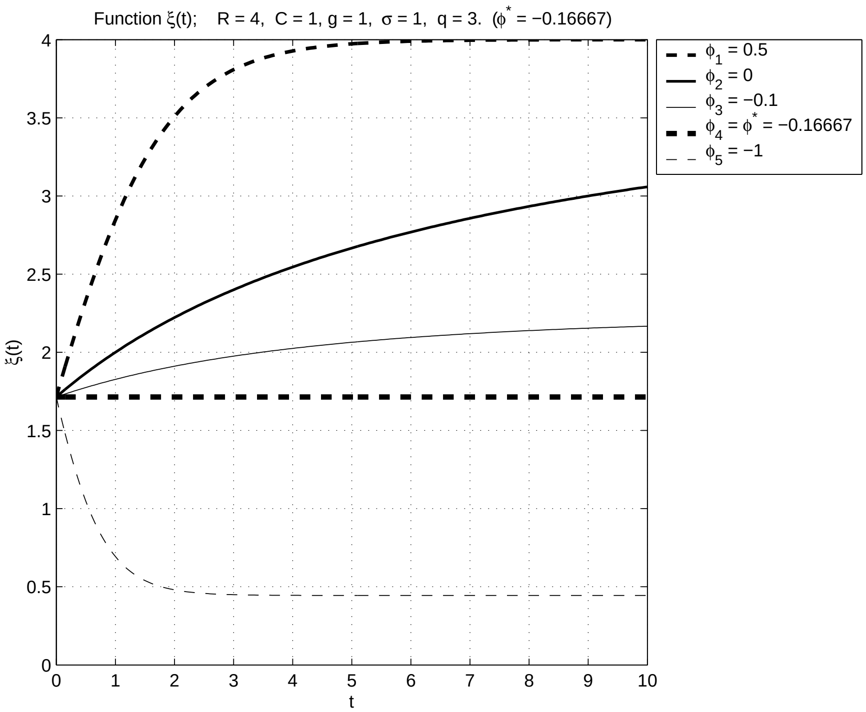


Figure 3.1: Function $\xi(t)$, plotted for several values of eigenvalue ϕ. Note that when $\phi = \phi^* \triangleq -\frac{g^2\sigma^2}{2q}$ we have that $\xi(t) \equiv \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}$, for all t.

limit for $\phi \to 0$. It is thanks to equations (3.6) and the knowledge of the probability distribution of T_k that we will be able to calculate the probability density of $P_{k+1|k+1}$ conditioned on event $\{P_{k|k} = q\}$, i.e. conditional density (3.4).

For this purpose, consider the following function of $t \geq 0$:

$$\xi(t) = \begin{cases} \dfrac{R}{C^2}\dfrac{q + g^2\sigma^2 t}{q + g^2\sigma^2 t + \frac{R}{C^2}} & \text{for } \phi = 0\,, \\ \\ \dfrac{R}{C^2}\dfrac{\left(q + \frac{g^2\sigma^2}{2\phi}\right)e^{2\phi t} - \frac{g^2\sigma^2}{2\phi}}{\left(q + \frac{g^2\sigma^2}{2\phi}\right)e^{2\phi t} - \frac{g^2\sigma^2}{2\phi} + \frac{R}{C^2}} & \text{for } \phi \neq 0\,, \end{cases} \tag{3.7}$$

so that we may write: $P_{k+1|k+1} = \xi(T_k)$. Function $\xi(t)$ is plotted, for several values of eigenvalue ϕ, in Figure 3.1; we only consider nonnegative values of argument t since r.v. T_k takes only nonnegative values. Note also that it only makes sense to choose values of q in interval $\left[0, R/C^2\right]$ since q has the meaning of error estimation variance $P_{k|k}$ in (3.6) and we have seen above that $\mathbb{P}\left[P_{k|k} \leq R/C^2\right] = 1$, for all $k \in \mathbb{N}$.

Function ξ has the following properties, which we will use for calculating the expression of conditional density (3.4) for different values of ϕ and λ:

1. For any $\phi \in \mathbb{R}$, we have that $\xi(0) = \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}$.

2. There exists a particular eigenvalue, namely $\phi^* \triangleq -\frac{g^2\sigma^2}{2q}$ (which is always *negative*) such that if $\phi = \phi^*$ then $\xi(t) \equiv \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}$, for all values of t.

3. $\xi(t)$ is monotone increasing for $\phi > \phi^*$ and monotone decreasing for $\phi < \phi^*$; in fact one can prove that $\xi'(t) > 0$ for $\phi > \phi^*$, $\xi'(t) < 0$ for $\phi < \phi^*$, and $\xi'(t) \equiv 0$ for $\phi = \phi^*$. Note that such derivative is *positive* for some *negative* values of ϕ.

4. For $t \to \infty$, we have the following limit:
$$\xi_\infty \triangleq \lim_{t\to\infty} \xi(t) = \begin{cases} \frac{R}{C^2} & \text{for } \phi \geq 0\,, \\ \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}} & \text{for } \phi < 0\,; \end{cases}$$
note that such limit *depends on* ϕ; for unstable systems and for $\phi = 0$ it is equal to R/C^2, whereas for asymptotically stable systems ξ_∞ is strictly less than R/C^2 and depends on the *particular* value ϕ takes. The limit above is a monotone non decreasing function of the eigenvalue, and it reaches its maximum for the first time for $\phi = 0$; for $\phi \to -\infty$ then $\xi_\infty \to 0$, but $\xi_\infty \neq 0$ for all values of ϕ.

We now have the necessary tools to prove the main result of this Chapter. We shall list both distributions and densities in various cases: in fact distributions
$$F(p|q) \triangleq \mathbb{P}\big[\, P_{j+1|j+1} \leq p \,\big|\, P_{j|j} = q \,\big]$$
are easier to compute, but we believe that the shape of densities are easier to interpret when plotted on a graph. One should not spend too much time analyzing the analytical expression of the listed functions, except maybe for the *support* of the probability density functions: an extensive discussion will follow in the next section, based on the plots of densities in several significant cases.

Theorem 3.1 *Let* $\phi^* = -\frac{g^2\sigma^2}{2}\frac{1}{q}$, *with* $q \in \big(0, \frac{R}{C^2}\big]$. *Then we will have that random variable* $P_{k+1|k+1}|\{P_{k|k} = q\}$ *has the following probability distributions and densities, depending on the value assumed by eigenvalue* ϕ*:*

- **Type I**, *for* $\phi < \phi^*$*:*

$$F(p|q) = \begin{cases} \left[\dfrac{\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)} \right]^{\frac{\lambda}{2\phi}} & \textit{for } \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}} \leq p < \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\,, \\ 0 & \textit{for } p < \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}\,, \\ 1 & \textit{for } p \geq \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\,; \end{cases} \tag{3.8}$$

$$f(p|q) = -\frac{\lambda}{2\phi}\frac{R^2}{C^4}\left(q+\frac{g^2\sigma^2}{2\phi}\right)\frac{\left[\left(\frac{R}{C^2}-p\right)\left(q+\frac{g^2\sigma^2}{2\phi}\right)\right]^{\frac{\lambda}{2\phi}-1}}{\left[\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)\right]^{\frac{\lambda}{2\phi}+1}} \cdot \cdot 1\left(\frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2-2\phi\frac{R}{C^2}} < p < \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\right); \tag{3.9}$$

- **Type II**, *for $\phi = \phi^*$:*

$$F(p|q) = \begin{cases} 0 & \textit{for } p < \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}, \\ 1 & \textit{for } p \geq \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}; \end{cases} \tag{3.10}$$

$$f(p|q) = \delta\left(p - \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\right); \tag{3.11}$$

- **Type III**, *for $\phi^* < \phi < 0$:*

$$F(p|q) = \begin{cases} 1-\left[\frac{\left(\frac{R}{C^2}-p\right)\left(q+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} & \textit{for } \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}} \leq p < \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2-2\phi\frac{R}{C^2}}, \\ 0 & \textit{for } p < \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}, \\ 1 & \textit{for } p \geq \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2-2\phi\frac{R}{C^2}}; \end{cases} \tag{3.12}$$

$$f(p|q) = \frac{\lambda}{2\varphi}\frac{R^2}{C^4}\left(q+\frac{g^2\sigma^2}{2\varphi}\right)\frac{\left[\left(\frac{R}{C^2}-p\right)\left(q+\frac{g^2\sigma^2}{2\varphi}\right)\right]^{\frac{\lambda}{2\varphi}-1}}{\left[\frac{g^2\sigma^2}{2\varphi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\varphi}\right)\right]^{\frac{\lambda}{2\varphi}+1}} \cdot \cdot 1\left(\frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}} < p < \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2-2\varphi\frac{R}{C^2}}\right); \tag{3.13}$$

- **Type IV**, *for $\phi = 0$ (Brownian motion):*

$$F(p|q) = \begin{cases} 1-\exp\left[-\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p}-q\right)\right] & \textit{for } \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}} \leq p < \frac{R}{C^2}, \\ 0 & \textit{for } p < \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}, \\ 1 & \textit{for } p \geq \frac{R}{C^2}; \end{cases} \tag{3.14}$$

$$
\begin{aligned}
f(p|q) &= \frac{\lambda}{2\varphi}\frac{R^2}{C^4}\left(\frac{R}{C^2}-p\right)^{-2}\cdot\exp\left[-\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p}-q\right)\right]\cdot\\
&\quad\cdot 1\left(\frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}<p<\frac{R}{C^2}\right); \qquad (3.15)
\end{aligned}
$$

- **Type V**, *for* $\phi > 0$*:*

$$
F(p|q) = \begin{cases} 1-\left[\dfrac{\left(\frac{R}{C^2}-p\right)\left(q+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} & \text{for } \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\le p<\frac{R}{C^2},\\ 0 & \text{for } p<\frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}},\\ 1 & \text{for } p\ge\frac{R}{C^2}; \end{cases} \qquad (3.16)
$$

$$
\begin{aligned}
f(p|q) &= \frac{\lambda}{2\varphi}\frac{R^2}{C^4}\left(q+\frac{g^2\sigma^2}{2\varphi}\right)\frac{\left[\left(\frac{R}{C^2}-p\right)\left(q+\frac{g^2\sigma^2}{2\varphi}\right)\right]^{\frac{\lambda}{2\varphi}-1}}{\left[\frac{g^2\sigma^2}{2\varphi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\varphi}\right)\right]^{\frac{\lambda}{2\varphi}+1}}\cdot\\
&\quad\cdot 1\left(\frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}<p<\frac{R}{C^2}\right). \qquad (3.17)
\end{aligned}
$$

In particular, the Type IV distribution (which is relative to random sampling of Brownian motion) may be obtained as the limit for $\phi\to 0^-$*, or* $\phi\to 0^+$*, of the Type III, or (respectively) Type V, distributions.*

It will be useful to introduce quantities:

$$
s_1 \triangleq \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2-2\phi\frac{R}{C^2}}, \qquad s_2 \triangleq \frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}; \qquad (3.18)
$$

they are, in the above cases (I to V), the extremes of the support of probability density $f(p|q)$ (see Figure 3.2). We shall return on this point in the next section.

Proof of Theorem 3.1. First of all note that, in general, we cannot apply Theorem 2.1 in order to calculate the above densities: function $\xi(t)$ has a negative singularity[3] if $\phi^* < \phi < \frac{g^2\sigma^2}{2}\frac{C^2}{R}$. In fact, it will be simpler to calculate probability

[3] Consider the case $\phi \neq 0$. The denominator of $\xi(t)$ is equal to zero when $t = t_0$, where

$$
t_0 = \frac{1}{2\phi}\log\frac{g^2\sigma^2-2\phi\frac{R}{C^2}}{g^2\sigma^2+2\phi q}
$$

when the logarithm exists. The argument of the logarithm is positive when its numerator and denominator have the same sign, which happens for $\phi\in\left(-\frac{g^2\sigma^2}{2}\frac{1}{q},\frac{g^2\sigma^2}{2}\frac{C^2}{R}\right)=\left(\phi^*,\frac{g^2\sigma^2}{2}\frac{C^2}{R}\right)$; if ϕ is equal to any of the two extremes of such interval one can easily check that no singularity occurs. The study of the sign of the above function yields that the singularity, when it occurs, is always negative (which should have already been clear from Figure 3.1).

distributions rather than densities just by computing:

$$F(p|q) = \mathbb{P}\big[P_{k+1|k+1} \le p \,\big|\, P_{k|k} = q\big] = \mathbb{P}\big[\xi(T_k) \le p \,\big|\, P_{k|k} = q\big],$$

where the dependence of ξ on q is defined by formulas (3.7). The corresponding densities are then obtained by differentiation.

The case $\phi < \phi^*$ is the only one for which $\xi(t)$, $t \ge 0$, is a monotone decreasing function (see Figure 3.1). If $p \in (s_1, s_2]$ then

$$\xi^{-1}(p) = \left\{ \frac{1}{2\phi} \log \frac{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}{\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right)} \right\}; \tag{3.19}$$

we shall call $t^\star$ the only element of set $\xi^{-1}(p)$ in this case; otherwise, if $p \notin (s_1, s_2]$, then $\xi^{-1}(p) = \emptyset$ (Figure 3.1 should clarify the situation). Therefore

$$\begin{aligned} F(p|q) &\triangleq \mathbb{P}[\, P_{k+1|k+1} \le p \,|\, P_{k|k} = q \,] = \\ &= \begin{cases} \mathbb{P}[\, T_k \ge t^\star \,|\, P_{k|k} = q \,] & \text{for } s_1 \le p < s_2\,, \\ 1 & \text{for } p \ge s_2\,, \\ 0 & \text{for } p < s_1\,. \end{cases} \end{aligned}$$

Random variable T_k is independent of $P_{k|k}$ (in fact $P_{k|k}$ is determined by $\{T_j\}_{j=0}^{k-1}$ and $v(t)$, $t \in (0, t_k)$, which are independent of T_k) hence we have that

$$\begin{aligned} \mathbb{P}[\, T_k \ge t^\star \,|\, P_{k|k} = q \,] &= \mathbb{P}[\, T_k \ge t^\star \,] = 1 - F_{T_k}(t^\star) = \\ &= 1 - \big(1 - e^{-\lambda t^\star}\big) \cdot 1(t^\star > 0) = e^{-\lambda t^\star}, \end{aligned}$$

since $T_k \sim \mathcal{E}(\lambda)$ and $t^\star > 0$. Substituting the value of $t^\star$ given in (3.19) into $e^{-\lambda t^\star}$ finally yields (3.8); differentiation then gives probability density (3.9). We call these, respectively, Type I probability distribution and density.

The case $\phi = \phi^*$ is obvious since ξ is identically equal to $s_1 = s_2 = \frac{R}{C^2} q \big(q + \frac{R}{C^2}\big)^{-1}$ for any value of t; hence r.v. $P_{k+1|k+1} | \{P_{k|k} = q\}$ is equal to s_2 with probability one, and the corresponding probability density is a Dirac delta function centered in s_2, which we call Type II density.

If $\phi^* < \phi < 0$ then the situation different from the first cases since ξ is monotone *increasing* (see Figure 3.1). Therefore we have that $\xi^{-1}(p)$ is given by (3.19) when $p \in [s_2, s_1)$; note that the extremes of such interval are reversed with respect to the first case ($\phi < \phi^*$). Also, $\xi^{-1}(p) = \emptyset$ when $p \notin [s_2, s_1)$. So we shall have

$$\begin{aligned} F(p|q) &= \mathbb{P}[\, P_{k+1|k+1} \le p \,|\, P_{k|k} = q \,] = \\ &= \begin{cases} \mathbb{P}[\, 0 \le T_k \le t^\star \,|\, P_{k|k} = q \,] & \text{for } s_2 \le p < s_1\,, \\ 1 & \text{for } p \ge s_1\,, \\ 0 & \text{for } p < s_2\,, \end{cases} \end{aligned}$$

where

$$\begin{aligned}\mathbb{P}[\,0 \le T_k \le t^\star \,|\, P_{k|k} = q\,] &= \mathbb{P}[\,0 \le T_k \le t^\star\,] \;=\; F_{T_k}(t^\star) \;= \\ &= \left(1 - e^{-\lambda t^\star}\right) \cdot 1(t^\star > 0) \;=\; 1 - e^{-\lambda t^\star}.\end{aligned}$$

Type III distribution (3.12) and density (3.13) are obtained, respectively, by substituting the expression for $t^\star$ in the above equation and by successive differentiation.

The important case of Brownian motion ($\phi = 0$) is peculiar since $\xi(t)$ asymptotically tends to R/C^2 as $t \to \infty$ for the "first time": as we noted in point 4 on page 21, ξ_∞ is an increasing function of ϕ and it reaches R/C^2 at $\phi = 0$, and "stays there" for higher values of ϕ. We will have that if $p \in \left[s_2, \frac{R}{C^2}\right)$ then

$$\xi^{-1}(p) = \left\{ \frac{1}{g^2\sigma^2} \left(\frac{R}{C^2} \frac{p}{\frac{R}{C^2} - p} - q \right) \right\} \tag{3.20}$$

and as before we shall call $t^\star$ the only element of $\xi^{-1}(p)$; if $p \notin \left[s_2, \frac{R}{C^2}\right)$ then $\xi^{-1}(p) = \emptyset$. So we shall have

$$\begin{aligned}F(p|q) &= \mathbb{P}\big[\,P_{k+1|k+1} \le p \,\big|\, P_{k|k} = q\,\big] \;= \\ &= \begin{cases} \mathbb{P}[\,0 \le T_k \le t^\star \,|\, P_{k|k} = q\,] & \text{for } s_2 \le p < \frac{R}{C^2}\,, \\ 1 & \text{for } p \ge \frac{R}{C^2}\,, \\ 0 & \text{for } p < s_2\,, \end{cases}\end{aligned}$$

where

$$\mathbb{P}[\,0 \le T_k \le t^\star \,|\, P_{k|k} = q\,] \;=\; 1 - e^{-\lambda t^\star}.$$

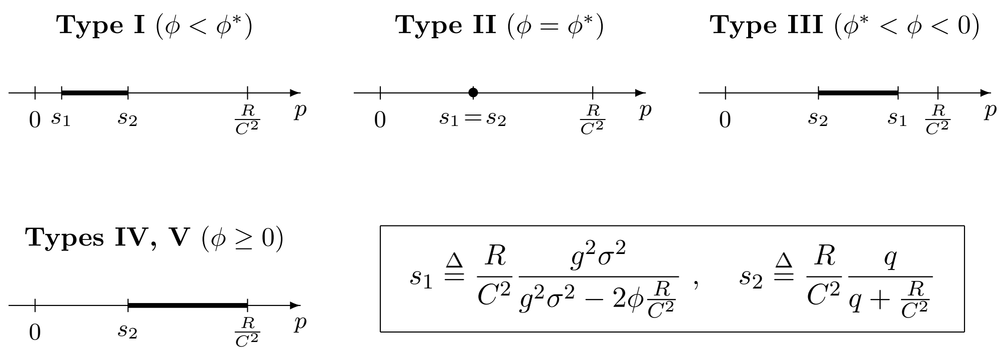
Now just substitute the expression for $t^\star$ given by (3.20) and differentiate to get Type IV distribution (3.14) and density (3.15).

Finally, case $\phi > 0$ is similar the case $\phi^* < \phi < 0$ except for the fact that $\xi_\infty = R/C^2$, so that $\xi^{-1}(p)$ is given by (3.19) for $p \in \left[s_2, \frac{R}{C^2}\right)$, and $\xi^{-1}(p) = \emptyset$ for $p \notin \left[s_2, \frac{R}{C^2}\right)$. The derivation of Type V distribution (3.16) and density (3.17) is at this point just a matter of repeating analogous calculations. $\square$

3.3 Analysis of conditional probability density

The previous section was dedicated to computing an analytical expression of conditional probability density (3.4). We saw, in particular, that it closely depends on the dynamics of the original continuous-time system, i.e. on eigenvalue ϕ.

We shall now analyze the density expressions we computed by first studying how their support depends on eigenvalue ϕ and conditioning value q; we shall subsequently study the shape of such densities, and analyze the influence that sampling intensity λ has on them.



$$s_1 \triangleq \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}\ , \qquad s_2 \triangleq \frac{R}{C^2}\frac{q}{q + \frac{R}{C^2}}$$

Figure 3.2: Support of conditional probability density $f(p|q)$. Note that such support does not depend on sampling intensity λ.

3.3.1 Density support

As we noted at the beginning of section 3.2 estimation error variance $P_{k|k}$, $\forall k \in \mathbb{N}$, is bounded by R/C^2. In fact this follows immediately from the form of the second of equations (1.1) and (2.1):

$$y(t_k) = Cx(t_k) + z(t_k),$$

which implies that the estimation error is in the worst case determined by noise $z(t_k)$, whose covariance matrix is R; in general we can do better than that since we are also using past information.

The support of conditional density $f(p|q)$ is shown in Figure 3.2 in the five cases of Types I to V. Note the important role played by s_1 and s_2, defied in (3.18): in all five cases the support is an *interval*, whose extremes are s_1 and s_2 in the first three cases, s_2 and R/C^2 in the last two.

Note how extreme s_1 is basically determined by eigenvalue ϕ, i.e. by the dynamics of continuous-time system (1.1), whereas s_2 is determined by q, i.e. by the *previous* value assumed by estimation error variance $P_{k|k}$. We should also remark the fact that sampling intensity λ has no influence on the *support* of conditional density $f(p|q)$, but only on its *shape* within the support, as we shall see in the next subsection.

Extreme s_1, which we only consider for negative values of ϕ (asymptotically stable systems), defined by the first of (3.18), is an *increasing* function of eigenvalue ϕ; this means that the more a stable system approaches instability (as ϕ tends zero from below) the higher will be the possible values of $P_{k+1|k+1}$, i.e. the *harder* it will be to estimate state x. In particular for Types IV ($\phi = 0$) and V ($\phi > 0$) extreme s_1 is substituted by R/C^2, which is the highest value estimation error variance can possibly assume.

Extreme s_2 is on the other hand determined by q, that is the value previously assumed by $P_{k|k}$; in practice such value summarizes previous history, i.e. it represents

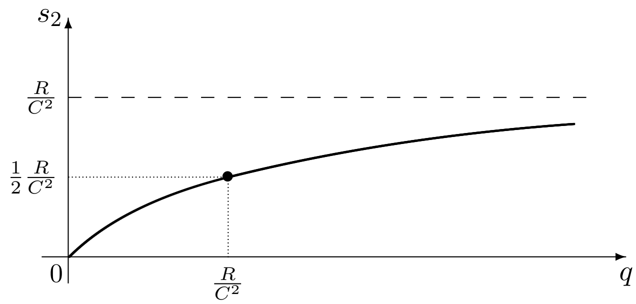


Figure 3.3: Dependence of s_2 on q.

a measure of the reliability of previous estimates. The dependence of s_2 on q is shown in Figure 3.3: since q can only assume values in interval[4] $[0, R/C^2]$, s_2 will always be lower than (or equal to) $\frac{1}{2}\frac{R}{C^2}$. Note that s_2 is an increasing function of q, which means that if the previous estimates was not good enough (i.e. the value of q is high) then the next estimation error variance will tend to assume high values, since we won't be able to use reliable past information.

Type I distributions represent the only case when we have that $s_1 < s_2 \leq \frac{1}{2}\frac{R}{C^2}$, which implies that the error variance will assume relatively low values. This happens when $\phi < \phi^*$ (which is negative) i.e. when the continuous time system is asymptotically stable with a relatively high degree of stability, since ϕ is negative but its absolute value is above a certain threshold. Hence, it is easier to perform state estimation for stable systems.

Type II distributions are somehow "singular", since the corresponding densities are Dirac delta functions centered in $s_1 = s_2$; it has to be noted that for a fixed system (i.e. for a fixed negative ϕ) they occur *with probability zero*, since they only occur when $P_{k|k}$ is exactly equal to $q = -\frac{g^2\sigma^2}{2\phi}$ (that corresponds to condition $\phi = \phi^*$) which is an event of probability zero since $P_{k|k}$ is a *continuous* random variable.

When $\phi^* < \phi < 0$ we have Type III densities and distributions, where $s_2 < s_1$; for increasing values of ϕ we have that s_1 "approaches" R/C^2, i.e. higher values of the error variance are allowed.

In the case $\phi = 0$, which corresponds to sampling Brownian motion, extreme s_1 finally "merges" with R/C^2, the highest possible value for the error variance: this corresponds to what we called Type IV distributions.

Finally, when $\phi > 0$ we have that the support of Type V densities is always given by interval $[s_2, R/C^2]$, independently of the (positive) value of ϕ. Note, however, that eigenvalue ϕ will still influence the shape of density function $f(p|q)$.

---

[4]With the only possible exception of $P_{0|0} \triangleq \mathrm{Var}[x(0)]$, which can be arbitrary.

Remark. Note that for a given one-dimensional system eigenvalue ϕ is fixed, which implies that extreme s_1 fixed. On the other hand $P_{k|k}$ varies in time, hence the value q assumes varies as well: this means that $\phi^* = -\frac{g^2\sigma^2}{2}\frac{1}{q}$ is variable in time too. Therefore if $\phi < 0$ we may have Types I, II or III at different times. Assume for instance that $P_{k|k} = q$, such that $\phi < \phi^*$ (which means that q is relatively high) so that we have a Type I density: then $P_{k+1|k+1}$ might assume a low enough value p so that $\phi > -\frac{g^2\sigma^2}{2}\frac{1}{p}$, which is the "new" value of ϕ^* (in fact p plays the role of q in the subsequent step: see section 3.1); so distribution

$$F(p_{k+2|k+2}|p_{k+1|k+1}) = \mathbb{P}\big[\, P_{k+2|k+2} \leq p_{k+2|k+2} \,\big|\, P_{k+1|k+1} = p_{k+1|k+1} \,\big],$$

with $p_{k+1|k+1} = p$, will be of Type III. On the other hand when $\phi = 0$ we will always have Type IV distributions and when $\phi > 0$ we will always have Type V distributions. However, the position of extreme s_2 will still vary in time in a random way.

3.3.2 Density plots in significant cases

We have plotted conditional density $f(p|q)$ in all five cases. We will now illustrate such plots, commenting how they are influenced by different choices of the system's eigenvalue ϕ and sampling intensity λ. The following parameters are common to all plots: $R = 4$, $C = 1$, $g = 1$, $\sigma^2 = 1$, $q = 3$, so that

$$\frac{R}{C^2} = 4 \qquad \text{and} \qquad \phi^* = -\frac{g^2\sigma^2}{2}\frac{1}{q} \simeq -0.16667\,;$$

the particular values of ϕ and λ are reported on each graph.

Type I densities ($\phi < \phi^*$). They are shown in Figure 3.4 (for a fixed sampling intensity $\lambda = 2$ and different values of $\phi < \phi^*$) and in Figure 3.5 (for a fixed eigenvalue $\phi = -1 < \phi^*$ and different sampling intensities).

In the first Figure higher values of ϕ tend to "squeeze" the probability density towards s_2 (which is determined by q) since s_1 is an increasing function of ϕ, as we discussed in the previous subsection.

In the second Figure the support of density $f(p|q)$ is fixed since we only vary λ, which has no influence on s_1 and s_2 (see (3.18)) but only affects the *shape* of the curve. We assist to an apparent *paradox*, which only occurs for Type I densities: in fact *higher values of* λ seem to shift the area below the curve *to the right*, i.e. higher sampling rates tend to increase the estimation error variance! How is this possible? The explanation is the following: when $\phi < \phi^*$ the continuous-time dynamics are relatively fast, meaning that state $x(t)$ quickly converges to zero; this implies that it is somehow convenient to "wait" a long time to get a new measurement (i.e. it is convenient to have lower sampling rates) since at that time state x will quite likely be very close to zero and it will be easier to formulate a correct state estimate.

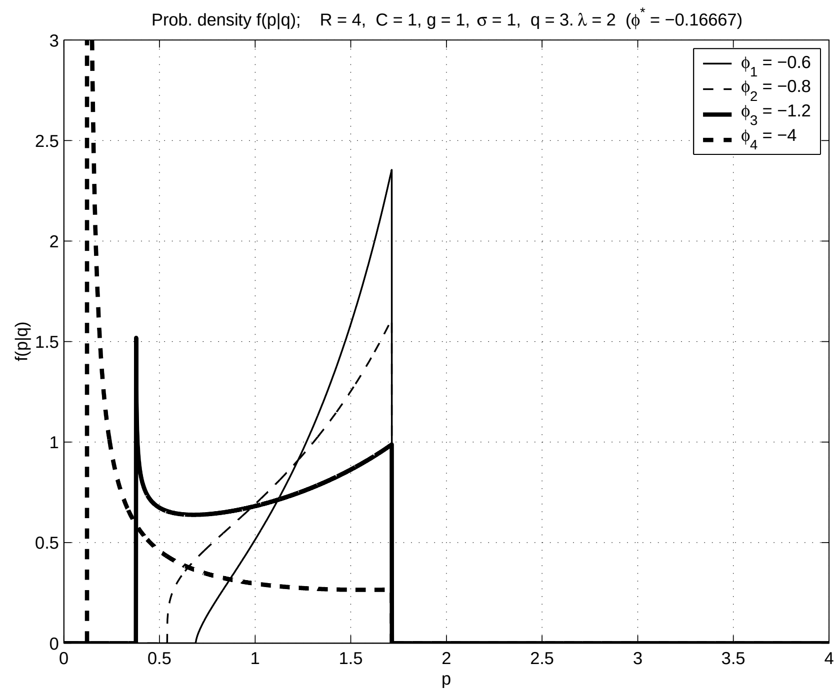


Figure 3.4: Type I densities, for a fixed value of λ.

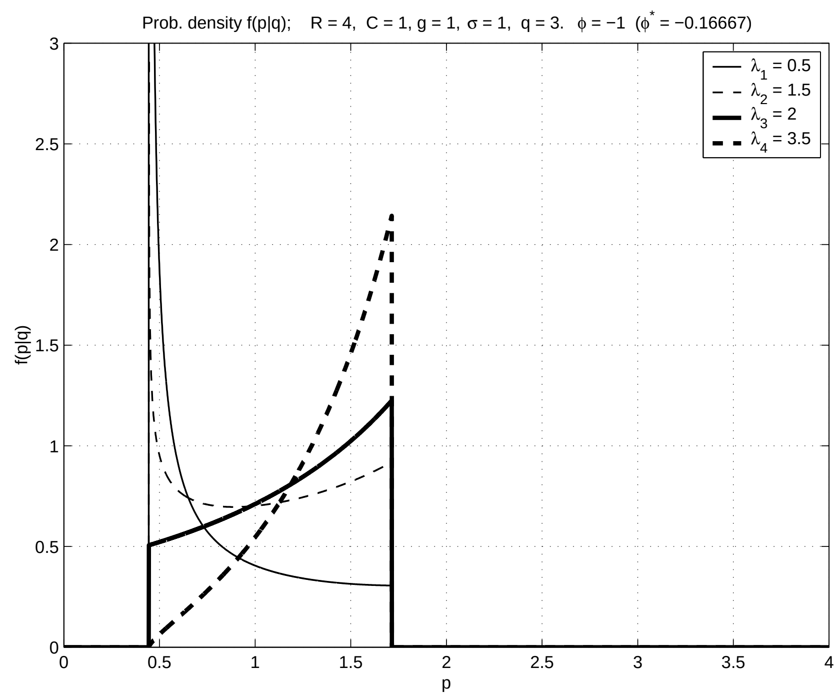


Figure 3.5: Type I densities, for a fixed value of ϕ (less than ϕ^*).

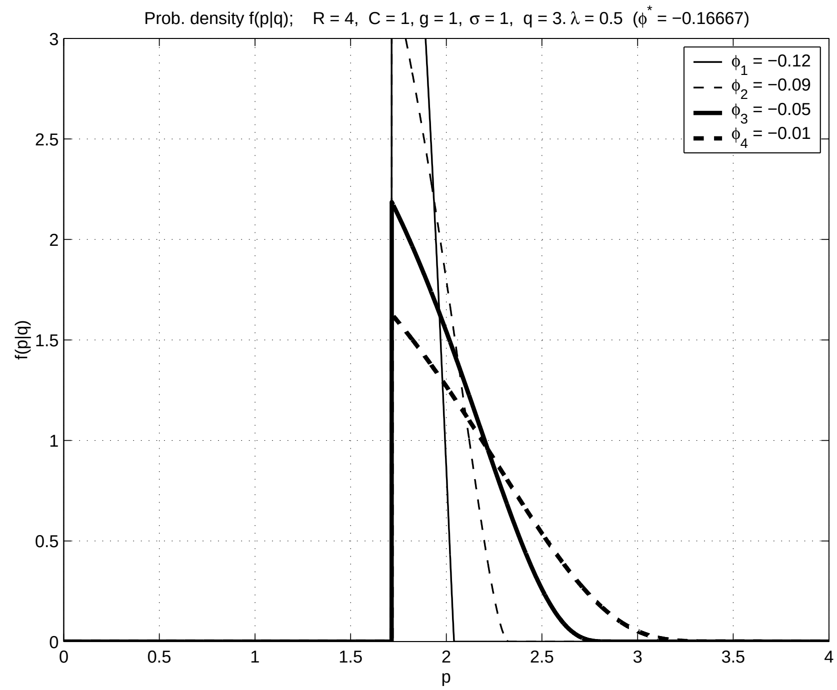


Figure 3.6: Type III densities, for a fixed value of λ.

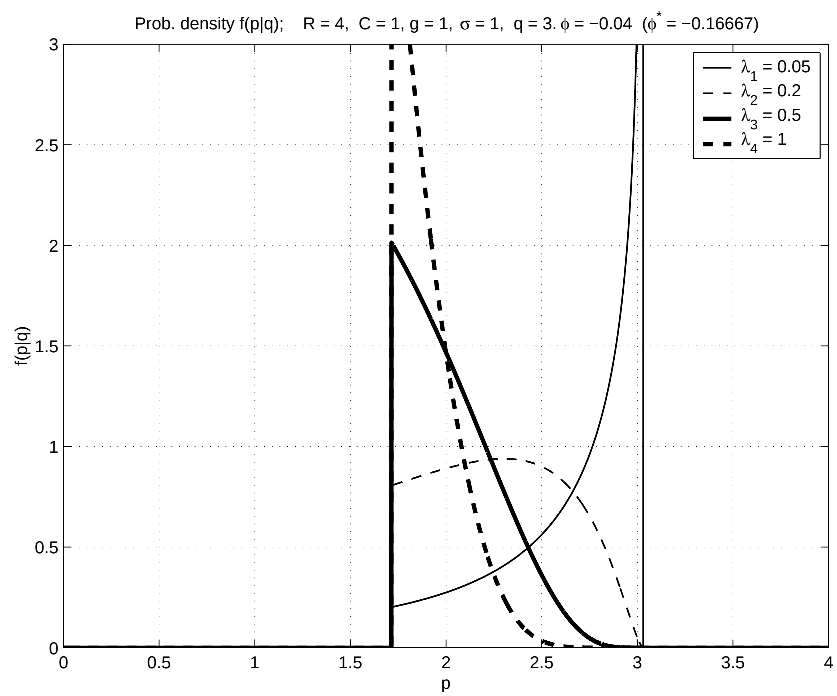


Figure 3.7: Type III densities, for a fixed value of ϕ.

Condition $\phi < \phi^*$ implies that $q > \frac{g^2\sigma^2}{2|\phi|}$, so we may equivalently interpret the situation saying that the prior knowledge on the state is little (i.e. q is relatively high), and since we cannot rely on it it is convenient to wait until $x(t)$ approaches zero before performing state estimation.

Note that, given an eigenvalue $\phi < 0$, if a Type I density occurs we will have with probability one that the value p that $P_{k+1|k+1}$ assumes will be lower than the value q assumed by $P_{k|k}$, since $p \leq s_2 < q$; so in the next step we will have that the "new" value of s_2 (i.e. the one relative to the support of the density of r.v. $P_{k+2|k+2}|\{P_{k+1|k+1} = p\}$) will be lower than the previous one: in fact it may happen that s_2 becomes smaller than s_1 (which is fixed, since it only depends on ϕ), and in that case we will have a Type III density. We shall prove rigorously in the next Chapter that, given a negative eigenvalue ϕ and a sampling intensity λ, if a Type I density $f(p_{k+1}|p_k)$ occurs in formula (3.3) then it will sooner or later "turn" into a Type III density and stick to that type, i.e. with probability one there exists a index $j > k$ such that all densities $f(p_{\ell+1}|p_\ell)$, $\ell \geq j$, are of Type III; such transition occurs in a random time whose mean is finite.[5] We will see below that the apparently paradoxical behavior of Type I densities does not occur for Type III densities: in fact increasing λ will *decrease* the estimation error variance, i.e. it will shift to the *left* the area below the graph of $f(p|q)$ (see Figure 3.7).

Type II densities ($\phi = \phi^*$). We did not plot Type II probability densities since they are simple Dirac delta functions, centered in $s_2 = \frac{R}{C^2} q\big(q + \frac{R}{C^2}\big)^{-1}$, *independently* of sampling intensity λ.

Type III densities ($\phi^* < \phi < 0$). These are plotted in Figure 3.6 (for a fixed sampling intensity $\lambda = 0.5$ and different values of ϕ, with $\phi^* < \phi < 0$) and in Figure 3.7 (for a fixed eigenvalue $\phi = -0.04 \in (\phi^*, 0)$ and different sampling intensities).

In the first Figure higher values of ϕ tend to expand the probability density towards $R/C^2 = 4$ by increasing the value of s_1 (see Figure 3.18), whereas s_2 is unchanged by ϕ. In other words, the more the system moves towards instability, the harder it is to estimate its state.

On the other hand, if we fix the value of $\phi \in (\phi^*, 0)$ and choose different sampling intensities λ then we have the situation depicted in Figure 3.7, which does not show the paradoxical behavior that is typical of Type I densities, thus being closer to what one would intuitively expect: higher sampling intensities reduce the estimation error variance by shifting the area below the graph of $f(p|q)$ to the left, i.e. towards lower values. Another way to interpret this fact is the following: $\phi^* < \phi < 0$ is equivalent to condition

$$0 < q < \frac{g^2\sigma^2}{2|\phi|} \,,$$

[5] Intuitively, this is the time that is necessary for state $x(t)$ to get close enough to zero and to start floating around it driven by noise $v(t)$.

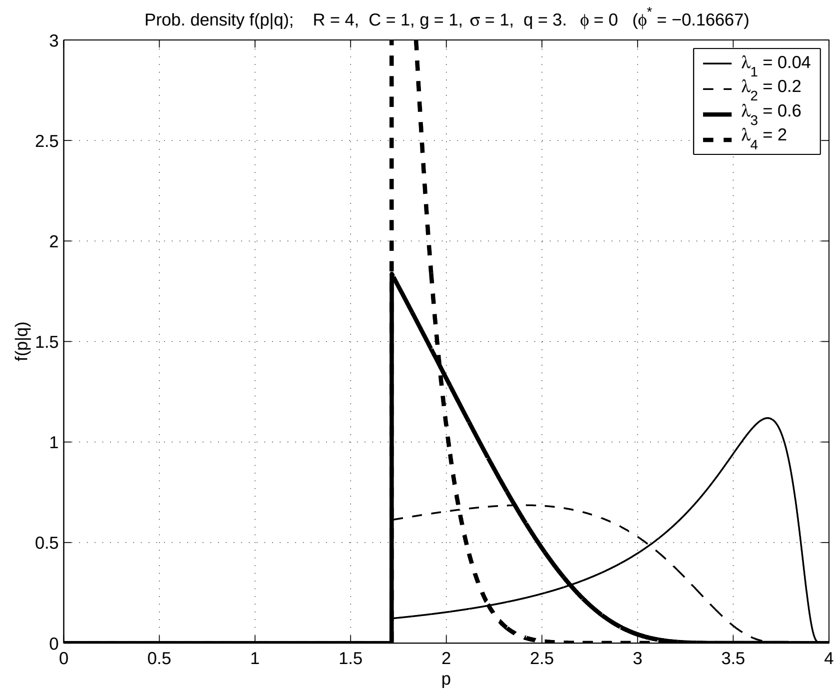


Figure 3.8: Type IV densities ($\phi = 0$), for different values of intensity λ.

so we may say that when the *previous* value q of the estimation error variance is smaller than the above threshold we have to rely on new measurements to estimate state x, thus a high sampling intensity is desirable. However, note once more that the value of λ does not influence the *support* of the conditional probability density, but only its values within a fixed support.

Type IV densities ($\phi = 0$). These densities, that correspond to the particular case of *Brownian motion*, are plotted in Figure 3.8 for different values of sampling intensity λ.

Note in Figure 3.18 that the support of conditional density $f(p|q)$ is $[s_2, R/C^2]$; as usual such support is independent of λ. The sampling intensity, however, has an influence on the shape of $f(p|q)$ in the sense that higher values of λ shift the area below the curve to the right: this is in accordance with the intuitive fact that sampling Brownian motion more frequently makes the tracking of its trajectory more accurate, by reducing the corresponding estimation error variance.

Type V densities ($\phi > 0$). These densities, that correspond to the random sampling of *unstable* dynamical systems, are are shown in Figure 3.9 (for a fixed sampling intensity $\lambda = 1$ and different values of $\phi > 0$) and in Figure 3.10 (for a fixed eigenvalue $\phi = 0.2$ and different sampling intensities). The support of $f(p|q)$

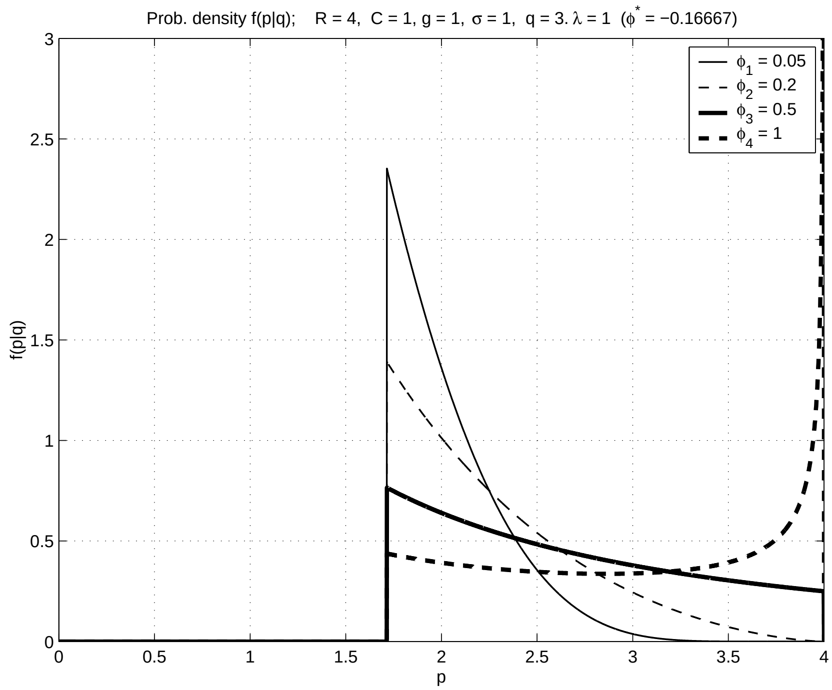


Figure 3.9: Type V densities, for a fixed value of λ.

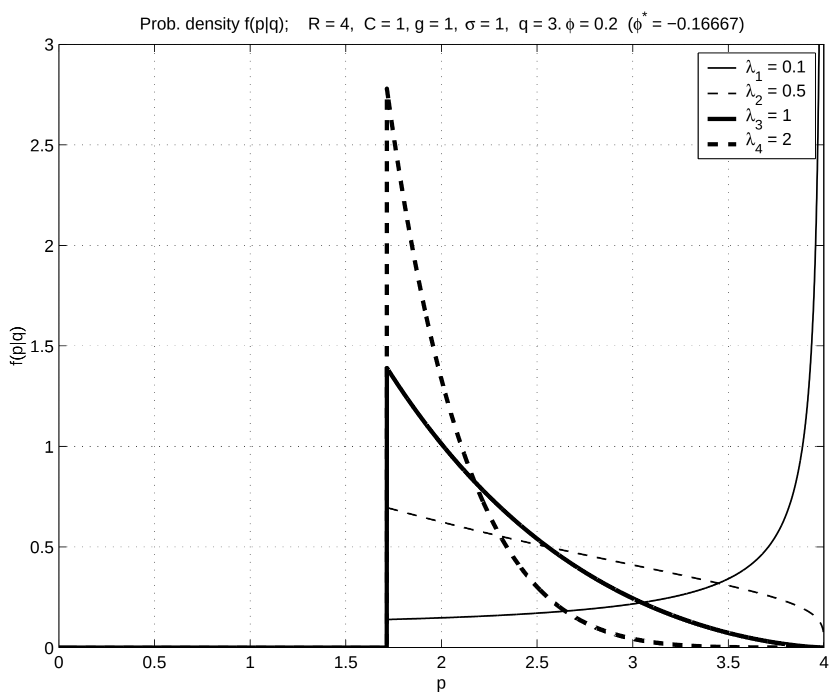


Figure 3.10: Type V densities, for a fixed value of ϕ.

is given by $[s_2, R/C^2]$, independently of the (positive) value of ϕ (and of sampling intensity λ).

Note in Figure 3.9 that the more unstable the continuous-time system is, the hardest it is to track its state x: in fact higher values of ϕ increase the probability of getting a high estimation error variance. On the other hand, Figure 3.10 shows that increasing the sampling intensity shifts the area below the graph of $f(p|q)$ to the left, thus making state estimation more accurate. This is in accordance with the intuitive idea that performing state estimation of an unstable system is easier when such system is sufficiently slow, or when measurements occur sufficiently often.

3.3.3 Plots of conditional mean

We found it useful, for the purposes of understanding how eigenvalue ϕ and sampling intensity λ influence the effectiveness of state estimation, to plot the *mean* of $P_{k+1|k+1}$ given $\{P_{k|k} = q\}$ as a function of such parameters.

We plotted conditional expectation

$$\mathbb{E}\big[P_{k+1|k+1} \,\big|\, P_{k|k} = q\big] \tag{3.21}$$

in Figures 3.11 and 3.12; for each (ϕ, λ) pair the expectation value was calculated *numerically*, simply by estimating the value of integral

$$\int_0^{\frac{R}{C^2}} p\, f(p|q)\; dp$$

in the five cases described in Theorem 3.1 and subsection 3.3.2. In the next section we shall see that in the case of unstable systems ($\phi > 0$) and Brownian motion ($\phi = 0$) it is actually possible to compute an *analytical* expression for the conditional expectation as a function of the system's parameters.

The first remarkable fact in Figures 3.11 and 3.12 is that when $\phi = \phi^*$ conditional mean (3.21) is equal to $\frac{R}{C^2} q\big(q + \frac{R}{C^2}\big)^{-1}$, *independently* of intensity λ; in fact when $\phi = \phi^*$ we have a Type II density (3.11), which is a Dirac delta function centered in s_2. For any value of the sampling intensity expectation (3.21) is an increasing function of eigenvalue ϕ, implying that the more unstable a system is, the harder it is to perform state estimation.

When ϕ is greater than ϕ^* conditional expectation (3.21) is an increasing function of sampling intensity λ. On the other hand, the opposite happens when $\phi < \phi^*$: this is a consequence of the paradoxical behavior of Type I distributions, which we thoroughly examined in the previous subsection.

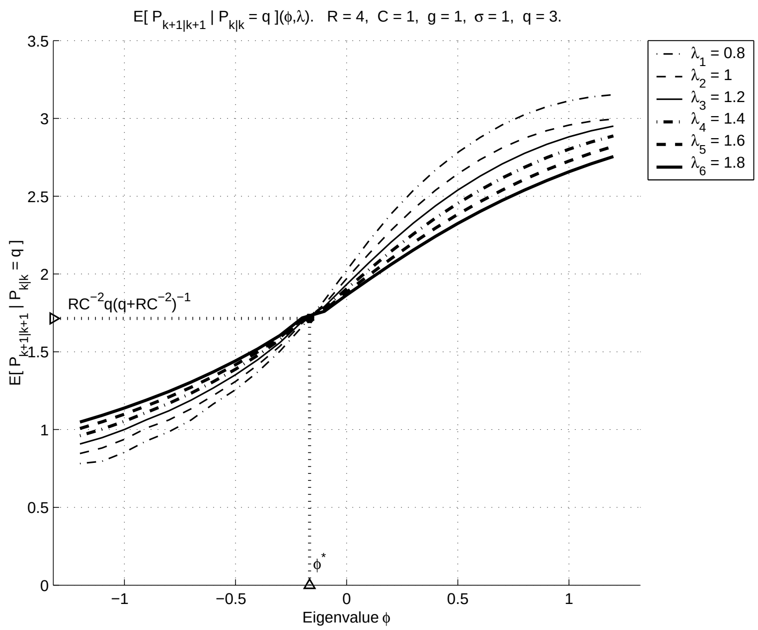


Figure 3.11: Conditional mean $\mathbb{E}[P_{k+1|k+1}|P_{k|k} = q]$ as a function of eigenvalue ϕ, for different sampling intensities.

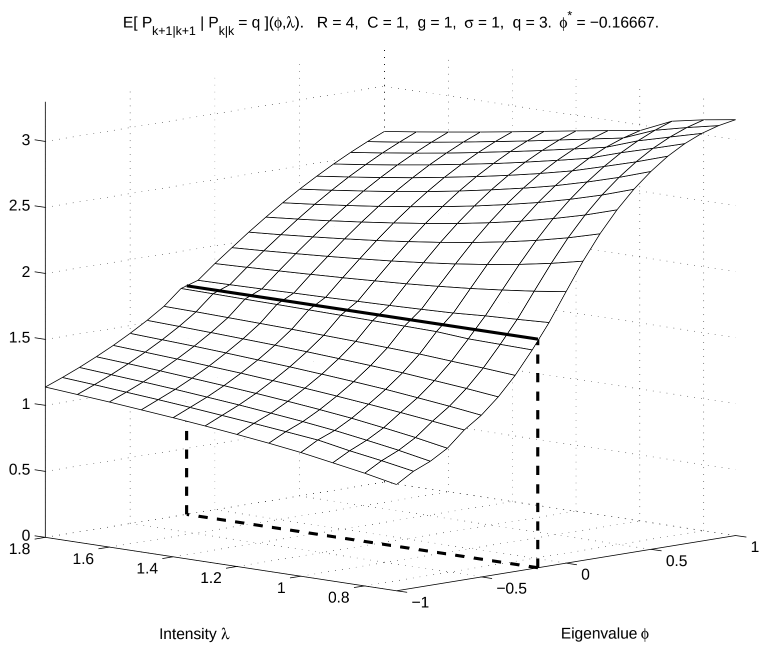


Figure 3.12: Three-dimensional plot of conditional mean $\mathbb{E}[P_{k+1|k+1}|P_{k|k} = q]$, as a function of eigenvalue ϕ and sampling intensity λ. The thick line corresponds to eigenvalue $\phi = \phi^*$.

3.4 An analytical expression for the conditional mean

The graphs plotted in Figures 3.11 and 3.12 were obtained by numerical computation. However, it is possible to obtain an analytical expression of conditional mean $\mathbb{E}[P_{k+1|k+1}|P_{k|k}=q]$ in the case of unstable systems and Brownian motion.

Theorem 3.2 *Assume $\phi > 0$. Then the expectation of $P_{k+1|k+1}$ given $\{P_{k|k}=q\}$ has the following form:*

$$\begin{aligned}\mathbb{E}\big[P_{k+1|k+1}\,\big|\,P_{k|k}=q\big] &= \\ &= \frac{R}{C^2}\left[1-\frac{R}{C^2}\frac{1}{\left(q+\frac{g^2\sigma^2}{2\phi}\right)\left(1+\frac{2\phi}{\lambda}\right)}\;{}_2F_1\!\left(1,1+\frac{\lambda}{2\phi},2+\frac{\lambda}{2\phi};\frac{\frac{g^2\sigma^2}{2\phi}-\frac{R}{C^2}}{q+\frac{g^2\sigma^2}{2\phi}}\right)\right],\end{aligned}$$

where ${}_2F_1(a,b,c;z)$ is the hypergeometric function.[6]

Theorem 3.3 *Assume $\phi = 0$ (Brownian motion). Then the expectation of $P_{k+1|k+1}$ given $\{P_{k|k}=q\}$ has the following form:*

$$\begin{aligned}\mathbb{E}\big[P_{k+1|k+1}\,\big|\,P_{k|k}=q\big] &= \\ &= \frac{R}{C^2}\left\{1-\frac{R}{C^2}\frac{\lambda}{g^2\sigma^2}\;\exp\!\left[\frac{\lambda}{g^2\sigma^2}\left(q+\frac{R}{C^2}\right)\right]\,\Gamma\!\left(0,\frac{\lambda}{g^2\sigma^2}\left(q+\frac{R}{C^2}\right)\right)\right\},\end{aligned}$$

where $\Gamma(z,b)$ is the incomplete gamma function.[7]

[6] Also known as *Gauss' hypergeometric function*, it is defined by series:

$${}_2F_1(a,b,c;z)=\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!}\,, \tag{3.22}$$

known as *hypergeometric series*; in the above definition *Pochhammer symbols* are used:

$$(a)_n = a(a+1)(a+2)\ldots(a+n-1).$$

The hypergeometric function is defined within its circle of convergence, which is the unit circle. However, the analytic continuation of (3.22) to the complex plane cut along the real axis from 1 to ∞ is given by *Euler's hypergeometric integral*:

$${}_2F_1(a,b,c;z)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\,dt\,, \tag{3.23}$$

where $\Gamma(\cdot)$ is the gamma function; formula (3.23) is valid if $\Re[c] > \Re[b] > 0$. The hypergeometric function is the solution solution of Gauss' hypergeometric differential equation, which arises in many classical problems of mathematical physics; for particular choices of arguments a, b, and c the hypergeometric function reduces to Chebyshev, Legendre or Jacobi polynomials in complex variable z. For further details see the classical texts [1] and [8], or [13] for a more modern approach with numerous examples taken from physics.

[7] The gamma function and the *incomplete* gamma function are respectively defined as:

$$\Gamma(z)=\int_0^{\infty} t^{z-1}e^{-t}\,dt \qquad\text{and}\qquad \Gamma(z,b)=\int_b^{\infty} t^{z-1}e^{-t}\,dt\,.$$

See [1, Chap. 6] for details.

Proof of Theorem 3.2. We may rewrite the second of equations (3.5) as follows:

$$P_{k+1|k+1} = \rho \frac{(P_{k|k} + \delta)\, e^{2\phi T_k} - \delta}{(P_{k|k} + \delta)\, e^{2\phi T_k} - \delta + \rho}\,,$$

where $\rho \triangleq \frac{R}{C^2}$, $\delta \triangleq \frac{g^2\sigma^2}{2\phi}$; since T_k is independent of $P_{k|k}$ we have that $f_{T_k|P_{k|k}} = f_{T_k}$, hence:

$$\begin{aligned}
\mathbb{E}\big[P_{k+1|k+1} \,\big|\, P_{k|k} = q\big] &= \int_{-\infty}^{\infty} P_{k+1|k+1}\; f_{T_k|P_{k|k}}(t|q)\; dt \;= \\
&= \rho\lambda \int_0^{\infty} \frac{(q+\delta)\, e^{2\phi t} - \delta}{(q+\delta)\, e^{2\phi t} - \delta + \rho}\, e^{-\lambda t}\; dt \;= \\
&= \rho\lambda \int_0^{\infty} \left[1 - \frac{\rho}{(q+\delta)\, e^{2\phi t} - \delta + \rho}\right] e^{-\lambda t}\; dt \;= \\
&= \rho\lambda \int_0^{\infty} e^{-\lambda t}\; dt - \rho^2\lambda \int_0^{\infty} \frac{e^{-\lambda t}}{(q+\delta)\, e^{2\phi t} - \delta + \rho}\; dt \;= \\
&= \rho\left\{1 - \rho\lambda \int_0^{\infty} \frac{e^{-\lambda t}}{(q+\delta)\, e^{2\phi t} - \delta + \rho}\; dt\right\}. \qquad (3.24)
\end{aligned}$$

Let's now compute integral:

$$\begin{aligned}
&\int_0^{\infty} \frac{-\lambda\, e^{-\lambda t}}{(q+\delta)\, e^{2\phi t} - \delta + \rho}\; dt \;\overset{(*)}{=}\; -\int_0^1 \frac{u^{\frac{2\phi}{\lambda}}}{(q+\delta) + (\rho - \delta)\, u^{\frac{2\phi}{\lambda}}}\; du \;= \\
&\overset{(**)}{=}\; -\frac{\lambda}{2\phi} \int_0^1 \frac{v^{\frac{\lambda}{2\phi}}}{(q+\delta) + (\rho-\delta)\, v}\; dv \;=\; -\frac{\lambda}{2\phi}\,\frac{1}{q+\lambda} \int_0^1 v^{\frac{\lambda}{2\phi}} \left(1 - \frac{\delta - \rho}{q+\delta}\, v\right)^{-1} dv \\
&=\; -\frac{\lambda}{2\phi}\,\frac{1}{q+\lambda} \int_0^1 v^{\left(1+\frac{\lambda}{2\phi}\right)-1} (1-v)^{\left(2+\frac{\lambda}{2\phi}\right)-\left(1+\frac{\lambda}{2\phi}\right)-1} \left(1 - \frac{\delta-\rho}{q+\delta}\, v\right)^{-1} dv\,;
\end{aligned}$$

step $(*)$ is justified by substitution $u = e^{-\lambda t}$ (note that since $\phi > 0$ the integral on the right-hand side of $(*)$ has to be between 0 and 1), whereas substitution $v = u^{\frac{2\phi}{\lambda}}$ yields step $(**)$. Comparing the last expression with Euler's hypergeometric integral:[8]

$${}_2F_1(a, b, c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1} (1-t)^{c-b-1} (1-tz)^{-a}\, dt\,,$$

which is only valid for $\Re[c] > \Re[b] > 0$, yields:

$$\begin{aligned}
&\int_0^{\infty} \frac{-\lambda\, e^{-\lambda t}}{(q+\delta)\, e^{2\phi t} - \delta + \rho}\; dt \;= \\
&\quad = \; -\frac{\lambda}{2\phi}\,\frac{1}{q+\lambda}\,\frac{\Gamma\big(1+\frac{\lambda}{2\phi}\big)\,\Gamma(1)}{\Gamma\big(2+\frac{\lambda}{2\phi}\big)}\;\; {}_2F_1\!\left(1, 1+\frac{\lambda}{2\phi}, 2+\frac{\lambda}{2\phi}; \frac{\delta-\rho}{q+\delta}\right); \qquad (3.25)
\end{aligned}$$

[8] See footnote 6.

note that condition $\Re[c] > \Re[b] > 0$ is satisfied since in our case $c = 2 + \frac{\lambda}{2\phi}$ and $b = 1 + \frac{\lambda}{2\phi}$, with $\phi > 0$; note also that since $\delta > 0$, $\rho > 0$ and $q > 0$ we have that the last argument of the hypergeometric function is (real and) less than one, hence it belongs to the complex plane cut along the real axis from 1 to ∞, which is the region where Euler's integral is well defined. Furthermore, by the well known relation

$$\Gamma(z+1) = z\,\Gamma(z)$$

we get:

$$\frac{\Gamma\big(1+\frac{\lambda}{2\phi}\big)}{\Gamma\big(2+\frac{\lambda}{2\phi}\big)} = \frac{1}{1+\frac{\lambda}{2\phi}}.$$

Finally, substituting (3.25) into (3.24) and then writing the expressions for ρ and δ yields expectation $\mathbb{E}[P_{k+1|k+1}|P_{k|k} = q]$, as expressed by the theorem. □

Proof of Theorem 3.3. We may rewrite the first of equations (3.5) as follows:

$$P_{k+1|k+1} = \rho \frac{P_{k|k} + \eta T_k}{P_{k|k} + \eta T_k + \rho},$$

where $\rho \triangleq \frac{R}{C^2}$ and $\eta \triangleq g^2\sigma^2$; since T_k is independent of $P_{k|k}$ we have that $f_{T_k|P_{k|k}} = f_{T_k}$, hence:

$$\begin{aligned}
\mathbb{E}\big[P_{k+1|k+1}\,\big|\,P_{k|k} = q\big] &= \int_{-\infty}^{\infty} P_{k+1|k+1}\ f_{T_k|P_{k|k}}(t|q)\ dt \ = \\
&= \rho\lambda \int_0^{\infty} \frac{q+\eta t}{q+\eta t+\rho}\, e^{-\lambda t}\ dt \ = \\
&= \rho\left\{1 - \rho\lambda \int_0^{\infty} \frac{e^{-\lambda t}}{q+\eta t+\rho}\ dt\right\}.
\end{aligned} \tag{3.26}$$

Let us calculate the integral that appears in (3.26):

$$\begin{aligned}
\int_0^{\infty} \frac{e^{-\lambda t}}{q+\eta t+\rho}\ dt \ &= \ \frac{1}{\eta}\int_0^{\infty} \frac{e^{-\lambda t}}{\eta t + \frac{q+\rho}{\eta}}\ dt \ = \\
&\overset{(\star)}{=} \frac{1}{\eta}\exp\left[\frac{\lambda}{\eta}(q+\rho)\right]\int_{\frac{\lambda}{\eta}(q+\rho)}^{\infty} x^{-1}\, e^{-x}\ dx \ = \\
&= \frac{1}{\eta}\exp\left[\frac{\lambda}{\eta}(q+\rho)\right]\Gamma\left(0, \frac{\lambda}{\eta}(q+\rho)\right)
\end{aligned} \tag{3.27}$$

where step $(\star)$ is obtained by substitution $x = \lambda\big(t + \frac{q+\rho}{\eta}\lambda\big)$ whereas in the last step we used the definition of the incomplete gamma function.[9] Finally, substituting

[9]See footnote 7.

formula (3.27) into (3.26) and using the expressions for η and ρ given above completes the proof. $\square$

The hypergeometric function takes simple forms in some particular cases. The next corollary exploits this fact and gives a simple expression for conditional expectation (3.21) when $\phi > 0$ and sampling intensity λ is a multiple of 2ϕ, i.e. $\lambda = 2n\phi$; note that we proved in Chapter 2 that these are significant cases since they represent the thresholds above which discrete noise Q_k in (2.1) does not have m-th order moments, for any $m \geq n$.

Corollary 3.4 *Assume $\phi > 0$, If there exists $n \in \mathbb{N}$ such that $\lambda = 2n\phi$ then the expectation of $P_{k+1|k+1}$ given $\{P_{k|k} = q\}$ may be written as follows:*

$$\begin{aligned}
\mathbb{E}\big[P_{k+1|k+1} \,\big|\, P_{k|k} = q\big] &= \\
&= \frac{R}{C^2}\left\{1 + \frac{R}{C^2}\frac{1}{\left(q + \frac{g^2\sigma^2}{2\phi}\right)\left(1 + \frac{2\phi}{\lambda}\right)}\frac{n+1}{z_0^{n-1}}\left[\sum_{j=1}^{n}\frac{z_0^j}{j!} + \log(1 - z_0)\right]\right\},
\end{aligned}$$

where $z_0 \triangleq \dfrac{\frac{g^2\sigma^2}{2\phi} - \frac{R}{C^2}}{q + \frac{g^2\sigma^2}{2\phi}}$.

Proof. Even though we are dealing with real functions it is convenient to use some tools provided by complex analysis [14]. By Theorem 3.2 it is sufficient to prove that:

$$ {}_2F_1(1, 1+n, 2+n; z) = -\frac{n+1}{z^{n-1}}\left[\sum_{j=1}^{n}\frac{z^j}{j!} + \mathrm{Log}(1 - z)\right], \tag{3.28}$$

for any z in $\mathcal{I} \triangleq \mathbb{C} \setminus \{z : \Im[z] = 0, \Re[z] \geq 1\}$, which is the complex region where the hypergeometric function is defined.[10] Note that this coincides with the region of definition of $\mathrm{Log}(\cdot)$, the principal value of the logarithm.

We shall proceed by proving the validity (3.28) within the circle of convergence of the hypergeometric series; the proof will be completed by applying the Identity Theorem for analytic functions (see [14, p. 187]). We have that:

$$ {}_2F_1(1, 1+n, 2+n; z) = \sum_{k=0}^{\infty}\frac{(1)_k(1+n)_k}{(2+n)_k}\frac{z^k}{k!};$$

the Pochhammer symbols in the above expression may be written as follows:

$$\begin{aligned}
(1+n)_k &= (1+n)(2+n)\dots(n+k) \\
(2+n)_k &= (2+n)(3+n)\dots(n+k+1) \\
(1)_k &= k!
\end{aligned}$$

[10]See footnote 6.

so that:

$$\frac{(1)_k(1+n)_k}{(2+n)_k}\frac{1}{k!} = \frac{1+n}{1+n+k}\,.$$

Assume for the moment that $z \in (-1,1) \subset \mathbb{R}$; we shall have that:

$$\begin{aligned}
{}_2F_1(1, 1+n, 2+n; z) &= \sum_{k=0}^{\infty} \frac{1+n}{1+n+k}\, z^k = \\
&= \frac{1+n}{z^{1+n}} \sum_{k=0}^{\infty} \frac{1}{1+n+k}\, z^{1+n+k} \overset{(\dagger)}{=} \frac{1+n}{z^{1+n}} \int_0^z \sum_{k=0}^{\infty} \zeta^{n+k}\, d\zeta = \\
&= \frac{1+n}{z^{1+n}} \int_0^z \left[\sum_{j=0}^{\infty} \zeta^j - \sum_{j=0}^{n-1} \zeta^j \right] d\zeta = \frac{1+n}{z^{1+n}} \int_0^z \left[\frac{1}{1-\zeta} - \sum_{j=0}^{n-1} \zeta^j \right] d\zeta = \\
&= \frac{1+n}{z^{1+n}} \left[\log \frac{1}{1-\zeta} - \sum_{j=0}^{n-1} \frac{\zeta^{j+1}}{j+1} \right]_0^z = -\frac{1+n}{z^{1+n}} \left[\log(1-z) + \sum_{j=1}^{n} \frac{z^j}{j} \right],
\end{aligned}$$

where in step (†) we have used integration by series [12, p. 151]. So (3.28) holds for any $z \in (-1,1) \subset \mathbb{R}$; by the Identity Theorem it holds for all $z \in \{\zeta \in \mathbb{C} : |\zeta| < 1\}$ as well, i.e. in the circle of convergence of the hypergeometric series, and also in $\mathcal{I}$, the extended region of existence of ${}_2F_1$ defined by Euler's hypergeometric integral; in particular, $\mathcal{I}$ contains all real negative values[11] of variable z. □

Having a simple expression for conditional expectation (3.21) is useful in the sense that it can help to choose sampling intensity λ (i.e. the number of sensors in an network) for an unstable system in a way that the expected estimation error variance is smaller than a given threshold: one could fix $q = R/C^2$ (which is the worst possible case) and then run an elementary computer program that finds the lowest $n \in \mathbb{N}$ so that $\mathbb{E}[P_{k+1|k+1}|P_{k|k} = q]$, whose expression is given by Corollary 3.4, is small enough. We shall return on the topic of bounding the estimation error variance in the next Chapter.

3.5 Conclusions

The state estimation process of the randomly sampled continuous-time dynamical system (1.1) has a random nature, in the sense that the estimation error variance $\{P_{k|k}\}_{k=1}^{\infty}$ is a stochastic process itself.

In this chapter we have computed analytical expressions of "transition" probability density (3.4), which by (3.3) is sufficient to provide a complete statistical description of the estimation error variance process $\{P_{k|k}\}_{k=1}^{\infty}$. In particular, we have examined in great detail how density $f(p|q)$ has a support and a shape that

[11] Note that z_0, defined in text of the corollary we have just proven, is always real and less than 1.

strictly depend on the dynamics of the original continuous-time system (represented by eigenvalue ϕ) and on the intensity λ of random sampling. We have proved that the state trajectory of unstable stochastic systems is relatively difficult to track, in the sense that high values of the estimation error variance are likely to occur, especially when the sampling rate is low. In the case of unstable systems, however, we have seen that if the system is "fast" then there are situations (characterized by Type I densities, which occur if ϕ is lower than a given threshold or –equivalently– if the prior variance is high) when it seems to be convenient to wait until the state approaches zero before performing state estimation.

In the last section we have computed analytical expressions for conditional expectation $\mathbb{E}[P_{k+1|k+1}|P_{k|k} = q]$ in the cases $\phi > 0$ (unstable systems) and $\phi = 0$ (Brownian motion). These expressions are useful when one has to pick a sampling intensity λ (which is proportional to the number of sensors in the physical situation of sensor networks) so that the mean variance is below a given threshold.

We are going to investigate other ways of bounding the estimation error in the next Chapter, where we shall formulate a criterion for the choice of λ that makes the estimation error variance arbitrarily low, with an arbitrary probability. We will use the analysis we performed in this Chapter, proving it provides a useful tool for the application we had in mind at the very beginning, i.e. the networks of sensors.

Chapter 4

Bounding the Estimation Error

As we anticipated in the previous Chapter we are now moving to the problem of limiting the estimation error variance by appropriately choosing a sampling intensity, which is proportional to the number of sensors in a network.

We will first formulate the problem in a rigorous way, then we shall solve it in different situations; in fact we shall see that the solutions will depend on the nature of the original continuous-time system, i.e. on the sign of eigenvalue ϕ. In particular, in the case of negative eigenvalue (asymptotically stable systems) we shall see that there are some situations where in order to properly bound the estimation error variance one has to wait until the state approaches zero.

4.1 Introduction

In Chapter 1 we motivated our research by proving that system (1.1) models a network of a large number of sensors. It is reasonable to think that it is possible to choose the number of such sensors; in other words, if T is the common sampling period of all sensors, one can control the magnitude sampling intensity λ by picking an appropriate number N of sensors by relation:

$$\lambda \simeq \frac{N}{T} \tag{4.1}$$

which we justified by Proposition 1.1 in section 1.2.

In the previous Chapter we have seen how Poisson intensity λ has a direct influence on the performance of the state estimation algorithm we formulated in Chapter 2, since the shape of conditional probability density (3.4) depends on the intensity of the sampling process, sometimes in a very subtle way (think of Type I probability densities). We then computed an analytical expression for the mean of $P_{k+1|k+1}$ given $\{P_{k|k} = q\}$ for $\phi \geq 0$; in the case $\lambda = 2n\phi > 0$ we were able to find a particularly simple expression for such expectation. This suggested a criterion for the choice of λ (i.e. of N in (4.1)) for unstable systems: the sampling intensity can be chosen in a way that the mean of $P_{k|k}$ is lower than a given threshold. Note that this procedure is limited to unstable systems.

Here we shall pursue a different objective. Let ϕ be the eigenvalue of system (1.1). Fix an arbitrary probability $\alpha \in (0,1)$ (close to 1) and an arbitrary estimation error variance p^*; we wish to find a sampling intensity λ^* such that:

$$\mathbb{P}\big[P_{k|k} \le p^*\big] > \alpha\,, \qquad \forall k \in \mathbb{N}\,,$$

for any choice of $\lambda > \lambda^*$. In the following sections we will formulate criteria to solve this problem; in particular, we shall see that the choice of λ^* depends on the nature of the original continuous-time dynamical system (1.1): in fact we will have different solutions for unstable systems, Brownian motion, and asymptotically stable systems. We shall proceed in this order, since the latter case presents peculiar subtleties and difficulties deriving from the existence of Type I densities, which we discussed in detail in the previous Chapter.

4.2 First case: unstable systems

We shall start by proving the following Lemma.

Lemma 4.1 *Assume $\phi > 0$. For all $p \in \big[0, \frac{R}{C^2}\big]$ the following inequality holds:*

$$\mathbb{P}\big[P_{k+1|k+1} \le p \,\big|\, P_{k|k} = q\big] \ge \mathbb{P}\big[P_{k+1|k+1} \le p \,\big|\, P_{k|k} = \tfrac{R}{C^2}\big],$$

for any choice of $q \in \big[0, \frac{R}{C^2}\big]$. In other words, the distribution $F(p|q)$ of random variable $P_{k+1|k+1}|\{P_{k|k} = q\}$ is greater than or equal to the distribution $F\big(p\big|\frac{R}{C^2}\big)$ of random variable $P_{k+1|k+1}|\{P_{k|k} = \frac{R}{C^2}\}$.

Proof. Since the supports of the probability densities of $P_{k+1|k+1}|\{P_{k|k} = q\}$ and $P_{k+1|k+1}|\{P_{k|k} = \frac{R}{C^2}\}$ are, respectively,

$$\left(\frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\,,\frac{R}{C^2}\right) \qquad \text{and} \qquad \left(\frac{1}{2}\frac{R}{C^2}\,,\frac{R}{C^2}\right),$$

and since the first one contains the second one, the proposition is trivial for $p \le \frac{1}{2}\frac{R}{C^2}$.

Hence, let $p \in \big(\frac{1}{2}\frac{R}{C^2}, \frac{R}{C^2}\big)$. Since $p \le R/C^2$ and $\phi > 0$ we have that:

$$\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right) \le \left(\frac{R}{C^2} - p\right)\left(\frac{R}{C^2} + \frac{g^2\sigma^2}{2\phi}\right); \tag{4.2}$$

furthermore we have that quantity:

$$\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right) = \frac{g^2\sigma^2}{2\phi}\left(\frac{R}{C^2} - p\right) + p\frac{R}{C^2} \tag{4.3}$$

is greater than zero, hence dividing (4.2) by the left-hand side of (4.3) does *not* change the direction of the inequality. Since $\phi > 0$, if we take the result of the previous operation to the power of $\frac{\lambda}{2\phi}$ then the direction of the inequality is still unchanged. One further step leads to relation:

$$1-\left[\frac{\left(\frac{R}{C^2}-p\right)\left(q+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} \geq 1-\left[\frac{\left(\frac{R}{C^2}-p\right)\left(\frac{R}{C^2}+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}};$$

comparing the above relation with the expression of Type V conditional probability distributions (3.16) one concludes that:

$$F(p|q) \geq F\big(p\big|\tfrac{R}{C^2}\big),$$

which is exactly what we wanted to prove. □

The previous Lemma states the intuitive fact that when $q = R/C^2$ we are in the worst possible case, in the sense that when this occurs the probability that $P_{k+1|k+1}$ assumes large values is highest. We are now ready to prove the following Theorem, that provides a criterion to choose an intensity λ^* that solves the problem we stated in the previous section. Such intensity can be immediately related to the number of sensors in a network by (4.1).

Theorem 4.2 *Assume $\phi > 0$. Choose $\alpha \in (0,1)$ and $p^* \in \left(\frac{1}{2}\frac{R}{C^2}, \frac{R}{C^2}\right)$ arbitrarily. Define:*

$$\lambda^* \triangleq 2\phi\,\log(1-\alpha)\left(\log\left[\frac{\left(\frac{R}{C^2}-p^*\right)\left(\frac{R}{C^2}+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p^*\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right]\right)^{-1}; \tag{4.4}$$

then we shall have that

$$\mathbb{P}\big[P_{k|k} \leq p^*\big] > \alpha\,, \qquad \forall k \in \mathbb{N}\,,$$

for any choice of $\lambda > \lambda^*$.

Proof. By Lemma 4.1 we have that if $F\big(p^*\big|\frac{R}{C^2}\big) > \alpha$ then $F(p^*|q) > \alpha$, $\forall q \leq \frac{R}{C^2}$. By (3.16) condition $F\big(p^*\big|\frac{R}{C^2}\big) > \alpha$ may be written as follows:

$$1-\left[\frac{\left(\frac{R}{C^2}-p^*\right)\left(\frac{R}{C^2}+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p^*\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} > \alpha\,,$$

that is,

$$\frac{\lambda}{2\phi}\,\log\left[\frac{\left(\frac{R}{C^2}-p^*\right)\left(\frac{R}{C^2}+\frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2}+p^*\left(\frac{R}{C^2}-\frac{g^2\sigma^2}{2\phi}\right)}\right] < \log(1-\alpha)\,.$$

Note that the logarithm on the left-hand side of the above equation is *negative* since the quantity in square brackets is less than one: in fact the its numerator may be written as:

$$\frac{g^2\sigma^2}{2\phi}\left(\frac{R}{C^2}-p^*\right)+\left(\frac{R}{C^2}-p^*\right)\frac{R}{C^2}\,,$$

whereas the terms of its denominator may be rearranged as follows:

$$\frac{g^2\sigma^2}{2\phi}\left(\frac{R}{C^2}-p^*\right)+p^*\frac{R}{C^2}\,,$$

so the denominator is greater than the numerator since $p^* \in \left(\frac{1}{2}\frac{R}{C^2},\frac{R}{C^2}\right)$. In the light of this we can conclude that when $\lambda > \lambda^*$, with λ^* given by (4.4), we shall have that

$$\mathbb{P}\big[P_{k+1|k+1} \leq p^* \,\big|\, P_{k|k} = q\big] > \alpha, \quad \forall k \in \mathbb{N}\,,$$

for any possible $q \leq \frac{R}{C^2}$. Therefore:

$$\begin{aligned}\mathbb{P}\big[P_{k+1|k+1} \leq p^*\big] &= \int_0^{\frac{R}{C^2}} \mathbb{P}\big[P_{k+1|k+1} \leq p^* \,\big|\, P_{k|k} = q\big]\; f_{P_{k|k}}(q)\; dq \;>\\ &> \alpha\int_0^{\frac{R}{C^2}} f_{P_{k|k}}(q)\; dq \;=\; \alpha\,,\end{aligned}$$

and the theorem is thus proven. □

A numerical example. In this example we have chosen $R = 4$, $C = 1$, $g = 1$, $\sigma = 1$, $q = 3$ and $\phi = 0.2$ (these are the same parameters that were used to plot the graphs in Figure 3.10 in the previous Chapter); in particular, $R/C^2 = 4$. We chose $p^* = 2.5$ and different values of probability α, obtaining the following values for λ^*:

α	0.9	0.99	0.999	0.9999
λ^*	5.686	11.373	17.059	22.745

As expected, λ^* increases as α approaches 1.

4.3 Second case: Brownian motion

Similar results hold in the case of Brownian motion, i.e. $\phi = 0$.

Lemma 4.3 *Let $\phi = 0$. For all $p \in \left[0, \frac{R}{C^2}\right]$ the following inequality holds:*

$$\mathbb{P}\big[P_{k+1|k+1} \leq p \,\big|\, P_{k|k} = q\big] \geq \mathbb{P}\big[P_{k+1|k+1} \leq p \,\big|\, P_{k|k} = \tfrac{R}{C^2}\big], \tag{4.5}$$

for any choice of $q \in \left[0, \frac{R}{C^2}\right]$; i.e. the distribution $F(p|q)$ of r.v. $P_{k+1|k+1}|\{P_{k|k} = q\}$ is greater than or equal to the distribution $F\big(p\big|\frac{R}{C^2}\big)$ of r.v. $P_{k+1|k+1}|\{P_{k|k} = \frac{R}{C^2}\}$.

Proof. As for Lemma 4.1, we only have to consider $p \in \left(\frac{1}{2}\frac{R}{C^2}, \frac{R}{C^2}\right)$ since the result is obvious for smaller values of p (when the right-hand side of (4.5) is zero).

Since $q \leq \frac{R}{C^2}$ we have that

$$\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p} - q \geq \frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p} - \frac{R}{C^2},$$

so

$$-\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p} - q\right) \leq -\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p} - \frac{R}{C^2}\right);$$

hence

$$1-\exp\left[-\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p} - q\right)\right] \geq 1-\exp\left[-\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p} - \frac{R}{C^2}\right)\right],$$

which holds for $p \in \left(\frac{1}{2}\frac{R}{C^2}, \frac{R}{C^2}\right)$. Comparing the last expression with Type IV distributions (3.14) yields $F(p|q) \geq F\left(p\middle|\frac{R}{C^2}\right)$. $\square$

Theorem 4.4 *Assume $\phi = 0$. Choose $\alpha \in (0,1)$ and $p^* \in \left(\frac{1}{2}\frac{R}{C^2}, \frac{R}{C^2}\right)$ arbitrarily. Define:*[1]

$$\lambda^* \triangleq -g^2\sigma^2\left[\frac{R}{C^2}\left(\frac{p^*}{\frac{R}{C^2}-p^*} - 1\right)\right]^{-1}\log(1-\alpha); \tag{4.6}$$

then we shall have that

$$\mathbb{P}\left[P_{k|k} \leq p^*\right] > \alpha, \qquad \forall k \in \mathbb{N},$$

for any choice of $\lambda > \lambda^*$.

Proof. By the previous Lemma we have that if $F\left(p^*\middle|\frac{R}{C^2}\right) > \alpha$ then $F(p^*|q) > \alpha$, $\forall q \leq \frac{R}{C^2}$. By (3.14) condition $F\left(p^*\middle|\frac{R}{C^2}\right) > \alpha$ may be written as follows:

$$1-\exp\left[-\frac{\lambda}{g^2\sigma^2}\frac{R}{C^2}\left(\frac{p^*}{\frac{R}{C^2}-p^*} - 1\right)\right] > \alpha,$$

i.e.

$$-\frac{\lambda}{g^2\sigma^2}\frac{R}{C^2}\left(\frac{p^*}{\frac{R}{C^2}-p^*} - 1\right) < \log(1-\alpha);$$

the expression in brackets on the left-hand side is positive, hence we get that condition $F\left(p^*\middle|\frac{R}{C^2}\right) > \alpha$ is equivalent to:

$$\lambda > -g^2\sigma^2\left[\frac{R}{C^2}\left(\frac{p^*}{\frac{R}{C^2}-p^*} - 1\right)\right]^{-1}\log(1-\alpha).$$

[1]Note that λ^* as defined in (4.6) is positive since $\log(1-\alpha) < 0$

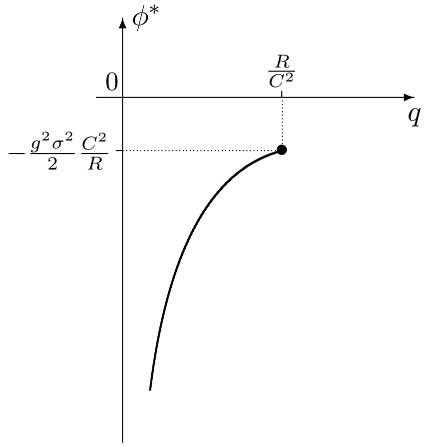


Figure 4.1: Dependence of ϕ^* on q.

So we can conclude that when $\lambda > \lambda^*$, with λ^* given by (4.4), we shall have that

$$\mathbb{P}\big[P_{k+1|k+1} \leq p^* \,\big|\, P_{k|k} = q\big] > \alpha, \quad \forall k \in \mathbb{N},$$

for any possible $q \leq \frac{R}{C^2}$. An argument identical to the one at the end of the proof of Theorem 4.2 allows us to complete this proof as well. □

4.4 Third case: asymptotically stable systems

The last case we analyze ($\phi < 0$) is the most complex one since we have seen in the previous Chapter that it involves three types of conditional densities $f(p|q)$: namely, Types I, II and III. The previous cases only involved one type of conditional density each (Type V for unstable systems and Type IV for Brownian motion). Furthermore, the subtle nature of Type I density requires particular attention.

Figure 4.1 illustrates the dependence of ϕ^* on q:

$$\phi^* = -\frac{g^2\sigma^2}{2}\frac{1}{q};$$

we know from Chapter 3 that whenever $\phi^* < \phi < 0$ we have a Type III distribution; if $\phi < \phi^*$ we have a Type I distribution; finally we have the singular case $\phi = \phi^*$ when $f(p|q)$, is a Dirac delta function.

Since q is a value that the estimation error variance $P_{k|k}$ assumes, it cannot be greater than R/C^2. Therefore Figure 4.1 suggests that we can distinguish two subcases about the sequence of conditional probability densities that appear in:

$$\begin{aligned} f^{(k+1)}(p_{k+1}, p_k, \ldots, p_1, p_0) &= \\ &= f_{k+1|k}(p_{k+1}|p_k) \cdot f_{k|k-1}(p_k|p_{k-1}) \cdot \ldots \cdot f_{1|0}(p_1|p_0) \cdot f_0(p_0), \end{aligned} \tag{4.7}$$

which we introduced at the beginning of the previous Chapter.

Subcase A: $-\frac{g^2\sigma^2}{2}\frac{C^2}{R} < \phi < 0$**.** Since we shall always have that $\phi > \phi^*$ (for any value of q) all conditional densities in chain (4.7) will be of Type III.

Subcase B: $\phi < -\frac{g^2\sigma^2}{2}\frac{C^2}{R}$**.** In this case Types I, II and III are all plausible in (4.7). However, since $P_{k|k}$ (for all k) is a *continuous* random variable, Type II conditional densities occur with probability zero: in fact they occur when ϕ^*, which by (4.1) depends on q (i.e. the value assumed by $P_{k|k}$), is exactly equal to ϕ. We shall see that if any of the densities $f(p_{k+1}|p_k)$ in (4.7) is of Type I, then there exists with probability one an index m such that *all* densities $f(p_{\ell+1}|p_\ell)$, with $\ell \geq m$, will be of Type III.

We shall now analyze these subcases in detail, and formulate for each one of them a criterion for the choice of λ^*.

4.4.1 Subcase A

When the following condition holds:

$$-\frac{g^2\sigma^2}{2}\frac{C^2}{R} < \phi < 0$$

all densities in chain (4.7) are of Type III. In particular, their support is given by

$$(s_2, s_1) = \left(\frac{R}{C^2}\frac{q}{q+\frac{R}{C^2}}\,,\ \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}\right)$$

as we studied in the previous Chapter (see Figure 3.2). Note in particular that the right extreme (s_1) is fixed since it depends on ϕ, whereas the left extreme (s_2) is in general different for the densities in (4.7) since it depends on the previous value of the estimation error variance.

In particular, for any value of q we have that

$$\mathbb{P}\big[P_{k+1|k+1} \leq s_1 \,\big|\, P_{k|k} = q\big] = 1, \quad \forall k \in \mathbb{N}\,, \tag{4.8}$$

hence $\mathbb{P}\big[P_{k+1|k+1} \leq s_1\big] = 1$, for all $k \in \mathbb{N}$. Therefore we are interested in finding a sampling intensity lower bound λ^* that further reduces the estimation error variance. As usual, we will start by proving an auxiliary Lemma. Note that this time the value of q is chosen in $[0, s_1]$ instead of $\big[0, \frac{R}{C^2}\big]$.

Lemma 4.5 *Let* $-\frac{g^2\sigma^2}{2}\frac{C^2}{R} < \phi < 0$*. For all* $p \in \big[0, \frac{R}{C^2}\big]$ *the following inequality holds:*

$$\mathbb{P}\big[P_{k+1|k+1} \leq p \,\big|\, P_{k|k} = q\big] \geq \mathbb{P}\big[P_{k+1|k+1} \leq p \,\big|\, P_{k|k} = s_1\big], \tag{4.9}$$

for any choice of $q \in [0, s_1]$*, i.e. the distribution* $F(p|q)$ *of r.v.* $P_{k+1|k+1}|\{P_{k|k} = q\}$ *is greater than or equal to the distribution* $F(p|s_1)$ *of r.v.* $P_{k+1|k+1}|\{P_{k|k} = s_1\}$*.*

Proof. Since the support of $f(p|q)$ is given by (4.8) we only have to prove (4.9) for $p \in \left(\frac{1}{2}\frac{R}{C^2}, s_1\right)$. We obviously have that:

$$\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right) \leq \left(\frac{R}{C^2} - p\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right); \tag{4.10}$$

note that $-\frac{g^2\sigma^2}{2}\frac{C^2}{R} < \phi < 0$ implies that $\frac{R}{C^2} + \frac{g^2\sigma^2}{2\phi} < 0$, so (since $s_1 < \frac{R}{C^2}$) the right-hand side of (4.10) is *negative* (this was not the case in the proof of Lemma 4.1). Furthermore expression:

$$\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right) \tag{4.11}$$

is *also* negative since $p < \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}$; hence dividing (4.10) by (4.11) yields:

$$\frac{\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)} \geq \frac{\left(\frac{R}{C^2} - p\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}. \tag{4.12}$$

Both sides of (4.12) are *positive*, so we can take a *negative* power (with exponent $\frac{\lambda}{2\phi}$) of both; this reverses the inequality once again and we finally get:

$$F(p|q) = \left[\frac{\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} \leq \left[\frac{\left(\frac{R}{C^2} - p\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} = F(p|s_1),$$

where we have used the expression for Type III distributions (3.12). □

Theorem 4.6 *Assume* $-\frac{g^2\sigma^2}{2}\frac{C^2}{R} < \phi < 0$. *Choose arbitrarily* $\alpha \in (0,1)$ *and*

$$p^* \in \left(\frac{R}{C^2}\frac{s_1}{s_1 + \frac{R}{C^2}}, s_1\right),$$

where $s_1 = \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}$. *Define:*

$$\lambda^* \triangleq 2\phi\log(1-\alpha)\left(\log\left[\frac{\left(\frac{R}{C^2} - p^*\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p^*\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}\right]\right)^{-1}; \tag{4.13}$$

then we shall have that

$$\mathbb{P}\left[P_{k|k} \leq p^*\right] > \alpha, \qquad \forall k \in \mathbb{N},$$

for any choice of $\lambda > \lambda^*$.

Note that this Theorem is very similar to Theorem 4.2, except for the interval where bounding variance p^* is chosen from: in particular, since

$$\frac{R}{C^2}\frac{s_1}{s_1+\frac{R}{C^2}} < \frac{1}{2}\frac{R}{C^2}$$

(because $s_1 < \frac{R}{C^2}$: see Figure 3.3 in the previous Chapter) one is now allowed to choose values of p^* that are smaller than $\frac{1}{2}\frac{R}{C^2}$. Also note that in expression (4.13) there is a factor where s_1 appears. The proof that follows is similar to the one of Theorem 4.2, except that one has to be much more careful with the signs of the quantities that are manipulated.

Proof. By Lemma 4.5 we have that if $F(p^*|s_1) > \alpha$ then $F(p^*|q) > \alpha$, $\forall q \leq s_1$. By (3.12) condition $F(p^*|s_1) > \alpha$ may be written as follows:

$$1 - \left[\frac{\left(\frac{R}{C^2} - p^*\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p^*\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}\right]^{\frac{\lambda}{2\phi}} > \alpha\,,$$

that is,

$$\frac{\lambda}{2\phi}\log\left[\frac{\left(\frac{R}{C^2} - p^*\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p^*\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}\right] < \log(1-\alpha)\,. \tag{4.14}$$

This time the logarithm on the left-hand side of the above equation is *positive* since the quantity in square brackets is *greater* than one: in fact the its numerator may be written as:

$$\text{Num} \triangleq \frac{g^2\sigma^2}{2\phi}\left(\frac{R}{C^2} - p^*\right) + \left(\frac{R}{C^2} - p^*\right)s_1\,,$$

whereas we may rearrange the terms of its denominator as follows:

$$\text{Den} \triangleq \frac{g^2\sigma^2}{2\phi}\left(\frac{R}{C^2} - p^*\right) + p^*\frac{R}{C^2}\,;$$

condition $p^* > \frac{R}{C^2}\frac{s_1}{s_1+\frac{R}{C^2}}$ implies

$$\left(\frac{R}{C^2} - p^*\right)s_1 < p^*\frac{R}{C^2}$$

so that Num < Den; but condition $p^* < s_1$ implies that Den < 0, so that $\frac{\text{Num}}{\text{Den}} > 1$.

Therefore the logarithm in (4.14) is positive; but ϕ is negative, therefore we have that condition $F(p^*|s_1) > \alpha$ is verified if and only if $\lambda > \lambda^*$, with λ^* given by (4.13). If this is the case we shall have that

$$\mathbb{P}\big[P_{k+1|k+1} \leq p^* \,\big|\, P_{k|k} = q\big] > \alpha, \quad \forall k \in \mathbb{N}\,,$$

for any possible $q \leq s_1$. Since $\mathbb{P}[P_{k|k} < s_1] = 1$ we conclude, with the same argument that we used at the end of the proof of Theorem 4.2, that $\mathbb{P}\big[P_{k+1|k+1} \leq p^*\big] > \alpha$, for all k. □

4.4.2 Subcase B

Suppose now that the following condition holds:

$$\phi < -\frac{g^2\sigma^2}{2}\frac{C^2}{R}\,;$$

in principle, probability density Types I, II and III are possible. However, the next Theorem states that if any of the conditional densities in chain (4.7) is of Type I then there exists a time after which all the subsequent densities will be of Type III.

Theorem 4.7 *Let $\phi < -\frac{g^2\sigma^2}{2}\frac{C^2}{R}$. Consider the following cases.*

- *There exists an index $k \in \mathbb{N}$ such that p_k, the value assumed by random variable $P_{k|k}$, satisfies inequality: $\phi \geq -\frac{g^2\sigma^2}{2}\frac{1}{p_k}$, so that conditional density $f(p_{k+1}|p_k)$ is of Type II or Type III. Then, w.p. 1, for all $j \geq k$ conditional densities $f(p_{j+1}|p_j)$ will be of Type III.*

- *There exists an index $k \in \mathbb{N}$ such that p_k, the value assumed by random variable $P_{k|k}$, satisfies inequality: $\phi < -\frac{g^2\sigma^2}{2}\frac{1}{p_k}$, so that conditional density $f(p_{k+1}|p_k)$ is of Type I. Then w.p. 1 there exists an index $m > k$ such that*

$$-\frac{g^2\sigma^2}{2}\frac{1}{p_m} < \phi\,,$$

i.e. $f(p_{m+1}|p_m)$ is of Type III; for all $j \geq m$ conditional densities $f(p_{j+1}|p_j)$ will be of Type III. Define random variable:

$$m_0 = \min_{m>k}\left\{m : -\frac{g^2\sigma^2}{2}\frac{1}{p_m} < \phi\right\}; \tag{4.15}$$

we will have that $\mathbb{E}\big[m_0\big|P_{k|k} = p_k\big] < \infty$.

Proof. It is a good idea to keep Figure 3.2 handy during the proof, which otherwise risks to be confusing. For convenience, define

$$s_2(i) \triangleq \frac{R}{C^2}\frac{p_i}{p_i + \frac{R}{C^2}}\,, \qquad i \in \mathbb{N}\,, \tag{4.16}$$

which represents extreme s_2 of conditional probability density $f(p_{i+1}|p_i)$ (see Figure 3.2 in Chapter 3). We have that $f(p_{i+1}|p_i)$ is of Type I iff $s_2(i) > s_1$, of Type II iff $s_2(i) = s_1$, of Type III iff $s_2(i) < s_1$. The "transition" from Type I to Type III densities occurs when there exists an index m such that $s_2(m) < s_1$.

Assuming there exists an $m \in \mathbb{N}$ such that $s_2(m) < s_1$, i.e. $f(p_{m+1}|p_m)$ is of Type III, then we will have that p_{m+1} (the value that r.v. $P_{m+1|m+1}$ assumes) will be less than s_1, so $s_2(m+1)$ will be less than s_1 as well; so $f(p_{m+2}|p_{m+1})$ will be of Type III. Hence once there exists an $m \in \mathbb{N}$ such that $f(p_{m+1}|p_m)$ is of Type III, all the subsequent conditional probability densities will be of Type III as well. With a similar argument one proves that if $f(p_{k+1}|p_k)$ is a Type II density then all the subsequent ones will be of Type III.

On the other hand, if $s_2(k) > s_1$ ($f(p_{k+1}|p_k)$ is of Type I) w.p. 1 then we will have that $p_{k+1} < s_2(k)$, but $s_2(k+1) < p_{k+1}$ so $s_2(k+1) < s_2(k)$. If $s_2(k+1) < s_1$ then $f(p_{k+2}|p_{k+1})$ is of Type III and we are done. Otherwise, we will have w.p. 1 that $s_2(k+2) < s_2(k+1)$, and so on. We have to prove that w.p. 1 there exists an index $m > k$ such that $s_2(m) < s_1$.

Condition $s_2(m) < s_1$ is equivalent to:

$$p_m < \vartheta_{s_1} \triangleq \frac{R}{C^2} \frac{s_1}{\frac{R}{C^2} - s_1} ;$$

note that the right-hand side of the above equation is greater than s_1. So assuming that $s_2(k) > s_1$ we will have that $s_2(k+1) < s_1$ (i.e. $f(p_{k+2}|p_{k+1})$ is of Type III) with probability $\mathbb{P}[s_1 < P_{k+1|k+1} < \vartheta_{s_1} | P_{k|k} = p_k] > 0$. Note that if p_i and p_j are such that $p_i > p_j$, $s_2(i) > s_1$ and $s_2(j) > s_1$ (so both $f(p_{i+1}|p_i)$ and $f(p_{j+1}|p_j)$ are of Type I) then one can prove that:

$$\mathbb{P}[s_1 < P_{j+1|j+1} < \vartheta_{s_1} | P_{j|j} = p_j] > \mathbb{P}[s_1 < P_{i+1|i+1} < \vartheta_{s_1} | P_{i|i} = p_i] \tag{4.17}$$

i.e. $F(\vartheta_{s_1}|p_j) > F(\vartheta_{s_1}|p_k)$ (the proof of this fact follows the main lines of the proof of Lemma 4.1, except that one has to use the expression for Type I distributions).

We will now show that the probability that $P_{j|j} > \vartheta_{s_1}$ for all $j > k$ is equal to zero. Assuming that $s_2(k) > s_1$ (which obviously implies that $p_k > \vartheta_{s_1}$), let us calculate:

$$\begin{aligned}
&\mathbb{P}\big[P_{j|j} > \vartheta_{s_1}, \forall j \in \{k+1, k+2, \dots, n\} \,\big|\, P_{k|k} = p_k\big] \;= \qquad (4.18)\\
&= \;\mathbb{P}\big[P_{n|n} > \vartheta_{s_1} \,\big|\, P_{n-1|n-1} > \vartheta_{s_1}, \dots, P_{k+1|k+1} > \vartheta_{s_1}, P_{k|k} = p_k\big] \cdot \\
&\quad \cdot \mathbb{P}\big[P_{n-1|n-1} > \vartheta_{s_1} \,\big|\, P_{n-2|n-2} > \vartheta_{s_1}, \dots, P_{k+1|k+1} > \vartheta_{s_1}, P_{k|k} = p_k\big] \cdot \\
&\qquad \dots \cdot \; \mathbb{P}\big[P_{k+1|k+1} > \vartheta_{s_1} \,\big|\, P_{k|k} = p_k\big] \;= \\
&= \;\mathbb{P}\big[P_{n|n} > \vartheta_{s_1} \,\big|\, P_{n-1|n-1} > \vartheta_{s_1}, P_{k|k} = p_k\big] \cdot \\
&\quad \cdot \mathbb{P}\big[P_{n-1|n-1} > \vartheta_{s_1} \,\big|\, P_{n-2|n-2} > \vartheta_{s_1}, P_{k|k} = p_k\big] \cdot \\
&\qquad \dots \cdot \; \mathbb{P}\big[P_{k+1|k+1} > \vartheta_{s_1} \,\big|\, P_{k|k} = p_k\big] \;< \\
&< \;\Big\{\mathbb{P}\big[P_{k+1|k+1} > \vartheta_{s_1} \,\big|\, P_{k|k} = p_k\big]\Big\}^{n-k},
\end{aligned}$$

where in the last step we have used inequality (4.17) $n-k-1$ times. So it is now clear that probability (4.18) tends to zero as $n \to \infty$, which means that with probability 1 there exists an index $m > k$ such that for all $j \geq m$ conditional densities $f(p_{j+1}|p_j)$ will be of Type III.

In order to prove the last part of the Theorem let us define:

$$\pi_0 \triangleq F(\vartheta_{s_1}|p_{k+1}) = \mathbb{P}\big[s_1 < P_{k+1|k+1} < \vartheta_{s_1} \,\big|\, P_{k|k} = p_k\big]$$

and consider random variable m_0 defined in (4.15). We will have:

$$\begin{aligned}
\mathbb{P}[m_0 = n|P_{k|k} = p_{k+1}] &= \\
&= \mathbb{P}\big[s_1 < P_{n|n} < \vartheta_{s_1}, P_{j|j} > \vartheta_{s_1}\ \forall j \in \{k+1,\ldots,n-1\} \,\big|\, P_{k|k} = p_k\big] = \\
&= \mathbb{P}\big[s_1 < P_{n|n} < \vartheta_{s_1} \,\big|\, P_{n-1|n-1} > \vartheta_{s_1}, P_{k|k} = p_k\big] \cdot \\
&\quad \cdot \mathbb{P}\big[P_{n-1|n-1} > \vartheta_{s_1} \,\big|\, P_{n-2|n-2} > \vartheta_{s_1}, P_{k|k} = p_k\big] \cdot \\
&\qquad \ldots \cdot\ \mathbb{P}\big[P_{k+1|k+1} > \vartheta_{s_1} \,\big|\, P_{k|k} = p_k\big] < \\
&< (1-\pi_0)^{n-k-1},
\end{aligned}$$

so:

$$\begin{aligned}
\mathbb{E}[m_0|P_{k|k} = p_k] &= \sum_{n=k+1}^{\infty} n\,\mathbb{P}[m_0 = n|P_{k|k} = p_k] \\
&< \sum_{n=k+1}^{\infty} n\,(1-\pi_0)^{n-k-1} = \frac{1+k\pi_0}{\pi_0^2}\,,
\end{aligned}$$

which is finite. □

Thus we have proven the fundamental result that when in (4.7) there is a probability density of Type I then after a while (in finite time) we will only have probability densities of Type III. This phenomenon corresponds to the intuitive idea that in the case we were performing state estimation on an asymptotically stable system ($\phi < 0$) it may be convenient to wait until its state $x(t)$ approaches zero in order to have better estimates, as we discussed in the previous Chapter. As the state approaches zero, Type I densities turn into Type III densities. At that point, in order to find λ^* we will just apply the same ideas we used in treating Subcase A in subsection 4.4.1.

In fact, the following Theorem holds; it differs from Theorem 4.6 only in the first and last lines.

Theorem 4.8 *Assume $\phi < -\frac{g^2\sigma^2}{2}\frac{C^2}{R}$. Choose arbitrarily $\alpha \in (0,1)$ and*

$$p^* \in \left(\frac{R}{C^2}\frac{s_1}{s_1 + \frac{R}{C^2}}\,, s_1\right),$$

where $s_1 = \frac{R}{C^2}\frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}$. Define:

$$\lambda^* \triangleq 2\phi\,\log(1-\alpha)\left(\log\left[\frac{\left(\frac{R}{C^2} - p^*\right)\left(s_1 + \frac{g^2\sigma^2}{2\phi}\right)}{\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p^*\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)}\right]\right)^{-1};$$

then any choice of $\lambda > \lambda^*$ we shall have that

$$\mathbb{P}\big[P_{k|k} \leq p^*\big] > \alpha\,, \qquad \forall k \in \mathbb{N}\,,$$

for all $k \geq m$, where m is any index such that $\phi > -\frac{g^2\sigma^2}{2}\frac{1}{p_m}$ (with p_m we intend the value taken by random variable $P_{m|m}$).

Proof. It is analogous to the proof of Theorem 4.6, with some special attention due to the fact that the Theorem only holds when the conditional probability densities have become of Type III. □

4.5 Conclusions

In this Chapter we have studied the problem of bounding the estimation error variance by acting on sampling intensity λ, i.e. on the number of sensors in a network in case this were the physical situation modelled by system (1.1), which we introduced at the very beginning of this thesis.

The particular problem we have solved is finding a lower bound λ^* for the sampling intensity such that, given a variance threshold p^* and a probability α, we will have $\mathbb{P}[P_{k|k} \leq p^*] > \alpha$, $\forall k$, whenever $\lambda > \lambda^*$. We should mention the fact that the conditions we have found on λ are *sufficient* but not *necessary*: in other words there may be lower values of λ^* would do the job. The refinement of our results is left for future work.

As we expected, the analysis we performed on asymptotically stable systems ($\phi < 0$) revealed some surprises. In fact in this case if the prior knowledge on the state is poor (i.e. if initially we have that $\phi < \phi^*$) one can hope to bound the estimation error variance only after the state has approached zero. At that point it will be possible to limit such variance by an appropriate choice of sampling intensity λ.

Chapter 5

Final remarks and future work

5.1 Summary of results

In this thesis we have studied the problem of random sampling of continuous-time (linear) dynamical systems, which arises in the context the analysis of large sensor networks. We have assumed that the sampling is generated by a Poisson process. To our knowledge it is the first time such a study has been performed, even in the simplest case of random sampling of Brownian Motion.

We have formulated a Kalman filter-based algorithm for state estimation; the main difference with ordinary Kalman filtering is that the sequence of estimation error covariance matrices is itself a random process, namely a homogeneous Markov process: therefore it cannot be computed off-line, and is subject to a statistical description. This is due to the fact that the variance of the noise that appears in the sampled state equation (obtained by sampling the continuous-time state equation in correspondence of the Poisson arrivals) is itself *random*, and so is the state matrix of the discrete system.

In the one-dimensional case we have performed a thorough analysis of the sampled system's random parameters, in function of the sampling intensity and the eigenvalue of the original continuous-time system. We have seen in particular that in the case of unstable dynamical systems the (random) variance of the noise in the system matrix may assume large values, especially if the rate of change of the state is high with respect to the sampling intensity. This suggests that *fast* unstable dynamical systems are hard to track when the sampling rate is too low, which is in accordance with physical intuition.

This idea was confirmed by the statistical description of the estimation error variance process, which we computed in Chapter 3. We also analyzed in depth the behavior of the estimator for asymptotically stable dynamical systems, which present particular subtleties that are are typical of this case. What happens is that when the prior knowledge on the state is poor, or (equivalently) when the system is relatively fast, it is convenient to wait until the state approaches zero in order to perform an efficient state estimation: this is the correct interpretation to the fact that in some situations a higher sampling rate does not seem to lower the state estimation error variance.

Finally, in the situation where it is possible to choose the sampling rate, e.g. by increasing the number of sensors in the case we were modelling a network of sensors, we have found a lower bound on the sampling intensity that allows to limit the error variance in an arbitrary way. We have found different results for unstable systems, Brownian motion, and asymptotically stable systems. In particular we have seen that in the latter case lower values of the estimation error variance are obtainable.

5.2 Future research

Since we believe that it is the first time that the subject of random sampling of continuous-time dynamical systems has been studied in depth, future research will largely depend on the success of our study on applications such as the analysis of sensor networks, and hopefully others as well. For this reason we will welcome feedback from the Electrical Engineering, Computer Sciences, Statistics, and Applied Mathematics communities.

In any case, the natural extension of our investigation should be the study of multi-dimensional, linear dynamical systems. The case of Brownian motion by itself should not be too onerous. The study of the general case should start from the Jordan form of the system matrix; we expect the periodic case (i.e. the presence of complex eigenvalues, which do not occur in one dimension) to present subtle and interesting behaviors. The idea of studying some simple classes of nonlinear dynamical systems is quite challenging as well.

Another interesting direction of research would be the refinement of the lower bounds for the sampling intensity we found in Chapter 4: in fact we only found sufficient conditions but not necessary ones. Finding stricter lower bounds, which would imply a deeper study of the sequence of conditional probability densities, would be desirable as these would allow to lower the number (i.e. the cost) of sensors in a network in order to achieve a given performance in state estimation.

Appendix A

Linear Matrix Equations

We will give a brief summary of results in linear matrix differential equations theory, which we used in section 2.1; for more details on this theory, see [2]. The method described in section A.2 is suggested in [11].

A.1 General form of solutions

Given vector differential equation: $\dot{x}(t) = A(t)x(t)$ with initial condition $x(t_0) = x_0 \in \mathbb{R}^n$, assuming that $A(t)$ is an $n \times n$ matrix whose elements are continuous functions of time t there always exists a matrix function $\Phi(t, t_0)$, called *transition matrix*, such that $x(t) = \Phi(t, t_0)x_0$ for all $t \in \mathbb{R}$. In general, it may be written in the so-called *Peano-Baker series* form:

$$\Phi(t, t_0) = \sum_{k=0}^{\infty} M_k(t, t_0), \quad \text{with}: \; M_0 = I, \quad M_k(t, t_0) = I + \int_{t_0}^{t} A(\sigma) M_{k-1}(\sigma, t_0)\, d\sigma$$

which, as a function of t, converges uniformly in any interval of the type $[t_0, t_1]$. When A is *constant* the transition matrix is simply given by an exponential matrix, i.e. $\Phi(t, t_0) = e^{A(t-t_0)}$.

Consider now the following linear matrix equation:

$$\dot{X}(t) = A_1(t)X(t) + X(t)A_2(t) + B(t), \tag{A.1}$$

where all matrices are $n \times n$. The following general result holds:

Theorem A.1 (Matrix Variation of Constants Formula) *Assume $\Phi_1(t, t_0)$ is the transition matrix for equation $\dot{x}(t) = A_1(t)x(t)$ and $\Phi_2(t, t_0)$ is the transition matrix for equation $\dot{x}(t) = A_2^T(t)x(t)$; then the solution of equation (A.1) with the initial value $X(t_0)$ is given by:*

$$X(t) \; = \; \Phi_1(t, t_0)X(t_0)\Phi_2^T(t, t_0) + \int_{t_0}^{t} \Phi_1(t, \sigma)B(\sigma)\Phi_2^T(t, \sigma)\, d\sigma\,. \tag{A.2}$$

In the particular case that A_1, A_2, and B are constant matrices, solution (A.2) simply takes the form:

$$X(t) \;=\; e^{A_1(t-t_0)} X(t_0)\, e^{A_2(t-t_0)} + \int_{t_0}^{t} e^{A_1(t-\sigma)} B\, e^{A_2(t-\sigma)}\, d\sigma\,;$$

so, if $A_1 = F$, $A_2 = F^T$, $B = GSG^T$ and $t_0 = 0$ (as in equation (2.5)), we will get:

$$\begin{aligned} X(t) &= e^{Ft} X(0)\, e^{F^T t} + \int_0^t e^{F(t-\sigma)} GSG^T\, e^{F^T(t-\sigma)}\, d\sigma \\ &= e^{Ft} X(0)\, e^{F^T t} + \int_0^t e^{F\tau} GSG^T\, e^{F^T\tau}\, d\sigma\,. \end{aligned}$$

This way, it is easily seen that Q_k, as expressed in formula (2.4) of Chapter 2, is indeed the solution $Q(t)$ of matrix differential equation (2.5) calculated it $t = T_k$, since in this case $X(0) = Q(0) = 0$ (which is the value that Q_k would assume in the case $T_k = 0$).

A.2 A numerical method

Consider matrix equation (2.5):

$$\dot{Q} = FQ + QF^T + GSG^T\,, \tag{A.3}$$

with initial condition $Q(0) = 0$. The corresponding solution calculated in T_k is covariance matrix Q_k, defined in (2.4). Equation (A.3) is a linear equation that may be rewritten in a more familiar form for n^2-dimensional vector

$$q \triangleq \mathrm{vec}(Q)\,,$$

where function vec$(\cdot)$ simply places the columns of matrix Q above each other, i.e.:

$$\mathrm{vec} : R^{n\times n} \to R^{n^2\times 1} : \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \mapsto \begin{bmatrix} a_{11} & \cdots & a_{n1} \,\big|\, \cdots \,\big|\, a_{1n} & \cdots & a_{nn} \end{bmatrix}^T.$$

In fact equation (A.3) is equivalent to the following one in variable q:

$$\dot{q} = \Phi q + s,\,, \qquad q(0) = 0\,, \tag{A.4}$$

where $s = \mathrm{vec}(GSG^T)$ and $\Phi \in R^{n^2\times n^2}$ is an appropriate matrix. Once the solution of this is known, Q_k is obtained in vector form as $\mathrm{vec}(Q_k) = q(T_k)$. Note that by (A.4) we may write

$$q(T_k) = e^{\Phi T_k} q(0) + \int_0^{T_k} e^{\Phi t} s\; dt = \int_0^{T_k} e^{\Phi t} s\; dt,$$

which may be calculate using MATLAB function `c2d.m`, which converts continuous linear state models into discrete ones.

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, New York, 1964.

[2] R. W. Brockett. *Finite Dimensional Linear Systems.* John Wiley & Sons, Inc., New York, 1970.

[3] G. Cariolaro and G. Pierobon. *Processi Aleatori.* Edizioni Libreria Progetto, Padova, Italy, third edition, 1994. In Italian.

[4] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons, 1991.

[5] E. Fornasini and G. Marchesini. *Appunti di Teoria dei Sistemi.* Edizioni Libreria Progetto, Padova, Italy, 1992. In Italian.

[6] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes.* Oxford University Press, Oxford, UK, third edition, 2001.

[7] M. I. Jordan and C. M. Bishop. *An Introduction to Graphical Models.* In preparation.

[8] M. J. Kampé de Fériet. *La fonction hypergéométrique.* Gauthier-Villars, Paris, France, 1937.

[9] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification, and Adaptive Control.* Prentice Hall, Englewood Cliffs, New Jersey, 1986.

[10] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, third edition, 1991.

[11] G. Picci. *Filtraggio di Kalman, Identificazione ed Elaborazione Statistica dei Segnali.* Edizioni Libreria Progetto, Padova, Italy, 1998. In Italian.

[12] W. Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, New York, third edition, 1976.

[13] J. B. Seaborn. *Hypergeometric Functions and Their Applications.* Springer-Verlag, New York, 1991.

[14] I. Stewart and D. Tall. *Complex Analysis*. Cambridge University Press, Cambridge, UK, 1983.