Random Sampling of a Continuous-time Stochastic Dynamical System

Mario Micheli[*‡] Michael I. Jordan[†]

[*]*Division of Applied Mathematics, Brown University, Providence, RI 02012, USA*
[‡]*Dipartimento di Elettronica e Informatica, Università di Padova, Padova, Italy*
[†]*Computer Science Division and Dept. of Statistics, UC Berkeley, Berkeley, CA 94720, USA*

`mariom@dam.brown.edu` `jordan@cs.berkeley.edu`

Abstract

We consider a dynamical system where the state equation is given by a linear stochastic differential equation and noisy measurements occur at discrete times, in correspondence of the arrivals of a Poisson process. Such a system models a network of a large number of sensors that are not synchronized with one another, where the waiting time between two measurements is modelled by an exponential random variable. We formulate a Kalman Filter-based state estimation algorithm. The sequence of estimation error covariance matrices is not deterministic as for the ordinary Kalman Filter, but is a stochastic process itself: it is a homogeneous Markov process. In the one-dimensional case we compute a complete statistical description of this process: such a description depends on the Poisson sampling rate (which is proportional to the number of sensors on a network) and on the dynamics of the continuous-time system represented by the state equation. Finally, we have found a lower bound on the sampling rate that makes it possible to keep the estimation error variance below a given threshold with an arbitrary probability.

1 Introduction

In this paper we briefly summarize the results described in [5] and [6], concerning the problem of state estimation for continuous-time stochastic dynamical systems in a situation where measurements are available at randomly-spaced time instants.

More specifically, we consider the following dynamical model:[1]

$$\begin{cases} \dot{x}(t) & = & Fx(t) + Gv(t) \\ y(t_k) & = & Cx(t_k) + z(t_k) \end{cases} \qquad t \in \mathbb{R},\, k \in \mathbb{N} \tag{1}$$

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 known time-invariant real matrices. In linear model (1) two different white,

[1]We will refer to the first equation in (1) as a *state equation*, and the second equation as a *measurement equation*. A formally correct way of writing the state equation would be:

$$dx = Fx\,dt + G\,dw,$$

which is the standard notation for stochastic differential equations (whose solutions are known as Itô processes, or diffusions [3]).

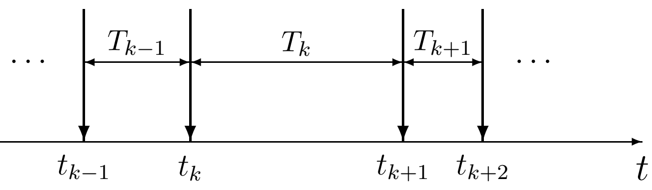

Figure 1: Random sampling process.

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

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

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 known constant positive definite matrices (in general, S may be just semipositive definite); we also assume that $v(\cdot)$ and $z(\cdot)$ are independent of each other.

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

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

are i.i.d. exponential random variables with known parameter λ, $T_k \sim \mathcal{E}(\lambda)$; i.e. the sampling is generated by a *Poisson process* [3] of intensity λ. 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 illustrates the random sampling process.

Given such a model (in which matrices F, G, C, S, R and intensity λ are known) we wish to estimate state $x(t)$ in an *on-line* manner at any time instant $t \in \mathbb{R}$, using the set of past measurements $\{y(t_k) : t_k \leq t\}$ and the knowledge of past Poisson arrivals $\{T_{k-1} : t_k \leq t\}$. At time t we only know the realization of the Poisson process *up to* time t, and the corresponding measurements. Note in particular that when $F = 0$ in (1) we are dealing with the problem of estimating randomly sampled *Brownian Motion*.

Application: Sensor Networks. The mathematical model we just described arises quite naturally in the analysis of sensor of networks. Assume in fact that the evolution of a physical process $x(t)$ may be described by the state equation in (1), and that a number N of identical sensors measure such process. Each one of them periodically yields a noisy measurement of $x(t)$ according to the measurement equation in (1) every T seconds. If the sensors are independent of each other and *not synchronized* then the process that is obtained by summing the arrivals of all sensors may be approximated, when N is large, by a Poisson Process with intensity $\lambda = N/T$; i.e., at any time instant the next arrival will occur after an exponential (hence memoryless) random time. For a more detailed and rigorous discussion of this aspect see [5] or [6].

Paper Summary. In the next section we describe a Kalman filter-based estimation algorithm that also yields, step by step, the covariance matrix of the estimation error. Such a matrix provides a measure of the effectiveness of our estimation algorithm. The sequence of estimation error covariance matrices is *stochastic* due to the random nature of the sampling process (as opposed to what happens in the case of ordinary Kalman filtering, where the same matrix sequence is deterministic): namely, it is a homogeneous Markov process. We also study the problem of estimating state $x(t)$ between two consecutive Poisson arrivals. We then give a brief description of the random parameters that appear in the discrete-time system obtained by sampling the state equation in correspondence with the Poisson arrivals.

In section 3, where our major results are described, we perform an analysis of the sequence of estimation error variances in the one-dimensional case. By exploiting the Markov property, we give a complete statistical description of such stochastic process: we study the "transition" conditional probability density, which plays the role of the transition matrix for a Markov chain. In particular, we analyze the subtle relation between the sampling rate, the (only) eigenvalue of state matrix F, and the estimation error variance.

Finally, in section 4 we briefly describe the possibility of bounding the estimation error variance below a given threshold with arbitrary probability, by an appropriate choice of sampling intensity λ. Note that when equations (1) model a network of sensors such intensity is proportional to the number of sensors ($\lambda \simeq N/T$, for large N): therefore choosing λ corresponds to picking an appropriate number of sensors. Section 5 is dedicated to final remarks and comments on possible future directions of reserach.

2 Estimation Algorithm

In order to estimate state x at the Poisson arrivals $\{t_k\}_{k\in\mathbb{N}}$ we consider the *sampled* version (see, e.g., [2]) of the state equation, where the samples are taken in correspondence with the Poisson arrivals.

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

$$\begin{cases} x(t_{k+1}) &= A_k x(t_k) + w(t_k) \\ y(t_k) &= C x(t_k) + z(t_k) \end{cases} \qquad k \in \mathbb{N}, \tag{2}$$

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

$$A_k = e^{F(t_{k+1}-t_k)} = e^{FT_k}, \tag{3}$$

and the vector

$$w(t_k) \;=\; \int_0^{T_k} e^{F\tau}\, G\, v(t_{k+1}-\tau)\, d\tau\,. \tag{4}$$

We should remark that A_k depends on random variable T_k, therefore it is a *random matrix.* Note that the randomness of noise $w(t_k)$ derives from its dependence from both continuous-time noise $v(t)$ and random variable T_k.

For on-line estimation purposes we are interested in calculating the mean and the 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; in other words, T_k is itself an observation on which we are basing our estimation. It is simple to verify that $\mathbb{E}\big[w(t_k)\,|\,T_k\big] = 0$. One can compute that the covariance matrix of $w(t_k)$ given T_k is given by:

$$Q_k \triangleq \mathbb{E}\big[w(t_k)w^T(t_k)\,\big|\,T_k\big] = \int_0^{T_k} d\tau\, e^{F\tau} GSG^T e^{F^T\tau}; \tag{5}$$

being a function of T_k, Q_k is a *random matrix* as well. In general, one can prove that random process $w(t_k)$, conditioned on $\{T_j\}_{j=1}^{\infty}$, is *white* Gaussian noise; in particular, $w(t_k)|T_k \sim \mathcal{N}(0, Q_k)$.

For a fixed T_k, solving integral (5) analytically is generally unfeasible. However, matrix Q_k may be obtained as the solution of the following linear matrix equation [1]:

$$\dot{Q}(t) = FQ(t) + Q(t)F^T + GSG^T, \tag{6}$$

with initial condition $Q(0) = 0$, calculated in T_k, i.e. $Q_k = Q(T_k)$. Equation (6) may be solved *numerically* on-line; see Appendix A of [5] for further details.

2.1 Estimation at Poisson arrivals

The natural way of performing state estimation for a discrete-time system like (2) is Kalman Filtering [4] [7]. However, one has to pay special attention to the fact that some of the parameters that are deterministic in ordinary Kalman Filtering are, in our case, *random*: namely, matrices A_k and Q_k. Also, time intervals $\{T_k\}$ are themselves *measurements*, as well as sequence $\{y(t_k)\}$.

In the light of this, define the following quantities[2] (note that, at time t_k, measurements up to $y(t_k)$ are known, whereas interarrival times up to T_{k-1} are known: refer to Figure 1):

$$\begin{aligned}
\hat{x}_{k|k} &\triangleq \mathbb{E}\big[x(t_k)\big|\{y(t_j), T_{j-1}\}_{j\le k}\big]\,, && (7)\\
P_{k|k} &\triangleq \mathrm{Var}\big[x(t_k)\big|\{y(t_j), T_{j-1}\}_{j\le k}\big]\,, && (8)\\
\hat{x}_{k+1|k} &\triangleq \mathbb{E}\big[x(t_{k+1})\big|\{y(t_j), T_j\}_{j\le k}\big]\,, && (9)\\
P_{k+1|k} &\triangleq \mathrm{Var}\big[x(t_{k+1})\big|\{y(t_j), T_j\}_{j\le k}\big]\,; && (10)
\end{aligned}$$

we should remark that estimator $\hat{x}_{k|k}$, as defined in (7), satisfies the following:

$$\hat{x}_{k|k} = \arg\min_{\hat{x}\in\mathcal{M}} \mathbb{E}\big[||x(t_k) - \hat{x}||^2\big],$$

where $\mathcal{M}$ is the set of all measurable functions of variables $\{y(t_j), T_{j-1}\}_{j\le k}$. An analogous property holds for $\hat{x}_{k+1|k}$, defined in (9). The corresponding Kalman Filter equations, which hold when the noise is Gaussian, are the following:

$$\begin{aligned}
\hat{x}_{k+1|k} &= A_k\hat{x}_{k|k} && (11)\\
P_{k+1|k} &= A_k P_{k|k} A_k^T + Q_k && (12)\\
\hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} + P_{k+1|k}C^T(CP_{k+1|k}C^T + R)^{-1}(y_{k+1} - C\hat{x}_{k+1|k}) && (13)\\
P_{k+1|k+1} &= P_{k+1|k} - P_{k+1|k}C^T(CP_{k+1|k}C^T + R)^{-1}CP_{k+1|k} && (14)
\end{aligned}$$

where we have written y_{k+1} instead of $y(t_{k+1})$. 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*.

We should now note the most significant differences between the above 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 $\{P_{k|k}\}_{k=0}^{\infty}$, which in the ordinary case is (for all k) completely deterministic and can be

[2]To be more precise, we should define $\hat{x}_{0|0} = \mathbb{E}[x(0)]$ and $P_{0|0} = \mathrm{Var}[x(0)]$; this way, definitions (7) and (8) are valid for $k \ge 1$, whereas (9) and (10) hold for $k \ge 0$.

computed off-line before measurements start, in our case is itself a *random process*. In fact, by using the independence hypotheses between $\{T_k\}_{k=1}^{\infty}$ and noise $v(t), t \in \mathbb{R}$, and the fact that since T_k's are iid and $v(t)$ is white and Gaussian, one can prove that $\{P_{k|k}\}_{k=0}^{\infty}$ is a *homogeneous Markov process*. See [5] and [6] for details.

Secondly, while in the ordinary case time update $k \to k+1$ (i.e. equations (11) and (12)) can be performed at time t_k, in the case of random sampling one has to wait time t_{k+1} since matrices A_k and Q_k are needed in equations (11) and (12): 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} (i.e. when T_k is known).

In section 3 we shall focus on the statistical description of stochastic process $\{P_{k|k}\}_{k=0}^{\infty}$ in the 1-D case; in particular, we will analyze its dependence on the continuous-time dynamics (represented by F) and Poisson sampling intensity λ.

2.2 Estimation between Poisson arrivals

Similar techniques may be applied to state estimation *between* two consecutive Poisson arrivals, i.e. when time elapses with no new measurements occurring. For this purpose, define:

$$\begin{aligned} \hat{x}_t &\triangleq \mathbb{E}\big[x(t) \,\big|\, \{y(t_j), T_{j-1} : t_j \leq t\}\big] , \\ P_t &\triangleq \mathrm{Var}\big[x(t) \,\big|\, \{y(t_j), T_{j-1} : t_j \leq t\}\big] . \end{aligned}$$

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

$$\hat{x}_t \equiv \hat{x}_{k|k} , \qquad P_t = P_{k|k} + GSG^T (t - t_k);$$

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

2.3 On matrices A_k and Q_k

Due to lack of space, we only briefly summarize the statistical description of the random matrices A_k and Q_k. These matrices have rather different behaviors according to the dynamics of the original continuous-time state equation in (1).

For example, in the 1-D case ($m = n = p = 1$), i.e. when both A_k and Q_k are random variables) if the (only) eigenvalue of matrix F is negative —that is, if the state equation in (1) is *stable*— then the *support* of the probability density of Q_k is bounded. On the other hand, if the original continuous-time system is *unstable* then such support is unbounded.

In fact, in this case random variable Q_k only has a finite number of finite moments: the n-th order moment $\mathbb{E}[Q_k^n]$ exists if and only if $\lambda > 2nF$, i.e. when the sampling intensity λ is high enough; in particular, when $\lambda \leq 2\phi$ random variable Q_k does not even have finite mean. Given the role that Q_k plays in equation (12) (and consequently in (14)), this suggests that in the case on unstable systems it will be harder to perform state estimation, since the Kalman Filter will tend to yield higher values for estimation error covariance matrix $P_{k|k}$. See [5] for a thorough discussion of this issue, including the analytical expressions of the probability densities of random variables A_k and Q_k.

3 Statistical Description of Estimation Error Variance

The effectiveness of the state estimation algorithm is described by $\{P_{k|k}\}_{k=0}^{\infty}$, i.e. the sequence of estimation error covariance matrices. We have seen already that it has the property of being a homogeneous Markov process.

Assume that the probability density 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 the distribution function:[3]

$$F^{(k+1)}(p_{k+1}, p_k, \ldots, p_1, p_0) \;=\; \mathbb{P}\big[\, P_{k+1|k+1} \leq p_{k+1}, \ldots, P_{0|0} \leq p_0 \,\big];$$

thanks to the Markov property we may write the corresponding probability density as follows:[4]

$$f^{(k+1)}(p_{k+1}, p_k, \ldots, p_1, p_0) \;=\; f_{k+1|k}(p_{k+1}|p_k) \cdot f_{k|k-1}(p_k|p_{k-1}) \cdot \ldots \cdot f_{1|0}(p_1|p_0) \cdot f_0(p_0), \quad (15)$$

where the meaning of symbols should be obvious. Since $\{P_{k|k}\}_{k=0}^{\infty}$ is a *homogeneous* Markov process each of the above conditional densities $f_{j+1|j}(\,\cdot\,|\,\cdot\,)$, $0 \leq j \leq k$, 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:

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

where[5]

$$f(p|q) \;\triangleq\; \frac{\partial}{\partial p}\; \mathbb{P}\big[\, P_{j+1|j+1} \leq p \,\big|\, P_{j|j} = q \,\big] \quad (16)$$

does not depend on index j but only on the value assumed by matrices p and q. Note than since process $\{P_{k|k}\}_{k=0}^{\infty}$ is homogeneous Markov, the knowledge of the above conditional probability density (together with prior density f_0) gives a complete statistical description of such process.

In [5] and [6] we compute an analytical expression for (16) in the one-dimensional case. We shall describe our results, starting with the *support* (with respect to variable p) of the

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

[4]We indicate with $f_0(\cdot)$ 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 Dirac delta function.

[5]Here symbol $\partial/\partial p$ indicates differentiation with respect to each single element of matrix p (in fact, we are performing n^2 differentiations). In the one-dimensional it just indicates ordinary partial differentiation with respect to variable p.

conditional density above.[6] We will then give the explicit expression for $f(p|q)$ and illustrate some density plots. In particular, we will analyze how the expression of conditional density (16) is very closely related on the stability of the original continuous time system (1) and on the intensity λ of the Poisson sampling process, sometimes in quite a surprising manner.

3.1 Conditional probability density support

Let us first introduce some notation for the one-dimensional case ($m = n = p = 1$). We will use scalar quantities: $\phi = F$, $g = G$, $\sigma^2 = S$. 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. Note that when $\phi < 0$ continuous-time dynamical system $\dot{x} = \phi x$ is asymptotically stable, when $\phi = 0$ it is simply stable, when $\phi > 0$ it is unstable; we shall refer to these cases later on.

In this subsection we shall talk about the support of conditional density function (16); in the next one we will give its explicit expression and illustrate some density plots. First of all note that, for a fixed q, the support of $f(\cdot|q)$ has to be contained in interval $[0, R/C^2]$. In fact the second of equations (1), i.e. $y(t_k) = Cx(t_k) + z(t_k)$, implies that the estimation error is in the worst case determined by noise $z(t_k)$, whose covariance matrix is R. If we did not know past history the probability density $f_k(p_k)$ of $P_{k|k}$ would be a delta function centered at R/C^2; in general we can do better than this since we are also using past information, hence the probability density $f_k(p_k)$ of $P_{k|k}$ is "spread" on interval $[0, R/C^2]$. Since, for all k, the probability density $f_{k+1}(p_{k+1})$ of $P_{k+1|k+1}$ may be computed as

$$f_{k+1}(p_{k+1}) = \int_0^{R/C^2} f(p_{k+1}|p_k) f_k(p_k)\, dp_k\,,$$

the support (in the p variable) of $f(p|q)$ must be contained in interval $[0, R/C^2]$ as well.

The support of $f(\cdot|q)$ depends on the system parameters; however, it does *not* depend on sampling intensity λ, which only influences the *shape* of $f(p|q)$ within its support. See Figure 2, which we will now explain in detail.

In general, we will have that the support of $f(\cdot|q)$ is an *interval* contained in $[0, R/C^2]$. A "singular" case occurs when eigenvalue ϕ is equal to the following *negative* number:

$$\phi^* \triangleq -\frac{g^2\sigma^2}{2q}$$

in which case $f(p|q)$ is a Dirac delta function centered at $p = \frac{R}{C^2} q\big(q + \frac{R}{C^2}\big)^{-1}$, independently of sampling intensity λ. We shall refer to this "singular" case as **Type II** conditional density. We should note that, for a fixed ϕ, this case occurs when $\phi = -\frac{g^2\sigma^2}{2q}$, which happens *with probability zero* since $P_{k|k}$ is a continuous random variable.

[6]Since density functions that are relative to absolutely continuous probability distributions are defined in the "almost everywhere" sense, we should specify what we mean by support. A right-neighborhood of x is a set of the type $\{y : x \leq y < \epsilon\}$, for some $\epsilon > 0$; a left-neighborhood is defined in an analogous manner. We will say that x belongs to the support of probability density $f(\cdot)$ if there exists a right-neighborhood *or* a left-neighborhood I_x of x such that $f(y) > 0$ for all $y \in I_x$. When density $f(\cdot)$ has Dirac delta functions (i.e. the corresponding random variable has a discrete component) we shall add the corresponding points to the support.

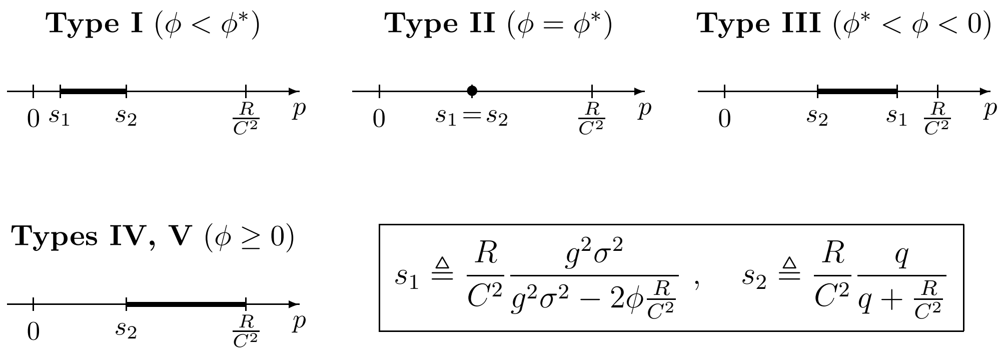

Figure 2: Support of conditional probability density $f(p|q)$ with respect to variable p. Note that it does not depend on sampling intensity λ.

It will be convenient to define the following quantities:

$$s_1 \triangleq \frac{R}{C^2} \frac{g^2\sigma^2}{g^2\sigma^2 - 2\phi\frac{R}{C^2}}, \qquad s_2 \triangleq \frac{R}{C^2} \frac{q}{q + \frac{R}{C^2}}, \tag{17}$$

which in most cases determine the *extremes* of the support of $f(\cdot|q)$. Note that s_1 is a function of eigenvalue ϕ and *not* a function of the previous estimation error variance value q; on the other hand, s_2 is a function of q and *not* a function of ϕ; however, none of them is a function of sampling intensity λ. Note also that when $\phi = \phi^*$ (Type II densities) we have that s_1 and s_2 coincide, and$f(\cdot|q)$ is a delta function centered at $s_1 = s_2$.

When $\phi < \phi^*$ (which, since $\phi^* < 0$, implies a relatively high degree of stability for the state equation of the continuous-time state equation) we have that $s_1 < s_2$ and the support of $f(\cdot|q)$ is in interval $[s_1, s_2]$. In such case we will talk of **Type I** conditional density functions; it is the only instance when $s_1 < s_2$. Note that as ϕ, which is negative, decreases (i.e. as we consider systems that are more and more stable) the value of s_1 decreases too, meaning that the variance of $P_{k+1|k+1}$ may assume lower values. This is in accordance with the idea that *stable* dynamical systems are easier to track.

When $\phi^* < \phi < 0$ we have **Type III** conditional densities. In this case $s_2 < s_1 < \frac{R}{C^2}$, and the support of $f(\cdot|q)$ is in interval $[s_2, s_1]$; for increasing values of ϕ we have that s_1 approaches R/C^2, i.e. as we approach instability 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 case corresponds to what we will call **Type IV** distributions.

Finally, when $\phi > 0$ (unstable systems) 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 the conditional density.

It is important to remark that s_2 is a monotone increasing function of q, it is equal to zero for $q = 0$ and converges to R/C^2 as $q \to \infty$; however, since q can only assume values in interval[7] $[0, R/C^2]$, we necessarily have that $s_2 \in [0, \frac{1}{2}R/C^2]$. Since q is the value previously assumed by the estimation error variance, it is a measure of the reliability of the

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

previous state estimates. The fact that s_2 is an increasing function of q means that if the previous estimate 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.

3.2 Conditional probability density plots

We will now explicitly show the expression for conditional density $f(\cdot|q)$ in the cases described in the previous subsection. Since Type II densities are trivial, we will not discuss them. We will also plot the conditional density is several significant cases. The following parameters are common to all plots: $R = 4$, $C = 1$, $g = 1$, $\sigma^2 = 1$, $q = 3$, so that $R/C^2 = 4$ and $\phi^* = -g^2\sigma^2/(2q) \simeq -0.16667$; the particular values of ϕ and λ are reported on each graph.

Type I densities ($\phi < \phi^*$). The explicit form of conditional probability density (16) is the following [5, §3.2]:

$$f(p|q) \;=\; -\frac{\lambda}{2\phi}\frac{R^2}{C^4}\left(q + \frac{g^2\sigma^2}{2\phi}\right)\frac{\left[\left(\frac{R}{C^2} - p\right)\left(q + \frac{g^2\sigma^2}{2\phi}\right)\right]^{\frac{\lambda}{2\phi}-1}}{\left[\frac{g^2\sigma^2}{2\phi}\frac{R}{C^2} + p\left(\frac{R}{C^2} - \frac{g^2\sigma^2}{2\phi}\right)\right]^{\frac{\lambda}{2\phi}+1}} \cdot 1(s_1 < p < s_2), \qquad (18)$$

where $1(\cdot)$ is the indicator function. Note that the minus sign in front of it makes sense since the third factor (in round parentheses) is negative for $\phi < \phi^*$. Type I densities are shown are shown in Figure 3 (for a fixed sampling intensity $\lambda = 2$ and different values of $\phi < \phi^*$) and in Figure 4 (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 (17)) 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! 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. 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 poor (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. We will briefly return on this case in section 3.3 and provide further clarification.

Type III densities ($\phi^* < \phi < 0$). The formula for $f(p|q)$ is in this case the same as in expression (18), except that it does *not* have the minus sign in front and the multiplying indicator function is $1(s_2 < p < s_1)$, i.e. the support is $[s_2, s_1]$. Type III densities are plotted in Figure 5 (for a fixed sampling intensity $\lambda = 0.5$ and different values of ϕ, with $\phi^* < \phi < 0$) and in Figure 6 (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, whereas s_2 is unchanged by ϕ. In other words, the

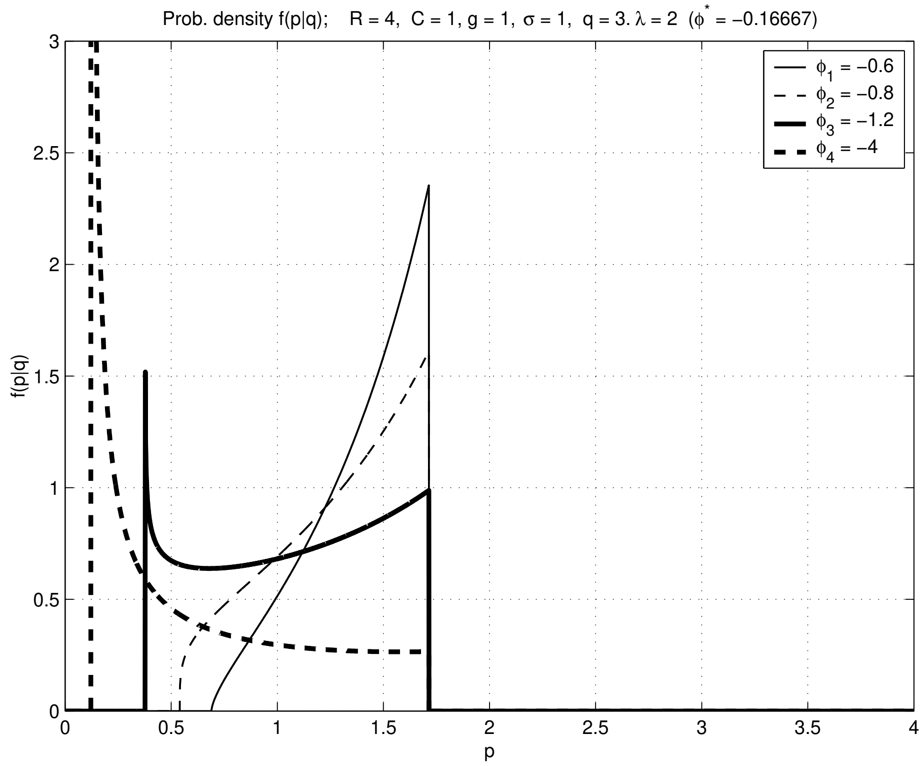

Figure 3: Type I densities, for a fixed value of λ.

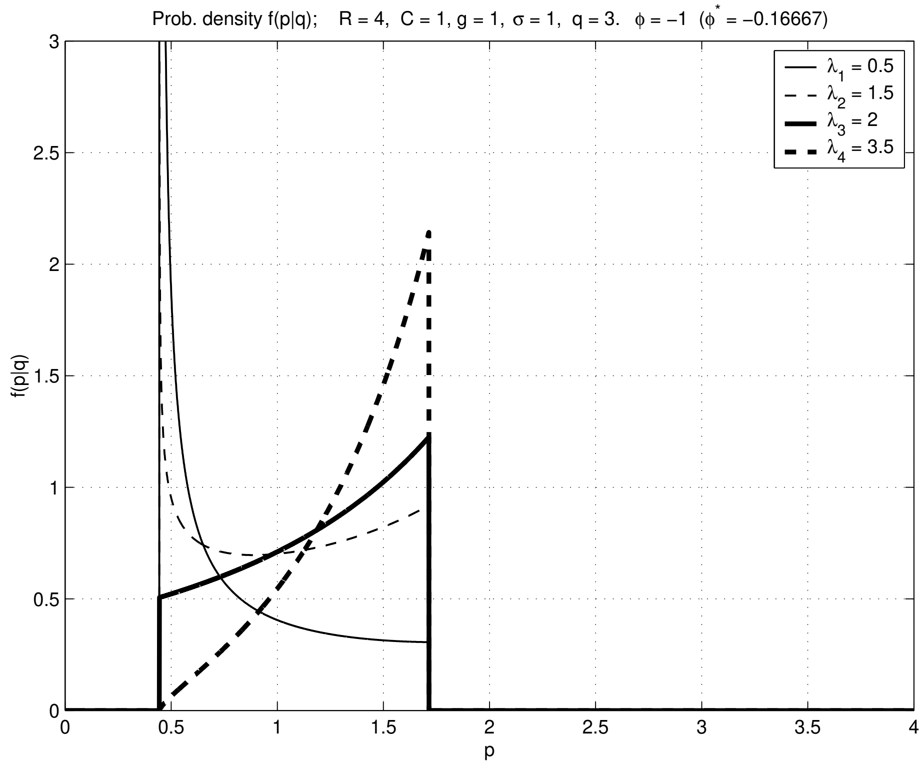

Figure 4: Type I densities, for a fixed value of ϕ (less than ϕ^*).

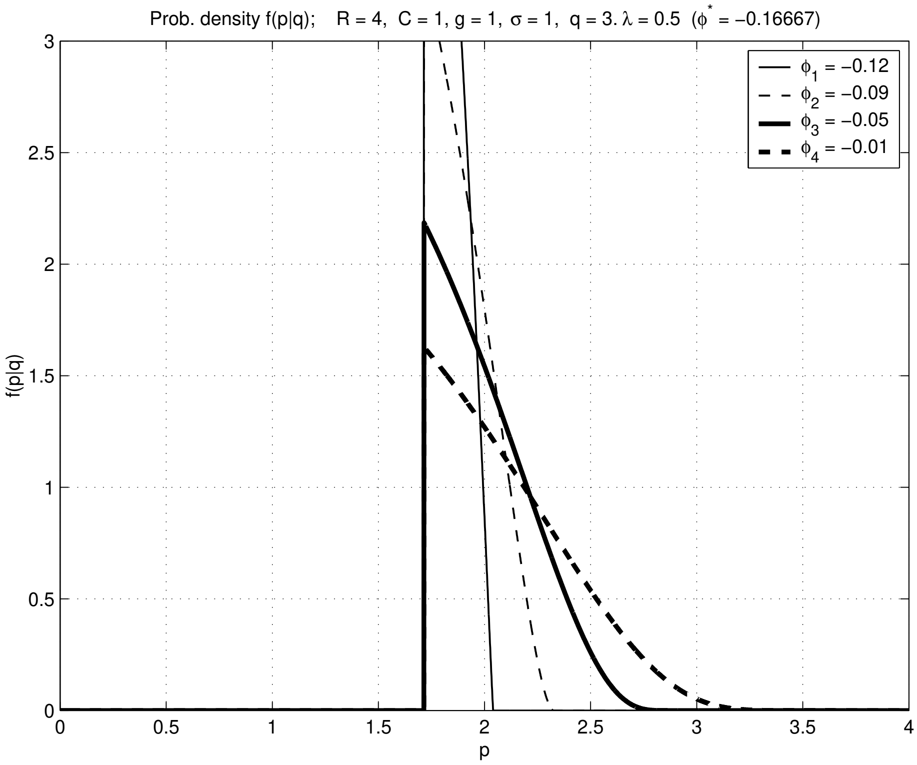

Figure 5: Type III densities, for a fixed value of λ.

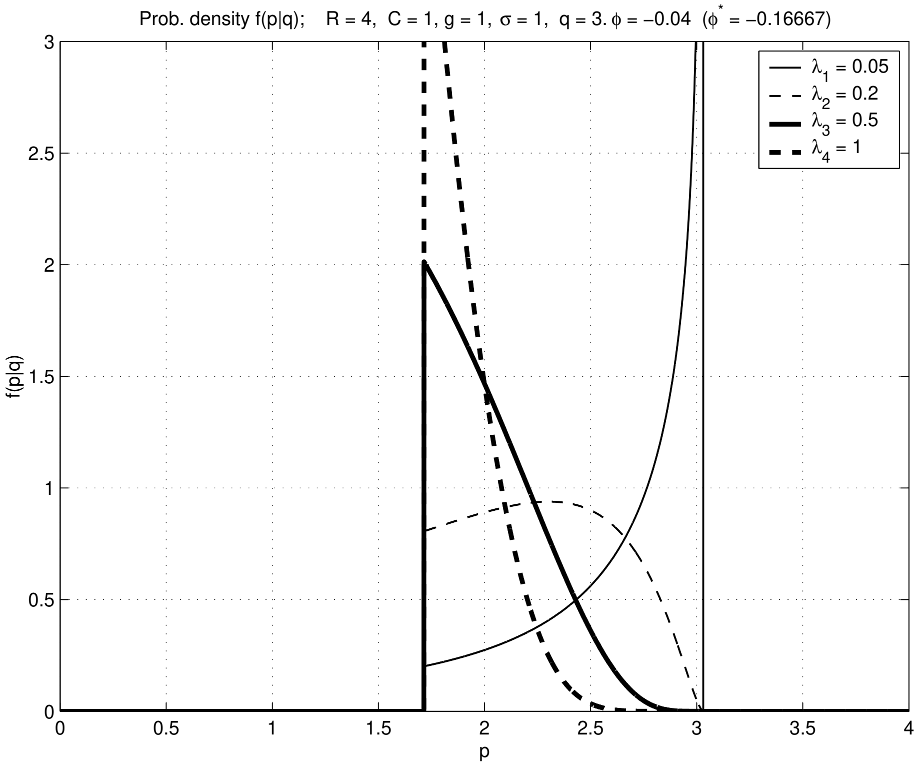

Figure 6: Type III densities, for a fixed value of ϕ.

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 intensities λ then we have the situation depicted in Figure 6, which does not show the paradoxical behavior that is typical of Type I densities, thus being closer to intuition: higher sampling intensities reduce the estimation error variance by shifting the area below the graph of $f(p|q)$ to the left.

Type IV densities ($\phi = 0$, Brownian motion). In this situation the expression for $f(p|q)$ is the following:

$$f(p|q) \;=\; \frac{\lambda}{2\varphi}\frac{R^2}{C^4}\left(\frac{R}{C^2}-p\right)^{-2}\exp\left[-\frac{\lambda}{g^2\sigma^2}\left(\frac{R}{C^2}\frac{p}{\frac{R}{C^2}-p}-q\right)\right]\cdot 1\left(s_2 < p < \frac{R}{C^2}\right).$$

We will not report any graph of this case here, due to lack of space; see [5], [6] for plots and more details. We will just say that the corresponding graphs are similar to the ones reported in Figure 6, except that support is given by interval $[s_2, R/C^2]$.

Type V densities ($\phi > 0$). In this case $f(p|q)$ has the same expression as in (18), except that (as for Type III densities) there is no minus sign, while the multiplying indicator function is $1\big(s_2 < p < \frac{R}{C^2}\big)$. These densities, that correspond to the random sampling of *unstable* dynamical systems, are are shown in Figure 7 (for a fixed sampling intensity $\lambda = 1$ and different values of $\phi > 0$) and in Figure 8 (for a fixed eigenvalue $\phi = 0.2$ and different sampling intensities). The support of $f(p|q)$ is given by $[s_2, R/C^2]$, independently of the (positive) value of ϕ (and of sampling intensity λ).

As Figure 7 illustrates, the more unstable the continuous-time system is, the hardest it is to track its state. On the other hand, Figure 8 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. In other words performing state estimation of an unstable system is easier when such system is sufficiently slow, or when measurements occur sufficiently often.

3.3 On the behavior of $f(\cdot|q)$ for $\phi < 0$

It is appropriate, at this point, to spend a few more words on the asymptotically stable case ($\phi < 0$); such case, as we have shown above, presents certain subtleties (characteristic of Type I densities) that do not appear when $\phi \geq 0$.

For a given one-dimensional system eigenvalue ϕ is fixed, which implies that extreme s_1 is fixed. On the other hand $P_{k|k}$ varies in time, hence the value that q assumes changes as well: this means that $\phi^* = -g^2\sigma^2/(2q)$ is variable in time too. Therefore if $\phi < 0$ we may have density Types I, II or III at different times. Assume for instance that $P_{k|k} = q$, such that $\phi < \phi^*$ (i.e. 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 $\phi > -g^2\sigma^2/(2p)$, which is the "new" value of ϕ^* (in fact p plays the role of q in the *subsequent* step: see the beginning of section 3); so density

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

with $p_{k+1} = p$, will be of Type III.

In fact we have shown rigorously [5] [6] that, for fixed values of $\phi < 0$ and λ, if a Type I density $f(p_{k+1}|p_k)$ occurs in formula (15) then it will sooner or later "turn" into a Type III density and stick to that type. More rigorously speaking, with probability one there exists a

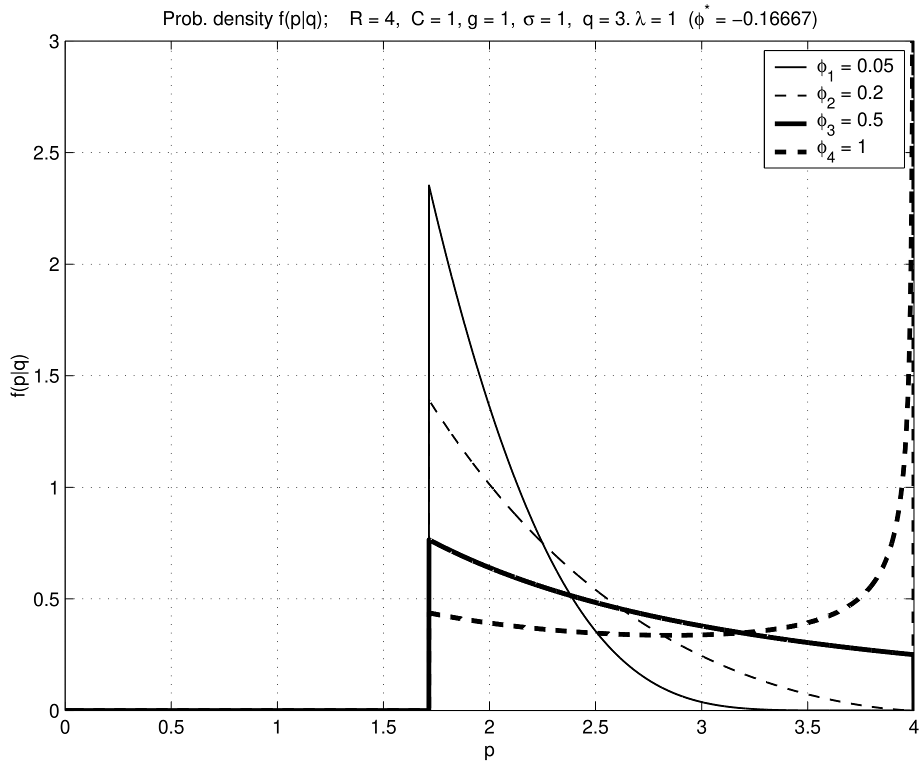

Figure 7: Type V densities, for a fixed value of λ.

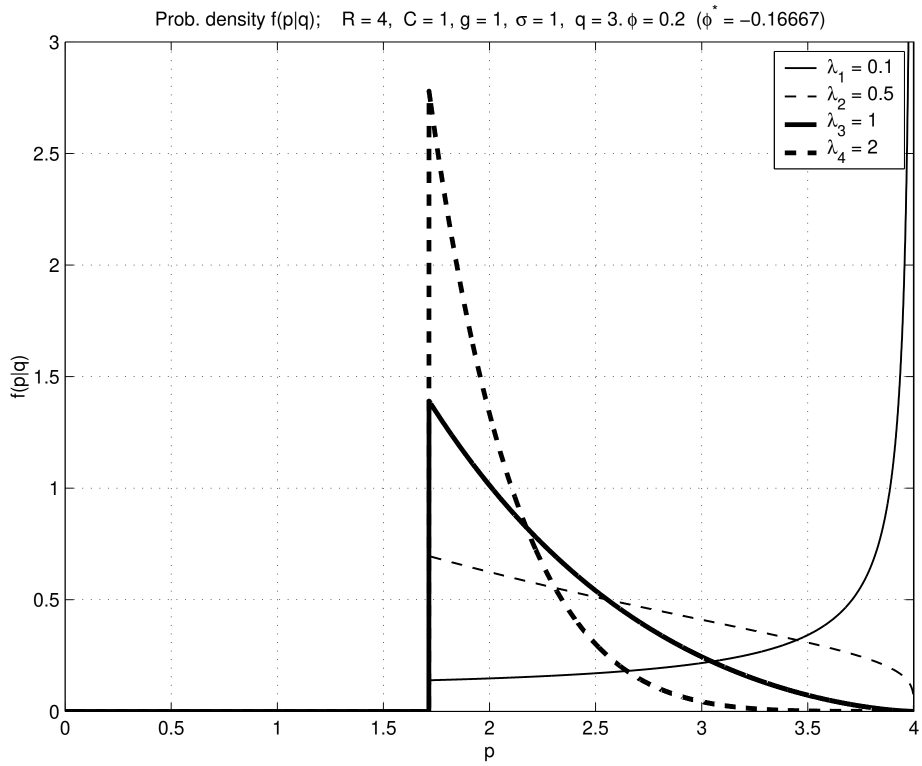

Figure 8: Type V densities, for a fixed value of ϕ.

index $j > k$ such that for all $\ell \geq j$ we will have $\phi^* = -g^2\sigma^2/(2p_\ell) < \phi$, so that all densities $f(p_{\ell+1}|p_\ell)$ will be of Type III; such transition occurs in a random time whose mean is finite. 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 previously. As the state approaches zero, Type I densities turn into Type III densities.

4 Bounding the Estimation Error

Recall the motivating example of a network with 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: $\lambda \simeq N/T$.

In the previous section we have seen how Poisson intensity λ has a direct influence on the performance of the state estimation algorithm we formulated in section 2, since the shape of conditional probability density (16) depends on the intensity of the sampling process (sometimes in a very subtle way: think of Type I probability densities).

In [5], [6] we have given an answer (although not an optimal one) to the following problem: "Let ϕ be the eigenvalue of system (1), fix an arbitrary probability $\alpha \in (0, 1)$ (close to 1) and an arbitrary estimation error variance p^*; find a sampling intensity λ^* such that: $\mathbb{P}\big[P_{k|k} \leq p^*\big] > \alpha, \forall k \in \mathbb{N}$, for any choice of $\lambda > \lambda^*$." The answer we have found depends on the sign of eigenvalue ϕ (there is also a solution for $\phi = 0$). For example, here is the proposition we have proved in the case of unstable systems.

Proposition. *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:*

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

then we shall have that

$$\mathbb{P}\big[P_{k|k} \leq p^*\big] > \alpha\,, \qquad \forall k \in \mathbb{N}\,, \tag{19}$$

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

Similar results hold for the case of Brownian motion ($\phi = 0$) and for asymptotically stable systems ($\phi < 0$); in the latter case lower values of p^* are admissible, however for (19) to hold one has to wait until conditional probability densities $f(p_{k+1}|p_k)$ "become" of type III (for k large enough: as noted in the previous section, this occurs in finite mean time). We should also note that our results are not optimal, in the sense that they provide *sufficient* but *not necessary* conditions for (19) to hold. We strongly suspect there may be lower values of λ^* that would imply the same expression. The refinement of our results is left for future work.

5 Conclusions and Future Work

In this paper we have presented recent work on random sampling of continuous-time stochastic dynamical systems. We have provided a Kalman-based state estimation algorithm and we have performed (in the 1-D case) a complete statistical description of the corresponding estimation error variance process. In particular, we have discussed the dependence of such a description on the dynamics of the original continuous-time system and on the intensity of the Poisson sampling process. Finally, we have briefly presented a result that makes it possible to bound the error variance of the state estimation process by choosing a suitable Poisson sampling intensity.

Research in this field has a number of natural future directions. The most prominent one is the extension of our study to multi-dimensional, linear dynamical systems: 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 phenomena. Also, in analogy with Markov chains, it is rather intuitive that there should be a *stationary* probability density $\pi(\cdot)$, such that, if we started with an error variance $P_{0|0}$ distributed according to it, the probability density of $P_{k|k}$ would not change in time; we find the problems of proving its existence or, more ambitiously, of finding $\pi(\cdot)$ in analytic form to be quite interesting ones. Another challenging problem would be the one of refining the results we briefly presented in section 4, regarding the bounding of the estimation error variance below a given threshold with arbitrary probability.

6 Acknowledgement

Research was supported by ONR MURI grant N00014-00-1-0637 and the Darpa SensIT program.

References

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

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

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

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

[5] M. Micheli. *Random Sampling of a Continuous-time Stochastic Dynamical System: Analysis, State Estimation, and Applications.* Master's Thesis, Department of Electrical engineering and Computer Sciences, UC Berkeley, 2001.

[6] M. Micheli and M. I. Jordan. Random sampling of a continuous-time stochastic dynamical system: Analysis, state estimation, and applications. Submitted to Journal of Machine Learning Research.

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