Robust Model-Free Software Sensors for the HIV/AIDS Infection Process

Hussain Alazki[1], Alexander Poznyak[2]

[1]Universidad Autonoma de Cuidad del Carmen, Campeche, Mexico
[2]CINVESTAV, DF, Mexico City, Mexico
Email: khussinaz@yahoo.com, apoznyak@ctrl.cinvestav.mx

Abstract

This paper considers the problem of the HIV/AIDS Infection Process filtering characterized by three compounds, namely, the number of healthy T-cells, the number of infected T-cells and free virus particles. Only the first and third of them can be measurable during the medical treatment process. Moreover, the exact parameter values are admitted to be also unknown. So, here we deal with an uncertain dynamic model that excludes the application of classical filtering theory and requires the application of robust filters successfully working in the absence of a complete mathematical model of the considered process. The problem is to estimate the number of infected T-cells based on the available information. Here we admit the presence of stochastic "white noise" in current observations. To do that we apply the Luenberger-like filter (software sensor) with a matrix gain, which should be adjusted at the beginning of the process in such a way that the filtering error would be as less as possible using the Attractive Ellipsoid Method (AEM). It is shown that the corresponding trajectories of the filtering error converge to an ellipsoidal set of a prespecified form in mean-square sense. To generate the experimental data sequences in the test-simulation example, we have used the well-known simplified HIV/AIDS model. The obtained results confirm the effectiveness of the suggested approach.

Keywords

HIV/AIDS Infection Model, Robust Filter, Stochastic System

1. Introduction

Many aspects of phenomena critical to our lives cannot be measured directly. Fortunately, models of these phenomena, together with more limited observations frequently allow us to make reasonable inferences about the state of the

systems that affect us. The process of using partial observations and a stochastic model to make inferences about an evolving system is known as stochastic state estimation (or filtering). In this paper, we consider the problem of the HIV/AIDS Infection Process filtering characterized by three compounds: the number of healthy T-cells, the number of infected T-cells and free virus particles. Only the first and third of them can be measurable during the medical treatment process. The problem is to estimate the number of infected T-cells (to create a software sensor) based on the available information. Here we admit the presence of stochastic "white noise" in current observations as well as in the dynamics of other components.

1.1. HIV/AIDS Infection Process

Human Immunodeficiency Virus (HIV) stands for human immunodeficiency virus. If left untreated, HIV can lead to the disease AIDS (acquired immunodeficiency syndrome). Unlike some other viruses, the human body can't get rid of HIV completely. So once you have HIV, you have it for life. That's why the problems of HIV/AIDS are very important from medical and human points of view. HIV attacks the body's immune system, specifically the CD4 cells (T-cells), which help the immune system fight off infections. If left untreated, HIV reduces the number of CD4 cells (T-cells) in the body (directly and indirectly destroys CD4 + T-cells), making the person more likely to get infections or infection-related cancers. Over time, HIV can destroy so many of these cells that the body can't fight off infections and disease. These opportunistic infections or cancers take advantage of a very weak immune system and signal that the person has AIDS, the last state of HIV infection. The medicine used to treat HIV is called antiretroviral therapy or ART. If taken the right way, every day, this medicine can dramatically prolong the lives of many people with HIV, keep them healthy, and greatly lower their chance of transmitting the virus to others. Today, a person who is diagnosed with HIV, treated before the disease is far advanced, and stays in treatment can live a nearly as long as someone who does not have HIV.

The dynamic HIV/AIDS have been studied by many researchers (see, for example, [1] [2]). Experimental data show that the treated disease progression rate is significantly varied between individuals: from two weeks up to 20 years. Figure 1 shows the natural history of HIV infection dynamics. One can see that if an individual has been received HIV virus as primary infection, a number of HIV virus will dramatically increase in the first 30 days (resulting CD4 + T-cell reduction). Then after the primary infection period, a body builds HIV antibodies for agent virus so that, the infection still stabilizes an approximate steady state. In fact, up to now no effective cure for HIV exists, but nevertheless by some proper treatment and medical care, HIV can be controlled. To realize these treatments on-line information data are extremely required. The special physic equipment for these on-line measurements is very expensive and unavailable in many health clinics. So, the problem of designing the cheap and easy realizable
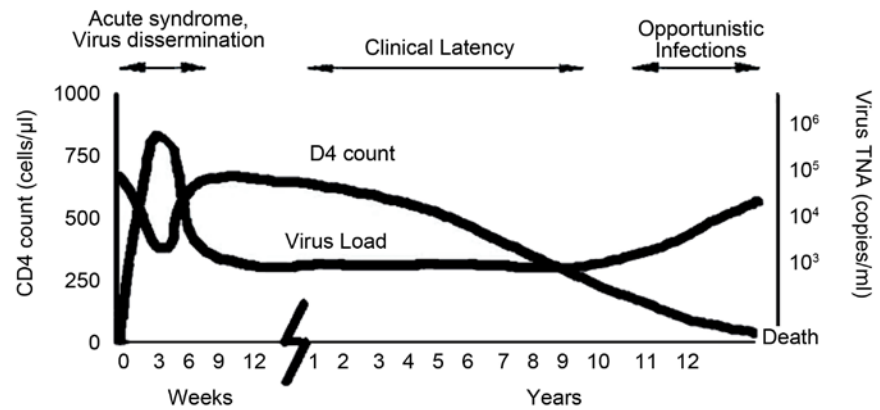
Figure 1. The HIV infection dynamics.

on-line *software sensors* seems to be very actual.

The model is proposed to include the activation process of CD4 + T-cells and their intervention in the HIV infection dynamics. Let T be the average number of CD4 + T-cells at time t, T^a the average number of CD4 + T-cells activated and specialized in identifying viral proteins, T^* the average number of infected CD4 + T-cells at time t and V the average HIV concentration at time t. Assume that cells increase at a constant rate σ and die in proportion to their size with death rate μ, so the average number of cells that die at time t is given by μT. Similarly, the specialized CD4 + T-cells are activated at a mean dTV, a term that describes the encounter of the T cells with infectious viral particles. It is assumed that specialized CD4 + T-cells also die at an average μT^a. If the population of Tes cells infected by the virus with a probability β, the principle of mass action allows to establish that the average number of cells that are infected at a time t is βTV, corresponding to a classical consideration in this type of models. In the same way CD4 + T-cells Specialized T^a are susceptible to infection and are infected at an average of ψT^a at a time t.

Based on the results of ([1] [2]) we may conclude that the current of an AIDS person state at time $t \geq 0$ can be characterized by the following three compounds:

- $x_1(t)$ -the number of healthy white blood cells (known as T-cells),
- $x_2(t)$ -the number of infected T-cells,
- $x_3(t)$ represents free virus particles (viruses are not classified as living organisms because they can not replicate without a help of a host cell).

The states $x_1(t)$ and $x_3(t)$ are measurable during the treatment process, but $x_2(t)$ not.

Problem. *Based on the available measurements $\{x_1(\tau), x_3(\tau) : \tau \in [0,t]\}$ obtain the state estimate $\hat{x}_2(t)$ which would be sufficiently closed (in some probabilistic sense) to the real state $x_2(t)$. In other words our aim is to design a, so-call, software sensor for the on-line estimation (or filtering) the unmeasured coordinate $x_2(t)$.*

Remark 1. Notice that any real mathematical model describing the exact behavior of the state vector $x(t) := (x_1(t), x_2(t), x_3(t))^\top$ is not available now.

Some simplified mathematical model, given in [1], we will use here to generate $x(t)$ during the test-numerical simulation but not for the filter designing.

1.2. Briefly on the Most Popular Filtering Methods

All classical filtering methods require the exact knowledge of the dynamic model which components are intended to be estimated. The main of them are as follows.

The Wiener (*frequency domain*) *filtering*. The origins of the filtering problem in discrete time can be traced back to the works [3] and [4]. In the continuous time case the first analysis of the optimal state estimation of dynamic systems in the presence of noise was given in [5]. The results were included in a classified National Defense Research Council report issued in January/February 1942. Subsequently declassified, it appeared as a book in 1949.

The Kalman-Busy-Stratonovich (*time-domain*) *filtering*. The next major development in stochastic filtering was the introduction of the *linear filter*. In this case, the signal satisfies a stochastic differential equation with linear coefficients and Gaussian initial condition. The linear filter can be solved explicitly in a finite-dimensional format: the distribution of the state estimate is shown to be Gaussian, and hence completely determined by its mean and its covariance matrix. These were the reasons for the linear filter's widespread success in the 1960s. Bucy and Kalman were the pioneers in this field. Kalman was the first to publish in a wide circulation journal. In [6], he solved the discrete time version of the linear filter. Bucy obtained similar results independently [7]. Stratonovich suggested the same filter, but for another type if stochastic integrals [8].

The extended Kalman filter (*EKF*). Following the success of the linear filter, scientists started to explore different avenues. Firstly they extended the application of the Kalman filter beyond the linear/Gaussian framework. The basis of this extension is the fact that, locally, all systems behave linearly. So, at least locally, one can apply the Kalman filter equation. This gave rise to a class of algorithm called the *extended Kalman filter* [9]. At the time of writing these algorithms, most of which are empirical and without theoretical foundation, are still widely used in a variety of applications.

The ensemble Kalman filter (*EnKF*). The progress in data assimilation is related with both increased computational power and the introduction of techniques that are capable of handling large amounts of data and more severe nonlinearities. The *EnKF* has been introduced to petroleum science recently [10] and, in particular, has attracted attention as a promising method for solving the history matching problem.

The singular evaluative interpolated Kalman filter (*SEIKF*). Inherent data and model uncertainties render the history-matching inverse problem extremely non-unique. Therefore, a reliable uncertainty quantification framework for predicting dynamic performance requires multiple models that match field production data. An efficient variant of the ensemble Kalman filter, namely, *Singular Evaluative Interpolated Kalman Filter* (*SEIKF*) is applied to the multi-

model history-matching problem [11].

Advanced developments in the time domain filtering. In the mid-1960s in [12], [13] there were derived and analyzed the stochastic models using the Itô (and not Stratonovich) calculus. In [14] there was provided the first rigorous derivation in the case of a general observation process where the signal and observation noises may be correlated. In 1968, in [15] there was introduced the innovation approach to linear filtering. This new method for deducing the filtering equations was extended in the early 1970s in [16] and [17]. Similarly, the other type of the filtering equation was introduced in the same period in [18] [19] [20] [21], is consequently referred to as the *Zakai* or the *Duncan-Mortensen-Zakai* equation. The stochastic partial differential equations (SPDEs) associated with the filtering equations were rigorously analyzed and extended in the late 1970s in [2] [22] [23]. This research was continued in [1] [24] [25]. Much of the work carried out in the 1990s has focussed on the numerical solution of the filtering problem.

H^∞ *stochastic filtering.* The problem of applying H^∞ filters on stationary, continuous-time, linear systems with stochastic uncertainties in the state-space signal model was considered in [26] [27] [28].

1.3. Main Contribution of the Paper

In this paper we follows the approach suggested in [29] [30].

- This paper considers the problem of designing a robust model-free filter for a respectively wide class of uncertain nonlinear stochastic system where classical filtering theory can not be applied.
- The behavior of these systems is given in Itô form and contains both a regular part, which assumed to be the Quasi-Lipschitz type but unknown exactly, as well as a stochastic part generated by a standard vector Wiener process.
- Filtering itself is suggested to be realized by a Luenberger-like filter with a matrix gain which should be adjusted in the beginning of the process in such a way that the filtering error would be as less as possible.
- It is shown that the corresponding trajectories of the filtering error converge (in the mean-square sense) to an ellipsoidal set of a prespecified form.
- We show that the HIV/AIDS infection process can be effectively realized by the suggested model-free technique based on the, so-called, *Attractive Ellipsoid Method* [31].

2. Filtering Problem for a Class of Nonlinear Systems and Its Solution

2.1. Class of Possible Dynamics for the Main State Variables of HIV Dynamics

The HIV dynamics (human immunodeficiency virus (HIV) causes the acquired immune deficiency syndrome, knows as AIDS) were analyzed in [32] [33] [34]. It captures the time rate of healthy, infected white blood cells (T-cells) and the

number of HIV viruses. The presented analysis is limited in the sense that it "cannot take into the account patient factors as physiological/genetic level, physicochemical factors at cell-protein-viral interactions level, and viral factors that relate to the various HIV strains and clades". Based on the corresponding numerical data it is possible to conclude that the dynamics of the HIV/AIDS Infection Process can be completely (without consideration the limiting factors mentioned above) described by the following systems of stochastic differential equations:

$$\dot{x}_t = f(x_t, t) + \sigma \mathcal{N}_t, x_0 = \overline{x}_0$$
$$\dot{y}_t = g(x_t, t) + \eta \mathcal{M}_t \tag{1}$$

where

$x_t = (x_{1,t}, \cdots, x_{n,t})^\top \in \mathbb{R}^n$ is the state vector of the system in time $t \geq 0$ which *may contain unmeasurable components*,

$y_t = (y_{1,t}, \cdots, x_{p,t})^\top \in \mathbb{R}^p$ is the output vector of the system in time $t \geq 0$ which is completely measurable components (available),

the vector functions $f : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^p$ and the matrices $\sigma : \mathbb{R}^m \to \mathbb{R}^n$ and $\eta : \mathbb{R}^k \to \mathbb{R}^p$ are nonlinear *mappings which are admitted to be exactly known a priory*,

$\mathcal{N}_t \in \mathbb{R}^m$ and $\mathcal{M}_t \in \mathbb{R}^k$ are standard vector white noises with independent components, that is,

$$\mathrm{E}\{\mathcal{N}_t\} = 0, \mathrm{E}\{\mathcal{M}_t\} = 0$$
$$\mathrm{E}\{\mathcal{N}_t \mathcal{N}_t^\top\} = I_{m \times m}, \mathrm{E}\{\mathcal{M}_t \mathcal{M}_t^\top\} = I_{k \times k} \tag{2}$$

Notice that in our case of the HIV/AIDS Infection Process we have

$$n = m = 3, p = k = 2 \tag{3}$$

Remark 2. The system (1) is written in the, so-called, **engineering format**. The rigorous mathematical description (which we are working with) is given in Appendix. It is represented by the system of stochastic differential equations of the Itô type.

2.2. The Class of Uncertainties

Here we suppose that the uncertain vector functions $f(x,t)$ and $g(x,t)$ belong to the class of the *quasi-Lipschitz* functions, *that is*,

$$f \in \mathcal{C}(A, c_0, c_1) := \left\{ f : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n \,\middle|\, \|f(x,t) - Ax\|^2 \leq c_0 + c_1 \|x\|^2 \quad \forall x \in \mathbb{R}^n, t \in \mathbb{R}_+ \right\}$$
$$g \in \mathcal{C}(C, c_2, c_3) := \left\{ g : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^p \,\middle|\, \|g(x,t) - Cx\|^2 \leq c_2 + c_3 \|x\|^2 \quad \forall x \in \mathbb{R}^n, t \in \mathbb{R}_+ \right\} \tag{4}$$

The physical sense of the "linear represents" A and C of the classes $\mathcal{C}(A, c_0, c_1)$ and $\mathcal{C}(C, c_2, c_3)$ is discussed in [30] and [31]. The parameters of these are supposed to be known. If a considered dynamics is bounded (exactly this case we have in our problem) then $c_1 = c_3 = 0$.

To clarify the parameters of the Quasi-Lipschitz mapping class:

-*The parameters A, c_0 and c_1 (the same as for C, c_2 and c_3) have the*

following interpretation:

-A characterizes the gradient of a linear mapping Ax and may be considered as an approximation of $\dfrac{\partial}{\partial x} f(x)$ within the region with large enough $\|x\|$,

-c_0 is the upper estimate of $\|f(0)\|^2$, or in other words, the upper estimate of possible velocity of the plant at the origin,

-c_1 characterizes the growth rates of $f(x)$ as $\|x\| \to \infty$ which is not faster than linear function of $\|x\|$.

The following properties will be required hereafter:

The pair of matrices (C, A) is observable.

The stochastic system (1) is quadratically stable, that is,

$$\limsup_{t \to \infty} E\left\{\|x_t\|^2\right\} \le X^+ < \infty \tag{5}$$

2.3. The Structure of the Filter

The state vector x_t will be estimated by the Luenberger-type filter

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{x}_t = A\hat{x}_t \mathrm{d}t + L\left(\mathrm{d}y_t - C\hat{x}_t \mathrm{d}t\right), \hat{x}_0 = \tilde{x}_0 \tag{6}$$

where $\hat{x}_t$ is referred to as the "*state estimate*" of x_t and $L \in \mathbb{R}^{n \times p}$ is called the *observer-gain matrix to be designed*. Here the initial condition of the dynamic filter $\tilde{x}_0$ (6) may differ from the initial conditions $\bar{x}_0$ of the system (1).

Now our problem can be formulated in the following manner: find the observer-gain matrix L which provides the closeness of $\hat{x}_t$ to x_t in some probabilistic sense. Below we will give the formal formulation.

2.4. The Best Selection of the Observer-Gain Matrix

The next definition will be in use hereafter.

Definition 1. The ellipsoid

$$\mathcal{E}(0, P) := \left\{x \in \mathbb{R}^n : x^\top P x \le 1\right\}$$

(*with the center point in* 0 *and the ellipsoidal matrix* $0 < P = P^\top \in \mathbb{R}^{n \times n}$ (*positive definite*)) *is said to be* **attractive** *for the stochastic trajectories* $\{x_t\}$, *defined on filtered probability space* $\left(\Omega, \mathcal{F}, \{\mathcal{H}_t\}_{t \ge 0}, \mathcal{P}\right)$ (*where σ-algebra* $\mathcal{H}_0$ *contains all the P-null sets from* $\mathcal{F}$):

1) **In mean-square sense**, *if*

$$\limsup_{t \to \infty} E\left\{x_t^\top P x_t\right\} \le 1$$

2) **With probability one** (*or almost sure*), *if for any time-subsequences* $\{t_k\}$ $(k = 0, 1, \cdots)$ *almost all random sequences* $\{x_{t_k}\}$ *leave this ellipsoid only a finite number of times, that is,*

$$\sum_{k=0}^{\infty} \chi\left(\omega : x_{t_k}^\top P x_{t_k} > 1\right) \overset{a.s.}{<} \infty$$

where $\chi(\mathcal{A})$ *is the characteristic function of the event* $\mathcal{A}$, *namely,*

$$\chi(A) = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{if if not} \end{cases}$$

Let us introduce the "measure of closeness" $J_t(L)$ as the filtering-error $e_t := x_t - \hat{x}_t$ weighted covariation matrix, namely,

$$J_t(L) := E\left\{ e_t^\mathsf{T} P_{filt} e_t \right\} \tag{7}$$

where the weighting matrix should be done as much as possible satisfying

$$0 < P_{filt} = P_{filt}^\mathsf{T} \tag{8}$$

(the biggest eigenvalues of P_{filt} correspond the smallest attractive ellipsoid semi-axis). So, we intend to obtain the gain-matrix L which asymptotically guarantee the boundedness (in mean-square sense) of all trajectories in an attractive region containing the origin, and, moreover, which minimize the "size" (in this case, the trace of the inverse ellipsoidal matrix) of the attractive ellipsoid, containing this bounded region.

Below we formulate the 1-st main result.

Theorem 1 (on the mean-square attractive ellipsoid). *If the assumptions H1-H2 are fulfilled and additionally, the following matrix inequality holds for some L, $P = P^\mathsf{T} > 0$ and some positive scalars α, ε*

$$W(P,L \mid \alpha, \varepsilon) := \begin{bmatrix} P\left(A - LC + \dfrac{\alpha}{2}I\right) + \left(A - LC + \dfrac{\alpha}{2}I\right)^\mathsf{T} P & PL & -P \\ (PL)^\mathsf{T} & -\varepsilon I_{p \times p} & 0 \\ -P & 0 & -\varepsilon I_{n \times n} \end{bmatrix} \le 0 \tag{9}$$

then the ellipsoid

$$\mathcal{E}(0, P_{filt}) := \left\{ e \in \mathbb{R}^n \mid e^\mathsf{T} P_{filt} e \le 1 \right\} \tag{10}$$

with the center in the origin $e = 0$ with

$$\begin{aligned} P_{filt} &= \frac{\alpha}{\beta(P,L,\varepsilon)} P \\ \beta(P,L,\varepsilon) &= \beta_0(P,L) + \gamma(\varepsilon) \\ \beta_0(P,L) &:= \operatorname{tr}\left\{ \|\eta\|^2 L^\mathsf{T} PL + \|\sigma\|^2 P \right\} \\ \gamma(\varepsilon) &:= \varepsilon\left[(c_0 + c_2) + (c_1 + c_3) X^+ \right] \end{aligned} \tag{11}$$

is **attractive in mean-square sense**, that is,

$$\left[\sqrt{J_t(L)} - 1 \right]_+^2 \xrightarrow[t \to \infty]{} 0 \tag{12}$$

where

$$[z]_+ := \begin{cases} z & \text{if } z \ge 0 \\ 0 & \text{if } z < 0 \end{cases} \tag{13}$$

The proofs of this and the next theorem are in Appendix.

The 2-nd main result is as follows:

Theorem 2 (on the best observer-gain matrix) *The best observer-gain ma*

matrix L^* is given by

$$L^* = \left(X^* \right)^{-1} Y^* \tag{14}$$

where X^* and Y^* are the solution of the following constraint optimization problem

$$\left. \begin{array}{l} S\left(X, H \mid \alpha, \varepsilon \right) := -\ln \left(\operatorname{tr} \left\{ \alpha^{-1} \left(\|\eta\|^2 H + \|\sigma\|^2 X \right) \right\} + \gamma(\varepsilon) \right) \to \max_{\alpha>0, \varepsilon>0, X>0, Y, H} \\[2mm] \begin{bmatrix} X\left(A + \dfrac{\alpha}{2}I \right) + \left(A^\top + \dfrac{\alpha}{2}I \right)X - YC - C^\top Y^\top & Y & -X \\ Y^\top & -\varepsilon I & 0 \\ -X & 0 & -\varepsilon I \end{bmatrix} \le 0 \\[2mm] \begin{bmatrix} H & Y^\top \\ Y & X \end{bmatrix} \ge 0 \end{array} \right\} \tag{15}$$

2.5. On the Numerical Solution of the Constraint Optimization Problem

The solution of the suboptimal constraint optimization problem (15) can be obtained by the *SEDUMI* and *YALMIP* toolboxes of MATLAB which effectively use the *Interior Point Method* (*see, for example, the details in* [31]). These toolboxes are used at each step of the following iterative procedure:

1) First, we fix the scalar parameters α, ε and solve the problem (15) varying only the matrix variables X, H and Y using the toolboxes mentioned above;

2) Second, for the found matrix variables X, H and Y we solve the problem (15) with respect to the scalar parameters α and ε;

3) Then the process iterated up to the iteration which has no solution (the toolboxes mentioned above provide this information);

4) Returning one-step back we declare this iterative approximation as a solution

$$\alpha^*, \varepsilon^*, X^*, H^* \text{ and } Y^*$$

of the considered optimization problem (15).

For the numerical implementation, then we consider the approximation of the Gaussian noise signal W_t by the noisy signal in Simulink is given in terms of a Gaussian noise signal generator W_t, using the following approximation:

$$\frac{\Delta W_t}{\Delta t} \approx \frac{W_t - W_{t-h}}{h}, 0 < h \ll 1$$

3. Filtering of HIV/AIDS Infection Model

Mathematical Model of HIV/AIDS

The Mathematical models of HIV dynamics (human immunodeficiency virus (HIV) causes the acquired immune deficiency syndrome, knows as AIDS) were derived several years ago. In this study the third-order model of HIV dynamics is considered. It captures the time rate of healthy and infected white blood cells

(T-cells) and the number of HIV viruses.

There are more complex models of HIV dynamics that can be found in the literature (see, for example, and. The methodology presented in these papers can be applied with minor modifications to the other models of HIV dynamics. However, as indicated by an anonymous reviewer of the manuscript, it should be emphasized that the presented analysis is limited in the sense that it "cannot take into the account patient factors as physiological/genetic level, physicochemical factors at cell-protein-viral interactions level, and viral factors that relate to the various HIV strains and clades."

Consider now the simplified nonlinear HIV-dynamics model (see [32] [33] [34])

$$
\begin{aligned}
\dot{x}_{1,t} &= \left[-d_0 x_{1,t} - \beta x_{1,t} x_{3,t} + s \right] dt + \sigma_1 \mathcal{N}_{1,t} \\
\dot{x}_{2,t} &= \left[-\mu_2 x_{2,t} + \beta x_{1,t} x_{3,t} \right] dt + \sigma_2 \mathcal{N}_{2,t} \\
\dot{x}_{3,t} &= \left[k x_{2,t} - \mu_1 x_{3,t} \right] dt + \sigma_3 \mathcal{N}_{3,t} \\
y_t &= C x_t, x_t = \left(x_{1,t}, x_{2,t}, x_{3,t} \right)^\top \\
C &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
\tag{16}
$$

The constant parameters in (16) are as follows:

- $s = 10$ mm^3 per day is the constant source of healthy T-cells (thymus);
- $d_0 = 0.02$ per day represents the death rate of healthy T-cells;
- $\mu_1 = 0.24$ per day represents the death rate of viruses;
- $\mu_2 = 2.4$ per day represents the death rate of infected T-cells;
- $b = 2.4 \times 10^{-5}$ per ($mm^3 \cdot$day) is the infectivity rate of free viruses;
- $k = 100$ per cell is the rate of virons (free virus particles) produced per infected T-cell.

Here C corresponds to the real situation when only the *first* and the *third* states, corrupted by the random noises, are available in time. The random variables $\mathcal{N}_{i,t} (i = 1, 2, 3)$ are the standard white noises. According to the description (18) we have

$$
f(x) = \begin{pmatrix} -d_0 x_1(t) - \beta x_1(t) x_3(t) + s \\ -\mu_2 x_2(t) + \beta x_1(t) x_3(t) \\ k x_2(t) - \mu_1 x_3(t) \end{pmatrix}, \sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}
$$

so that $n = 3, p = 2, k = m = 3$. This implies

$$
\begin{bmatrix}
dy_t = C dx_t = C f(x_t) dt + C \sigma dW_t \\
g(x) = C f(x), \eta = C \sigma \\
d\tilde{B}_t = dW_t
\end{bmatrix}
$$

1) In (18) select $A = A_1$ as

$$
A_1 = \begin{bmatrix}
-0.02 & 0 & -2.4 \times 10^{-5} \\
2.4 \times 10^{-5} & -2.4 & 2.4 \times 10^{-5} \\
0 & 100 & -2.4
\end{bmatrix}
$$

and

$$c_0 = c_2 = 0, c_1 = c_3 = 1$$

$$\sigma_i = 1/\sqrt{3}, \eta_i = 1/\sqrt{2}$$

Here $\mathcal{N}_{i,t}$ is a standard white noise, modeled in Simulink by a pseudo-random generator such that $\sup_t |\mathcal{N}_{i,t}| = 1$. The optimization procedure described above gives

$$\alpha^* = 0.00326, \beta^* = 0.567, \varepsilon^* = 1.51$$

$$H = \begin{bmatrix} 1.3524 & 0.0000 \\ 0.0000 & 2.0680 \end{bmatrix}$$

The matrix parameters P and L, obtained by the application of the suggested approach and realizing the robust output linear controller, are as follows:

$$P^* = \begin{bmatrix} 0.0033 & 0.0000 & 0.0000 \\ 0.0000 & 0.0409 & -0.0030 \\ 0.0000 & -0.0030 & 0.0035 \end{bmatrix}, L^* = \begin{bmatrix} 8.2089 & 0.0000 \\ 0.0000 & 3.0783 \\ 0.0000 & 17.5199 \end{bmatrix} \tag{17}$$

The **Figures 2-4** show the state trajectories x_1, x_2, x_3 and their estimates $\hat{x}_1, \hat{x}_2$ and $\hat{x}_3$.

The corresponding zoom-images are given in **Figure 5** and **Figure 6**. It can be seen from these figures that the filtering process has a *good performance* maintaining the filter estimates close to their real states.

Figure 7 and **Figure 8** show the entrance of the corresponding state estimation errors to the attractive ellipsoid $\mathcal{E}\left(0; \dfrac{\alpha^*}{\beta^*} P^*\right)$ in different plains:

$(e_1; e_2)$ and $(e_1; e_3)$.

We also have

$$\Theta_1 := \operatorname{tr}\{P_{filt}\} = \operatorname{tr}\left\{\dfrac{\alpha^*}{\beta^*} P^*\right\} = 0.0477$$
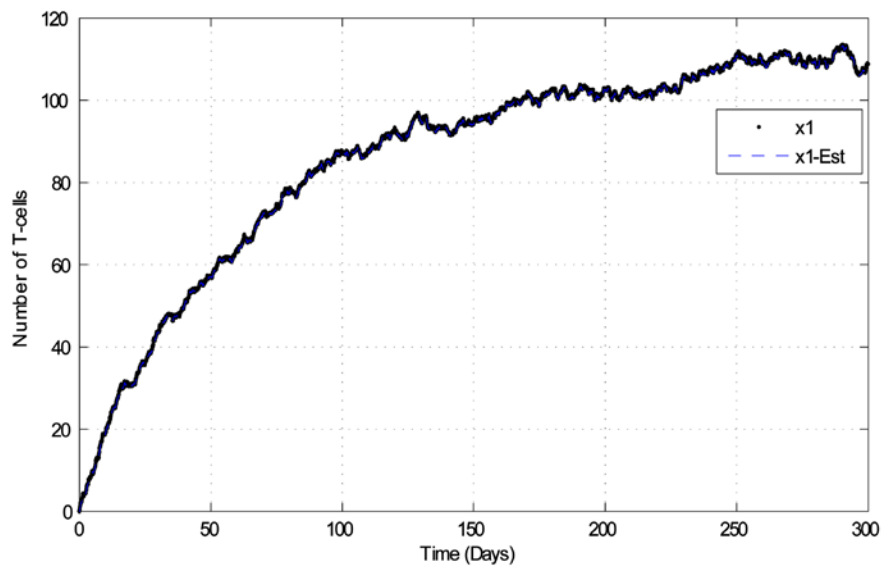

Figure 2. The state x_1 and its estimate $\hat{x}_1$.
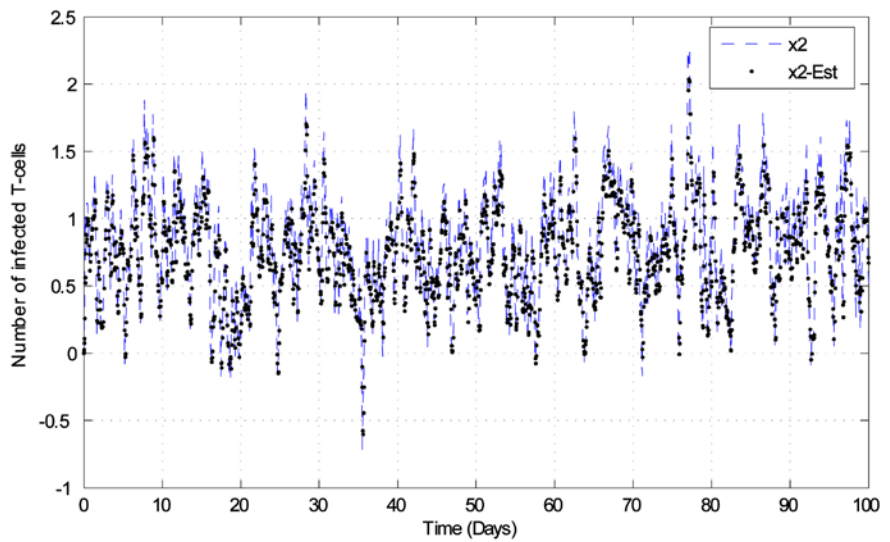
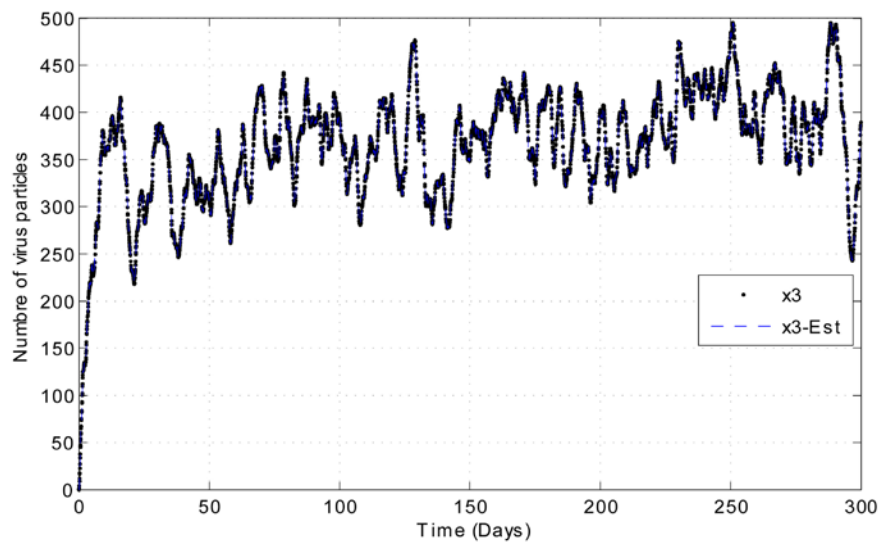**Figure 3.** The state x_2 and its estimate $\hat{x}_2$.

Figure 4. The state x_3 and its estimate $\hat{x}_3$.

2) To demonstrate that the suggested approach is sufficiently robust with respect to selection of matrix A, let us repeat the numerical example with another filter corresponding the following linear represents with $A = A_2$:

$$A_2 = \begin{bmatrix} -0.0122 & 0.0000 & -1.53 \times 10^{-5} \\ -1.53 \times 10^{-5} & -1.53 & -1.53 \times 10^{-5} \\ 0.0000 & 89.5 & -1.53 \end{bmatrix}$$

The results are as follows:

$$\alpha^* = 0.00235, \ \beta^* = 0.698, \ \varepsilon^* = 1.42$$

$$H = \begin{bmatrix} 1.3223 & 0.0000 \\ 0.0000 & 1.9044 \end{bmatrix}$$
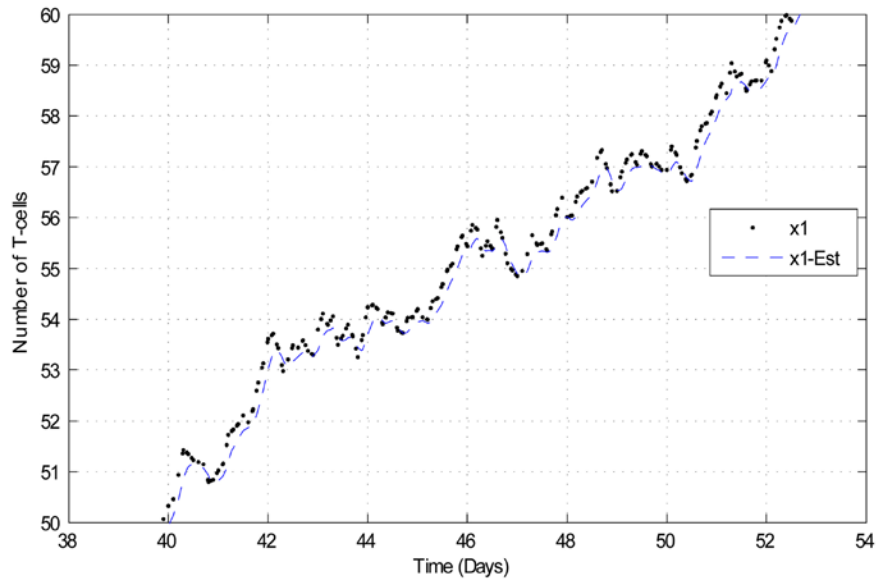
and

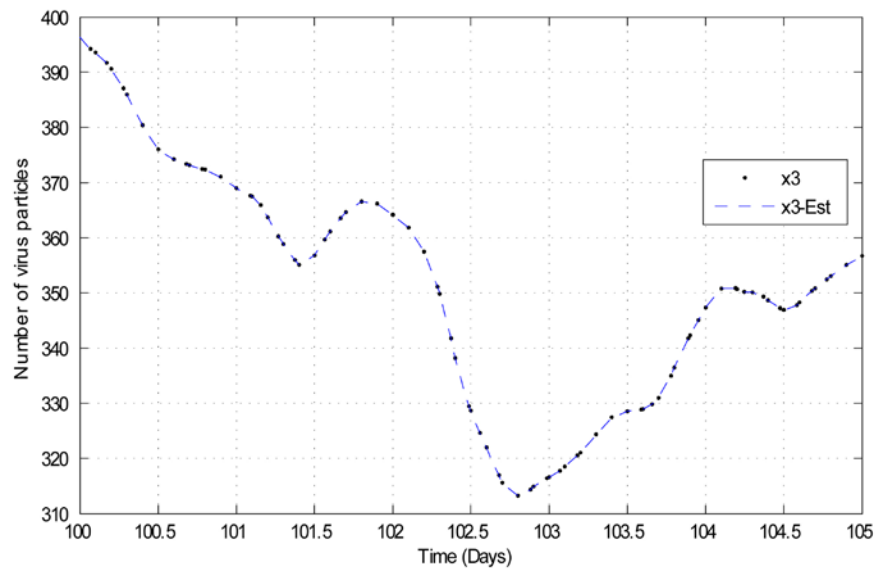**Figure 5.** Zoom-image of x_1 and its estimate $\hat{x}_1$.

Figure 6. Zoom-image of x_3 and its estimate $\hat{x}_3$.

$$\Theta_2 := \mathrm{tr}\left\{ P_{filt} \right\} = \mathrm{tr}\left\{ \frac{\alpha^*}{\beta^*} P^* \right\} = 0.0486$$

One can see that

$$\Theta_1 \simeq \Theta_2$$

4. Conclusions

In this paper, the proposed Attractive Ellipsoid Method can be successfully applied to the filtering process of nonlinear uncertain stochastic models given in Itô form, where the Luenberger-like filter, whose gain matrix should be designed, is suggested to be applied to the estimation process and the Itô calculus
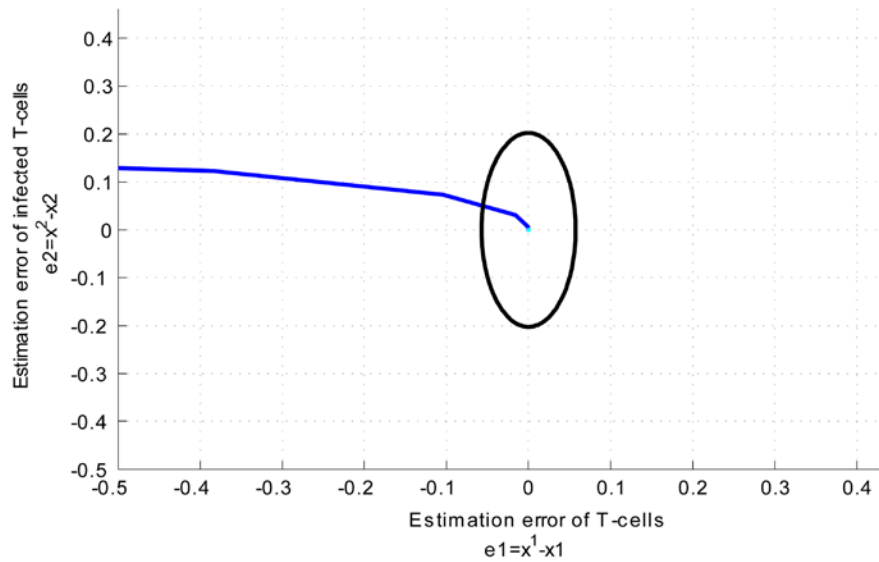
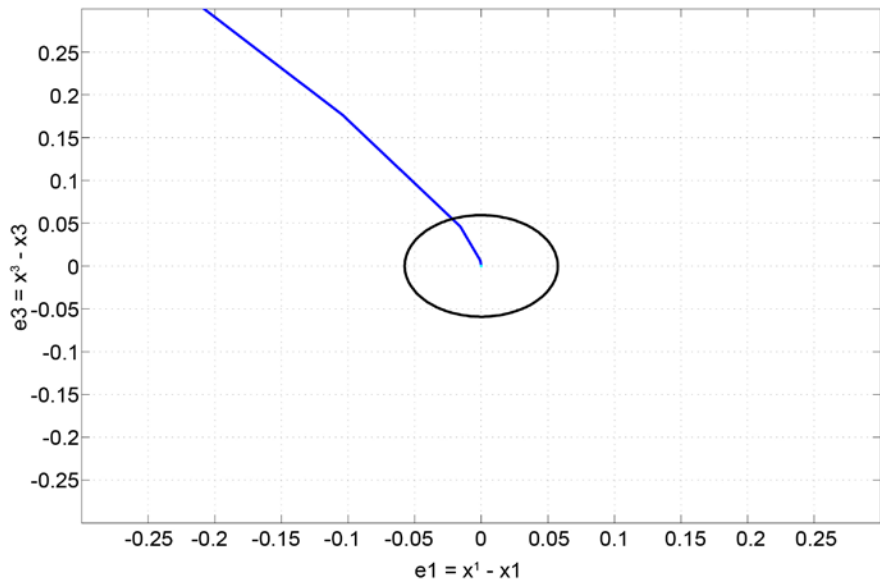**Figure 7.** The attractive ellipsoid in the plain $(e_1; e_2)$.

Figure 8. The attractive ellipsoid in the plain $(e_1; e_3)$.

should be used to derive the corresponding attractive ellipsoids where almost all trajectories of the state estimation errors converge.

To minimize the size of this ellipsoid the standard technique, under the LMI constraints, may be applied, also the suggested method is respectively robust with respect to the selection of the linear represents participating in the filter structure. Finally, the well-working of the suggested method is illustrated by the application to the filtering of the HIV/AIDS infection model.

References

[1]   Chow, P., Khasminskii, R. and Liptser, R. (1977) Tracking of Signal and Its Derivatives in Gaussian White Noise. *Stochastic Processes and Their Applications*, **69**,

259-273.

[2] Clark, J.M. (1978) The Design of Robust Approximations to the Stochastic Differential Equations of Nonlinear Filtering. In: Skwirzynski, J.K., Ed., *Communication Systems and Random Process Theory*, Sijthoff & Noordhoff, Alphen aan den Rijn, No. 25, 721-734.

[3] Kolmogorov, A. (1939) Sur l'interpolation et extrapolation des suites stationnaires. *Comptes Rendus de l'Académie des Sciences*, **208**, 2043.

[4] Krein, M. (1945) On a Generalization of Some Investigations of g. szego, w. m. smirnov, and a. n. kolmogorov. *Doklady Akademii Nauk SSSR*, **46**, 91-94.

[5] Wiener, N. (1949) Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications. MIT Press, Cambridge.

[6] Kalman, R. (1960) A New Approach to Linear Filtering and Prediction Problem. *ASME Transactions, Part D, Journal of Basic Engineering*, **82**, 35-45. https://doi.org/10.1115/1.3662552

[7] Kalman, R. and Bucy, R. (1961) New Results in Linear Filtering and Prediction Theory. *ASME Transactions, Part D, Journal of Basics Engineering*, **83**, 95-103. https://doi.org/10.1115/1.3658902

[8] Stratonovich, R. (1966) A New Representation for Stochastic Integrals and Equations. *SIAM Journal on Control and Optimization*, **4**, 362-371. https://doi.org/10.1137/0304028

[9] Jun, J.H., Zidong, W., Huijun, G. and Stergioulas, L.K. (2012) Extended Kalman Filtering with Stochastic Nonlinearities and Multiple Missing Measurements. *Automatica*, **48**, 2007-2015.

[10] Lorentzen, K., Looper, M. and Blake, J. (2001) Relativistic Electron Microbursts during the Gem Storms. *Geophysical Research Letters*, **28**, 2573-2576. https://doi.org/10.1029/2001GL012926

[11] Liang, B., Alpak, F., Sepehrnoori, K. and Delshad, M. (2007) A Singular Evolutive Interpolated Kalman Filter for Rapid Uncertainty Quantification. *Society of Petroleum Engineers, Source: SPE Reservoir Simulation Symposium*, Houston, 26-28 February 2007, Document ID SPE-106170-MS. https://doi.org/10.2118/106170-MS

[12] Kushner, H. (1967) Approximations of Nonlinear Filters. *IEEE Transactions on Automatic Control*, **12**, 546-556. https://doi.org/10.1109/TAC.1967.1098671

[13] Kushner, H. (1964) On the Differential Equations Satisfied by Conditional Densities of Markov Processes with Applications. *SIAM Journal on Control*, **2**, 106-119.

[14] Shiryaev, A. (1967) Some New Results in the Theory of Controlled Random Processes. In: Prague, A., Ed., *Transactions of the 4th Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*, Prague, 31-203. (In Russian)

[15] Kailath, T. (1968) An Innovations Approach to Least-Squares Estimation. I: Linear Filtering in Additive White Noise. *IEEE Transactions on Automatic Control*, **13**, 646-655. https://doi.org/10.1109/TAC.1968.1099025

[16] Frost, P. and Kailath, T. (1971) An Innovations Approach to Least Squares Estimation. *IEEE Transactions on Automatic Control*, **16**, 217-226. https://doi.org/10.1109/TAC.1971.1099704

[17] Fujisaki, M., Kallianpur, G. and Kunita, B.H. (1972) Stochastic Differential Equations for the Non Linear Filtering Problem. *Osaka Journal of Mathematics*, **9**, 19-40.

[18] Duncan, T. (1970) On the Absolute Continuity of Measures. *The Annals of Mathematical Statistics*, **41**, 30-38. https://doi.org/10.1214/aoms/1177697185

[19] Duncan, T. (1970) Likelihood Functions for Stochastic Signals in White Noise. *Information and Control*, **16**, 303-310.

[20] Mortensen, R.E. (1966) Stochastic Optimal Control with Noisy Observations. *International Journal of Control*, **1**, 455-464.
https://doi.org/10.1080/00207176608921439

[21] Zakai, M. (1969) On the Optimal Filtering of Diffusion Processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **11**, 230-243.
https://doi.org/10.1007/BF00536382

[22] Pardoux, E. (1979) Stochastic Partial Diffusion Equations and Filtering of Diffusion Processes. *Stochastics*, **3**, 127-167.

[23] Krylov, N.V. and Rozovskii, B.L. (1977) The Cauchy Problem for Linear Stochastic Partial Differential Equations. *Izv. Akad. Nauk SSSR Ser. Mat.*, **41**, 1329-1347.

[24] Davis, M. (1977) Linear Estimation and Stochastic Control. Chapman and Hall, New York.

[25] Liptser, R. and Shiryaev, A. (2001) Statistics of Random Processes: I, II General Theory, Volumes 5 and 6 of Stochastic Modelling and Applied Probablility. 2nd Edition, Springer, New York.

[26] Gershon, E., Limebee, D., Shaked, U. and Yaesh, I. (2001) Robust H-Inf Filtering of Stationary Continuous-Time Linear Systems with Stochastic Uncertainties. *IEEE Transactions on Automatic Control*, **46**, 1788-1793.
https://doi.org/10.1109/9.964692

[27] Ugrinovskii, V. and Petersen, I. (2002) Robust Filtering of Stochastic Uncertain Systems on an Infinite Time Horizon. *International Journal of Control*, **75**, 614-626.
https://doi.org/10.1080/00207170210134219

[28] Zhang, W., Chen, B.S. and Tseng, C.S. (2005) Robust H-Inf Filtering for Nonlinear Stochastic Systems. *IEEE Transactions on Signal Processing*, **53**, 589-597.
https://doi.org/10.1109/TSP.2004.840724

[29] Lozada-Castillo, N., Alazki, H. and Poznyak, A. (2013) Robust Control Design through the Attractive Ellipsoid Technique for a Class of Linear Stochastic Models with Multiplicative and Additive Noises. *IMA Journal of Mathematical Control and Information*, **30**, 1-19. https://doi.org/10.1093/imamci/dns008

[30] Alazki, H. and Poznyak, A.S. (2013) A Class of Robust Bounded Controllers Tracking a Nonlinear Discrete-Time Stochastic System: Attractive Ellipsoid Technique Application. *Journal of the Franklin Institute*, **350**, 1008-1029.

[31] Poznyak, A., Polyakov, A. and Azhmyakov, V. (2014) Attractive Ellipsoids in Robust Control. Birkhauser-Springer.

[32] Perelson, A. and Nelson, P. (1996) Mathematical Analysis of HIV-I Dynamics in Vivo. *SIAM Review*, **41**, 3-44. https://doi.org/10.1137/S0036144598335107

[33] Ko, J., Kim, W. and Chung, C. (2006) Optimized Structural Interpretation for HIV Therapy and Its Performance Analysis on Controllability. *Transactions on Biomedical Engineering*, **53**, 380-386. https://doi.org/10.1109/TBME.2005.869651

[34] Kremling, A. and Saez-Rodriguez, J. (2007) Systems Biology—An Engineering Perspective. *Journal of Biotechnology*, **129**, 329-351.

5. Appendix

5.1. The Mathematical Description of the Considered System: The Itô-Type Model

Consider a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{H}_t\}_{t \geq 0}, \mathcal{P}\right)$ where σ-algebra $\mathcal{H}_0$ contains all the P-null sets from $\mathcal{F}$, the filtration $\{\mathcal{H}_t\}_{t \geq 0}$ is right continuous, that is, $\mathcal{H}_{t+} := \bigcap_{s > t} \mathcal{H}_s = \mathcal{H}_t$. Let $\left\{(W_t, \mathcal{F}_t)\right\}_{t \geq 0}$, $\left\{(\tilde{B}_t, \tilde{\mathcal{F}}_t)\right\}_{t \geq 0}$ be two m and k dimensional independent standard Brownian motions, $\mathcal{F}_t, \tilde{\mathcal{F}}_t \subset \mathcal{F}$ and $m, k \in \mathbb{N}$. That is, $\mathcal{F}_t$ is independent of $\tilde{\mathcal{F}}_t$ for all $t \geq 0$. $\mathcal{H}_t \subset \mathcal{F}$ stands for the smallest σ-algebra containing σ-algebras $\mathcal{F}_t$ and $\tilde{\mathcal{F}}_t$.

We are interested in the following nonlinear stochastic differential equations

$$\left. \begin{aligned} \mathrm{d}x_t &= f\left(x_t\right)\mathrm{d}t + \sigma \mathrm{d}W_t \\ \mathrm{d}y_t &= g\left(x_t\right)\mathrm{d}t + \eta \mathrm{d}\tilde{B}_t \\ x_t &\in \mathbb{R}^n, y_t \in \mathbb{R}^p \end{aligned} \right\} \tag{18}$$

Here x_t and y_t are the state vector and the measured output of the model (18) at the time $t \geq t_0 = 0$. The functions $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^p$ are nonlinear mappings and $\sigma \in \mathbb{R}^{n \times m}$ and $\eta \in \mathbb{R}^{p \times k}$. The variables $W_t, \tilde{B}_t$ are *standard m and k vector valued Brownian motion* characterizing external random perturbations to the vector-state dynamics satisfying (18); in engineering applications it is referred as a "white noise" external perturbation affecting the given model dynamics. The vector valued Brownian motions W_t and $\tilde{B}_t$ have the following properties:

$$\mathrm{E}\left\{ \int_{\tau = t_0}^{t} s\left(x_\tau\right)\mathrm{d}W_\tau \right\} = 0, \mathrm{E}\left\{ \int_{\tau = t_0}^{t} s\left(x_\tau\right)\mathrm{d}\tilde{B}_\tau \right\} = 0$$

for any integrable function $s(x)$ and

$$E\left\{ \int_{\tau = t_0}^{t} \mathrm{d}W_\tau \mathrm{d}W_\tau^\top \right\} = \left(t - t_0\right) I_{m \times m}, E\left\{ \int_{\tau = t_0}^{t} \mathrm{d}\tilde{B}_\tau \mathrm{d}\tilde{B}_\tau^\top \right\} = \left(t - t_0\right) I_{k \times k}$$

$$E\left\{ \int_{\tau = t_0}^{t} \mathrm{d}\tilde{B}_\tau \mathrm{d}W_\tau^\top \right\} = 0_{k \times m}$$

5.2. Proof of Theorem 1

First, let us represent the system (18) in the, so-called, quasi-linear format:

$$\left. \begin{aligned} \mathrm{d}x_t &= \left[A x_t + \Delta f\left(x_t\right) \right]\mathrm{d}t + \sigma \mathrm{d}W_t \\ \mathrm{d}y_t &= \left[C x_t + \Delta g\left(x_t\right) \right]\mathrm{d}t + \eta \mathrm{d}\tilde{B}_t \\ \Delta f\left(x\right) &:= f\left(x\right) - Ax, \Delta g\left(x\right) := g\left(x\right) - Cx \end{aligned} \right\} \tag{19}$$

For the estimation error $e_t := x_t - \hat{x}_t$ which (in view of (6) and (19)) we have

$$\begin{aligned} \mathrm{d}e_t &= r_t \mathrm{d}t + \mathrm{d}\mathring{W}_t \\ r_t &:= \left(A - LC\right)e_t + \left(L\Delta g\left(x_t\right) - \Delta f\left(x_t\right)\right) \\ \mathrm{d}\mathring{W}_t &:= \left(L\eta \mathrm{d}\tilde{B}_t + \sigma \mathrm{d}W_t\right) = \begin{bmatrix} L\eta & \sigma \end{bmatrix}\begin{bmatrix} \mathrm{d}\tilde{B}_t \\ \mathrm{d}W_t \end{bmatrix} \end{aligned} \tag{20}$$

For the storage function

$$V(e) = e^\mathsf{T} P e, P = P^\mathsf{T} > 0$$

using the Itô formula (see, for example, [?]) we obtain

$$dV(e_t) = 2e_t^\mathsf{T} P de_t + \mathrm{E}\left\{\mathrm{tr}\left\{\mathring{W}_t \mathring{W}_t^\mathsf{T} P\right\}\right\} dt$$

$$= 2e^\mathsf{T} P\left[r_t dt + d\mathring{W}_t\right] + \mathrm{E}\left\{\mathrm{tr}\left\{\begin{bmatrix}L\eta & \sigma\end{bmatrix}\begin{bmatrix}d\tilde{B}_t d\tilde{B}_t^\mathsf{T} & d\tilde{B}_t dW_t^\mathsf{T} \\ dW_t d\tilde{B}_t^\mathsf{T} & dW_t dW_t^\mathsf{T}\end{bmatrix}\begin{bmatrix}L\eta & \sigma\end{bmatrix}^\mathsf{T} P\right\}\right\} dt$$

In the integral form this relation can be expressed as

$$V(e_{t+\Delta t}) - V(e_t)$$

$$= 2\int_{\tau=t}^{t+\Delta t} e_\tau^\mathsf{T} P\left[r_\tau d\tau + d\mathring{W}_\tau\right] + \int_{\tau=t}^{t+\Delta t} \mathrm{E}\left\{\mathrm{tr}\left\{\begin{bmatrix}L\eta & \sigma\end{bmatrix}\begin{bmatrix}d\tilde{B}_\tau d\tilde{B}_\tau^\mathsf{T} & d\tilde{B}_\tau dW_\tau^\mathsf{T} \\ dW_\tau d\tilde{B}_\tau^\mathsf{T} & dW_\tau dW_\tau^\mathsf{T}\end{bmatrix}\begin{bmatrix}L\eta & \sigma\end{bmatrix}^\mathsf{T} P\right\}\right\} d\tau$$

$$= 2\int_{\tau=t}^{t+\Delta t} e_\tau^\mathsf{T} P\left[r_\tau dt + d\mathring{W}_\tau\right] + \mathrm{tr}\left\{\begin{bmatrix}L\eta & \sigma\end{bmatrix}\begin{bmatrix}I_{k\times k} & 0_{k\times m} \\ 0_{m\times k} & I_{m\times m}\end{bmatrix}\begin{bmatrix}L\eta & \sigma\end{bmatrix}^\mathsf{T} P\right\}\Delta t$$

$$= 2\int_{\tau=t}^{t+\Delta t} e_\tau^\mathsf{T} P\left[r_\tau dt + d\mathring{W}_\tau\right] + \mathrm{tr}\left\{\begin{bmatrix}L\eta & \sigma\end{bmatrix}\begin{bmatrix}L\eta & \sigma\end{bmatrix}^\mathsf{T} P\right\}\Delta t$$

$$= 2\int_{\tau=t}^{t+\Delta t} e_\tau^\mathsf{T} P\left[r_\tau dt + d\mathring{W}_\tau\right] + \mathrm{tr}\left\{\left[L\eta\eta^\mathsf{T} L^\mathsf{T} + \sigma\sigma^\mathsf{T}\right] P\right\}\Delta t$$

Using the property of the Itô integral

$$\mathrm{E}\left\{\int_{\tau=t_0}^{t} f(x_\tau) d\mathring{W}_\tau\right\} = 0$$

applying the operator $\mathrm{E}\{\cdot\}$ of the mathematical expectation to both side of the previous identity, dividing by Δt and taking $t \to \infty$, for

$$V_t := \mathrm{E}\left\{V(e_t)\right\}$$

we get

$$\dot{V}_t = \mathrm{E}\left\{2e_t^\mathsf{T} P\left[(A - LC)e_t + \begin{bmatrix}L \vdots -I\end{bmatrix}\begin{pmatrix}\Delta g(x_t) \\ \Delta f(x_t)\end{pmatrix}\right]\right\}$$

$$\quad + \mathrm{tr}\left\{\left[L\eta\eta^\mathsf{T} L^\mathsf{T} + \sigma\sigma^\mathsf{T}\right] P\right\}$$

$$= \mathrm{E}\left\{e_t^\mathsf{T}\left[P(A - LC) + (A - LC)^\mathsf{T} P\right]e_t + 2e_t^\mathsf{T} P\begin{bmatrix}L \vdots -I\end{bmatrix}\begin{pmatrix}\Delta g(x_t) \\ \Delta f(x_t)\end{pmatrix}\right\}$$

$$\quad + \mathrm{tr}\left\{\left[L\eta\eta^\mathsf{T} L^\mathsf{T} + \sigma\sigma^\mathsf{T}\right] P\right\}$$

Since

$$\mathrm{tr}\left\{\left[L\eta\eta^\mathsf{T} L^\mathsf{T} + \sigma\sigma^\mathsf{T}\right] P\right\} = \mathrm{tr}\left\{P^{1/2}\left[L\eta\eta^\mathsf{T} L^\mathsf{T} + \sigma\sigma^\mathsf{T}\right] P^{1/2}\right\}$$

$$= \mathrm{tr}\left\{P^{1/2} L\eta\eta^\mathsf{T} L^\mathsf{T} P^{1/2} + P^{1/2}\sigma\sigma^\mathsf{T} P^{1/2}\right\}$$

$$\leq \beta_0(P, L)$$

we obtain

$$\dot{V}_t \leq \mathrm{E}\left\{z_t^\mathsf{T} W_0 z_t\right\} + \beta_0$$

where

$$z_t = \begin{pmatrix} e_t^\top & \Delta g\left(x_t\right)^\top & \Delta f\left(x_t\right)^\top \end{pmatrix}^\top$$

and

$$W_0\left(P,L\right) := \begin{bmatrix} P\left(A-LC\right)+\left(A-LC\right)^\top P & PL & -P \\ \left(PL\right)^\top & 0 & 0 \\ -P & 0 & 0 \end{bmatrix}$$

$$= W\left(P,L\mid\alpha,\varepsilon\right) - \alpha e_t^\top P e_t + \varepsilon\left[\left(c_0+c_2\right)+\left(c_1+c_3\right)X^+\right]$$

Then, finally we get

$$\dot{V}_t \le E\left\{z_t^\top W z_t\right\} - \alpha V_t + \beta\left(P,L,\varepsilon\right) \tag{21}$$

and if $W \le 0$, from (21) we may conclude that

$$\dot{V}_t \le -\alpha V_t + \beta\left(P,L,\varepsilon\right)$$

Then from (21) for the function

$$\hat{V}_t := \left[\sqrt{V_t}-\mu\right]_+^2, \mu > 0$$

it follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{V}_t = \left[\sqrt{V_t}-\mu\right]_+ \frac{\dot{V}_t}{\sqrt{V_t}} \le \left[\sqrt{V_t}-\mu\right]_+ \frac{-\alpha V_t+\beta}{\sqrt{V_t}}$$

$$= -\alpha\left[\sqrt{V_t}-\mu\right]_+ \frac{\left(V_t-\dfrac{\beta}{\alpha}\right)}{\sqrt{V_t}}$$

Taking $\mu = \sqrt{\dfrac{\beta}{\alpha}}$ from the last differential inequality we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{V}_t \le -\alpha\left[\sqrt{V_t}-\mu\right]_+ \left(\sqrt{V_t}-\mu\right)\frac{\left(\sqrt{V_t}+\mu\right)}{\sqrt{V_t}}$$

$$= -\alpha\left[\sqrt{V_t}-\mu\right]_+^2 \frac{\left(\sqrt{V_t}+\mu\right)}{\sqrt{V_t}} = -\alpha\hat{V}_t\frac{\left(\sqrt{V_t}+\mu\right)}{V_t} \le -\alpha\hat{V}_t$$

implying

$$\int_{t=0}^{\infty}\hat{V}_t\mathrm{d}t < \infty \quad\text{and}\quad \hat{V}_t \le \hat{V}_0 e^{-\alpha t} \to 0 \quad\text{when } t\to\infty$$

In view of

$$\left[\sqrt{J_t\left(L\right)}-1\right]_+^2 = \frac{1}{\mu}\hat{V}_t \xrightarrow[t\to\infty]{} 0$$

we conclude the proof.

5.3. Proof of Theorem 2

The "best" gain matrix L of the filter is a solution of the following optimization problem:

$$\frac{\alpha}{\beta(P,L)}\operatorname{tr}\{P\} \to \max_{\alpha>0,\varepsilon>0,P>0,L} W(P,L\,|\,\alpha,\varepsilon)\leq 0$$

This problem is equivalent to the following one

$$\ln \operatorname{tr} \frac{\alpha}{\beta(P,L)}P = -\ln\left(\frac{\beta(P,L)}{\alpha}\right)+\ln\operatorname{tr}\{P\}\to \max_{\alpha>0,\varepsilon>0,P>0,L}W(P,L\,|\,\alpha,\varepsilon)\leq 0$$

Notice that

$$-\ln\left(\frac{\beta(P,L)}{\alpha}\right)=-\ln\left(\beta_0(P,L)+\frac{\varepsilon}{\alpha}\Big[(c_0+c_2)+(c_1+c_3)X^+\Big]\right)$$

$$=-\ln\left(\alpha^{-1}\operatorname{tr}\left\{\|\eta\|^2 L^\mathsf{T}PL+\|\sigma\|^2 P\right\}+\frac{\varepsilon}{\alpha}\Big[(c_0+c_2)+(c_1+c_3)X^+\Big]\right)$$

$$\geq -\ln\left(\operatorname{tr}\left\{\alpha^{-1}\left(\|\eta\|^2 H+\|\sigma\|^2 P\right)\right\}+\gamma(\varepsilon)\right)$$

$$:= S(P,H\,|\,\alpha,\varepsilon)$$

where the matrices H satisfies

$$L^\mathsf{T}PL=\left(L^\mathsf{T}P\right)P^{-1}\left(PL\right)\leq H$$

or, equivalently, by the Shour's lemma

$$\begin{bmatrix} H & L^\mathsf{T}P \\ PL & P \end{bmatrix}\geq 0$$

Finally, in new variables

$$X := P, Y := PL$$

our problem (with the supporting functional ($S(P,H\,|\,\alpha,\varepsilon)$)) has the form (15). Theorem is proven.