Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

State estimation of nonlinear systems with Markov state reset

Stefano Battilotti

***Abstract*— We present a novel observer design for a class of single-output nonlinear systems with Markov jumps. The Markov jump process interferes with a deterministic nonlinear dynamics at random times and retains its state for a certain amount of time (dwell time). The estimation process is reset at these random times, depending on the reset values of the state process, and then evolves as a deterministic estimate of the state process itself. The novelty is given by the reset mechanism adopted for the estimation process itself, depending on the reset values of the state process. We prove that, as long as the mathematical expectation of the dwell times has a positive lower bound and the transition rate of the jump process at the first exit time out of any point is bounded, the state estimation error of the switching dynamics asymptotically converges to zero with probability one. The state estimate over each dwell time is designed using the novel technique of "output immersion".**

***Index Terms*— Stochastic systems, state reset, observer design.**

I. Introduction

In this paper we investigate the observer design problem for the class of stochastic systems

$$\begin{aligned} x^+(t) &= Ax(t) + f(x(t), r(t)), \\ \mu(t) &= (Cx(t), r(t)), \zeta_k \le t < \zeta_{k+1}, k = 0, 1, \dots, \infty, \\ x(\zeta_k) &= q(\zeta_k) := g(x(\zeta_k^-), q(\zeta_k^-)), \end{aligned} \quad (1)$$

where $\zeta_0 = 0$, (C, A) is an observable pair, $x(t)$ is the *state process*, $x^+(t)$ denotes the right-hand derivative of $x(t)$ and $x(t^-)$ the left-hand limit of $x(t)$, $\mu(t)$ is the *measured output process*, $y(t) = (r(t), q(t))$ is the *interfering process* and $f(x, r)$ and $g(x, q)$ are for each fixed (r, q) real-valued locally Lipschitz continuous functions. We will assume without loss of generality that (C, A) is in the observability form. The *interfering* process $y(t)$ is a jump Markov process defined on some probability space $\{\mathscr{Y}, \mathfrak{L}, \mathbf{P}_y\}$, where

- $\mathscr{Y} = \mathscr{R} \times \mathscr{Q}$, $\mathscr{R} \subset \mathbb{R}^s$ and $\mathscr{Q} \subset \mathbb{R}^n$ are finite sets endowed with the discrete topology (i.e. induced by the metric $\rho(v, w) = 1$ if $v \ne w$ and $\rho(v, w) = 0$ else)
- $\mathfrak{L}$ is a σ-algebra which contains all the singletons in $\mathscr{Y}$, i.e. the sets consisting of a single element in $\mathscr{Y}$
- $\mathbf{P}_y$ is a conditional probability given y defined on $\mathfrak{L}$
- the components $y_i(t)$ of $y(t)$ are right-continuous trajectories with Markov times $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \cdots$, $\zeta_n \uparrow \infty$ such that $y_i(t) = y_{ik}$ for $t \in [\zeta_k, \zeta_{k+1})$ (here $\uparrow$ means "tends monotonically increasing"; for a definition of Markov times see 1.4 of [3]). The random time ζ_k represents the time at which each $y_i(t)$ changes its state y_{ik} into $y_{i,k+1}$ and $\zeta_{k+1} - \zeta_k$ is the time for which the jump process $y(t)$ retains its state (*dwell time*).

The jump process $y(t)$ interferes with the process $x(t)$ through the maps $f(x(t), r(t))$ and the state reset $x(\zeta_k) = g(x(\zeta_k^-), q(\zeta_k^-))$, $i = 1, \dots, n$. The process $(x(t), y(t))$ is described by $\mathbf{P}_{x,y}$ its conditional probability given (x, y). This probability, defined according to (25), depends on the conditional probabilities $\mathbf{P}_x^{(y)}$ given x, associated to each value y of $y(t)$, and defines on turn the transition probabilities $P_{x,y}(u, w)$ of the process $(x(t), y(t))$ from the state (x, y) into (u, w). As it will be seen in the statement of our main result, we will assume that *instantaneous transition rates* $\partial P_{x,y}(u, w)$ of the process $(x(t), y(t))$ from (x, y) into (u, w) are well defined and bounded.

Our study

- is limited to a deterministic evolution of the continuous state for each fixed value $y(t) = y \in \mathscr{Y}$ (the *frozen* system)
- it does not allow transitions triggered by conditions on the state (*guards*)
- assumes that the process $r(t)$ (together with the first component of the state process $Cx(t) = x_1(t)$) can be measured.

The first restriction is assumed only for simplicity and in the appendix it is shown how to generalize our study to the case of continuous evolution driven by stochastic differential equations. The second restriction allows to exclude Zeno phenomena and it will be relaxed in a future work. The third restriction is also assumed for simplicity and more generally we can assume to measure only $x_1(t)$. In this case an observer should be designed also for estimating $r(t)$. Our model is similar to the one considered in [4] (see also numerous references therein).

Observers for linear systems (1) with no state reset were proposed in [2] and for feedback linearizable nonlinear systems in [1]. In [4] the infinitesimal generator of the process $(x(t), y(t))$ is calculated and used for the analisys of TCP flows. In this paper, we propose an observer for (1) as a result of observers designed for each frozen system. We assume that the state process of (1) evolves in some compact set. Our contribution does not introduce any new observability notion, rather it gives a new design tool for observers with state reset. As to the observer design for each frozen system we introduce a novel framework based on the notion of output immersion of a system, which consists of immersing its output space into some larger space in such a way to increase the number of its outputs up to as many as the number of the states. This immersion has as a result the

S. Battilotti is with the Dipartimento di Informatica e Sistemistica "Antonio Ruberti", Università di Roma "La Sapienza" battilotti@dis.uniroma1.it

peculiarity of decoupling the original n-dimensional system into n one-dimensional systems, each with its own state and output and for which a one-dimensional observer can be easily designed. Stepping back in the immersion process, one obtains an observer for the system before immersion as a result of the n one-dimensional observers. This mainly determines a "chained" structure for the observer. Moreover, following [7], on account of the nonlinear nature of the dynamics we saturate the estimation process over the same compact set in which the state process is assumed to evolve. However, we establish more general conditions (see (2)-(3) and remark 2.2) under which an observer can be designed.

We define a probability space for the process $z(t) = (x(t), \xi(t), y(t))$ where $\xi(t)$ is the estimation process. In particular, one of our contributions with respect to [4] is to define a transition probability for this process, using the conditional probability $\mathbf{P}^{(y)}_{x,\xi}$ given (x, ξ) for the process $(x(t), \xi(t), y)$, and to calculate the characteristic operator of $z(t)$. This transition probability chains in a natural way the conditional probabilities $\mathbf{P}^{(y)}_{x,\xi}$ given (x, ξ) so that from the properties of the state estimation error for each process $(x(t), \xi(t), y)$ it is easy to obtain asymptotic properties of the state estimation error for the switching process $z(t)$. In particular, we prove that, as long as the mathematical expectation of the dwell times has a positive lower bound and the transition rate of the jump process at the first exit time out of any point is bounded, the state estimation error of the switching dynamics asymptotically converges to zero with probability one.

A. Notations

- $\mathbb{R}^+$ (resp. $\mathbb{R}^{\geq}$) denotes the set of real positive (resp. nonnegative) numbers. $\mathbb{R}^n$ is the vector space of (column) vectors with n elements. For any vector $x \in \mathbb{R}^n$ we denote by x_i the i-th element of x. Also for any vector function $f(\cdot)$ we will denote by $f_i(\cdot)$ its i-th component.
- $\|v\|$ denotes the euclidean norm of a vector $v \in \mathbb{R}^n$.
- For any functions $h_j : \mathscr{D} \to \mathbb{R}^{\geq}$, $j = 1, 2$, we write $h_1 \preceq h_2$ (resp. $h_1 \succeq h_2$) if there exists $a \geq 1$ such that $h_1(s) \leq a h_2(s)$ (resp. $h_1(s) \geq a h_2(s)$) for all $s \in \mathscr{D}$.
- $x^+(t)$ denotes the right-hand derivative of $x(t)$ while $x(t^-)$ the left-hand limit of $x(t)$.
- $\uparrow$ means "tends monotonically increasing", while $\downarrow$ means "tends monotonically decreasing".

II. Assumptions and main result

In this section we discuss the main assumptions and we state the main result of this paper. Our first assumption (H1) is on the state trajectory $x(t)$ we want to estimate:
(H1) *The state process $x(t)$ has a compact phase space $\mathscr{X} = \{x \in \mathbb{R}^n : |x_i| \leq c^{\delta_i}\}$, with $c > 1$ and $\delta_i \in \mathbb{R}^{\geq}, i = 1, \dots, n$.*

Remark *2.1:* Each state trajectory $x(t)$ is assumed to stay for all times with probability one in some compact set $\mathscr{X}$. For our convenience, the compact set $\mathscr{X}$ has the form above. □

Let $\delta_{n+1} \in \mathbb{R}^{\geq}$ and $\theta_{il} \in [-\infty, \infty)$, $i, l = 1, \dots, n$, be numbers such that

$$\theta_{il} < \begin{cases} 2\sum_{h=l}^{n-1}(3^{h-l}-1)(\delta_h - \delta_{h+1}) \\ +(1+3^{n-l})\delta_{n+1} \qquad \text{if } i = n \\ 2\sum_{h=l}^{n-1}(3^{h-l}-1)(\delta_h - \delta_{h+1}) \\ -2\sum_{h=i+1}^{n-1}(3^{h-i-1}-1)(\delta_h - \delta_{h+1}) \\ +(3^{n-l} - 3^{n-i-1})\delta_{n+1} \text{ if } i = 1, \dots, n-1, \end{cases} \tag{2}$$

with

$$\delta_{n+1} \geq \delta_i - \delta_{i-1}, i = 2, \dots, n. \tag{3}$$

Our assumption (H2) is the following:
(H2) *For each $i = 1, \dots, n$ there exist $a_{il} : \mathscr{X} \times \mathscr{X} \to \mathbb{R}^{\geq}$, $l = 1, \dots, n$, such that*

$$[f_i(x, r) - f_i(w, r)]^2 \leq \sum_{l=1}^{n} (x_l - w_l)^2 a_{il}(x, w), \tag{4}$$

$$a_{il}(x, w) \preceq c^{\theta_{il}}, \tag{5}$$

for all $c > 1$, $x, w \in \mathscr{X}$ and $r \in \mathscr{R}$.

Remark *2.2:* Condition (4) requires that the "incremental rate" of each function $f_i(x, r)$ with increment $x_l - w_l$, $l = 1, \dots, n$, are bounded uniformly in r by some function $a_{il}(x, w)$: this condition is always satisfied by locally Lipschitz functions $f_i(x, r)$, $r \in \mathscr{R}$ a finite set. Therefore, in (5) we are restricting the "incremental rates" of each function $f_i(x, r)$ over a given compact set. Note that if $a_{il}(x, w) \equiv 0$ for all $l > i$ and $i = 1, \dots, n$ ([7]) then (5) can be always satisfied for any given $\delta_i \in \mathbb{R}^{\geq}$, $i = 1, \dots, n$, i.e. for any given $\mathscr{X}$, by choosing any $c > 1$ and a sufficiently large δ_{n+1}.

Assumption (H2) have the following physical interpretation. As it will be shown in the proof of our main result, the numbers $\delta_h - \delta_{h+1}$, $h = 1, \dots, n-1$, and δ_{n+1} determine the gains Δ_l, $l = 1, \dots, n-1$, and, respectively, Δ_n of the observer which estimates the state trajectory $x(t)$. On the other hand, the numbers $\delta_h - \delta_{h+1}$, $h = 1, \dots, n-1$, and δ_{n+1} determine a bound for the numbers θ_{il} in (2) and, therefore, for the incremental rates $a_{nl}(x, w)$, $l = 1, \dots, n$, over $\mathscr{X} \times \mathscr{X}$ in (5). On account of (2) is possible to increase this upper bound by *increasing* either δ_{n+1} or $\delta_h - \delta_{h+1}$, $h = l, \dots, n-1$, i.e. either Δ_n or Δ_l. This is exactly the interpretation of conditions (2) for $i = n$.

We end with few comments on the conditions (5) for $i = 1, \dots, n-1$. If $l - 1 < i$ it is possible to increase the upper bound for $a_{il}(x, w)$ in (5) over $\mathscr{X} \times \mathscr{X}$ by *decreasing* δ_{n+1} or *increasing* $\delta_l - \delta_{l+1}, \dots, \delta_i - \delta_{i+1}$, i.e. increasing $\Delta_l - \Delta_{i+1}$. If $i < l - 1$ it is possible to increase this upper bound by *decreasing* either δ_n or $\delta_{i+1} - \delta_{i+2}, \dots, \delta_{l-1} - \delta_l$, i.e. increasing $\Delta_l - \Delta_{i+1}$ or, by (3), Δ_{n+1}. Therefore, a trade-off should be sought in the selection of $\Delta_1, \dots, \Delta_n$ for increasing as much as possible the upper bounds for *all* $a_{il}(x, w)$, $l = 1, \dots, n$, over $\mathscr{X} \times \mathscr{X}$. Farther is the map $f(x, r)$ from having a "triangular" structure, more difficult is the achievement of this trade-off. This is exactly the interpretation of conditions (2) for $i = 1, \dots, n-1$. □

Also, let $\theta'_{il} \in [-\infty, \infty)$, $i, l = 1, \ldots, n$, be numbers such that

$$\theta'_{il} < \begin{cases} 2\sum_{h=l}^{n-1}(3^{h-l}-1)(\delta_h - \delta_{h+1}) \\ +3^{n-l}\delta_{n+1} & \text{if } i = n \\ 2\sum_{h=l}^{n-1}(3^{h-l}-1)(\delta_h - \delta_{h+1}) \\ -4\sum_{h=i+1}^{n-1}(3^{h-i-1}-1)(\delta_h - \delta_{h+1}) \\ +(3^{n-l} - 3^{n-i-1}2)\delta_{n+1} & \text{if } i = 1, \ldots, n-1. \end{cases}$$

Conditions similar to (2) are imposed on the reset maps $g(x, q)$ in (1):

(H3) *there exist $b_{il} : \mathscr{X} \times \mathscr{X} \to \mathbb{R}^{\geq}$, $l = 1, \ldots, n$, such that for some*

$$[g_i(x, q) - g_i(w, q)]^2 \leq \sum_{l=1}^{n}(x_l - w_l)^2 b_{il}(x, w), \quad (6)$$

$$b_{il}(x, w) \preceq c^{\theta'_{il}}, \quad (7)$$

for all $c > 1$, $x, w \in \mathscr{X}$ and $q \in \mathscr{Q}$.

Our last assumption (H4) is on the existence of a minimum dwell-time. Let $\mathbf{P}_{x,y}$ be the conditional probability given (x, y) defined according to (25) for the process $(x(t), y(t))$ and let $\mathbf{E}_{x,y}$ the corresponding mathematical expectation.

(H4) *for all $x \in \mathscr{X}$ and $y \in \mathscr{Y}$*

$$0 < \inf_k \mathbf{E}_{x,y}\{\zeta_{k+1} - \zeta_k\}. \quad (8)$$

Remark *2.3:* A positive minimum expected dwell time is required. This avoids a number of infinite switching in every finite time interval and has a clear physical meaning. □

Before stating the main result of this paper we need few more notations and objects. Let

$$\alpha_i(s) = \begin{cases} \phi_i\left[-1 + 2^{\frac{3}{2}}\beta\left(\dfrac{s}{\phi_i}\right)\right] & \text{if } s \geq \phi_i \\ r & \text{if } -\phi_i \leq s \leq \phi_i \\ \phi_i\left[1 + 2^{\frac{3}{2}}\beta\left(\dfrac{s}{\phi_i}\right)\right] & \text{if } s \leq -\phi_i \end{cases} \quad (9)$$

where $\beta(s) = \dfrac{s}{\sqrt{1+s^2}}$ and $\phi_i \in \mathbb{R}^+$, $i = 1, \ldots, n$. The function $\alpha_i : \mathbb{R} \to (-2\phi_i, 2\phi_i)$ is a *continuously differentiable* bounded odd function and its main properties, extensively used in this paper, are: for all $r, s \in \mathbb{R}$ $[\alpha_i(s) - \alpha_i(s-r)]^2 \preceq \alpha_i^2(r)$, $\alpha_i^2(s+r) \preceq \alpha_i^2(s) + \alpha_i^2(r)$, $\alpha_i^2(s) \leq 4\phi_i^2$. Any other (even simpler) function α'_i with the same properties as α_i can be used as well in the foregoing results.

Define $\alpha(s) = (\alpha_1(s_1), \ldots, \alpha_n(s_n))$ with $s = (s_1, \ldots, s_n)$ and $g(s) = (g(s_1), \ldots, g(s_n))$ and for each $\Delta_i > 0$, $i = 1, \ldots, n$, let $\mathscr{V} \subset \mathbb{R}^n$ be the finite set defined as follows: for each $q \in \mathscr{Q}$ the points ξ such that

$$\begin{aligned} \xi_1 &= q_1, \\ \xi_i &= q_i + \Delta_i[\alpha_{i-1}(\xi_{i-1}) - q_{i-1}], i = 2, \ldots, n, \end{aligned} \quad (10)$$

are in $\mathscr{V}$. The set $\mathscr{V}$ will be endowed with the discrete topology.

Theorem *2.1: Assume* **(H1)**-**(H4)**. *There exist $c > 1$ and $\Delta_i, \phi_i > 0$, $i = 1, \ldots, n$, such that if*

(L1) *the process $\xi(t) = (\xi_1(t), \ldots, \xi_n(t))$ satisfies*

$$\begin{aligned} &\xi_1^+(t) = \alpha_2(\xi_2(t)) + f_1(\alpha(\xi(t)), r(t)) + \Delta_1[x_1(t) - \xi_1(t)], \\ &\xi_2^+(t) = \alpha_3(\xi_3(t)) + f_2(\alpha(\xi(t)), r(t)) \\ &+\Delta_2[\alpha_1^+(\xi_1(t)) - \xi_2(t) - f_1(\alpha(\xi(t)), r(t))], \\ &\vdots = \vdots \\ &\xi_n^+(t) = f_n(\alpha(\xi(t)), r(t)) \\ &+\Delta_n[\alpha_{n-1}^+(\xi_{n-1}(t)) - \xi_n(t) - f_{n-1}(\alpha(\xi(t)), r(t))], \\ &\xi_1(\zeta_k) = g_1(\alpha(\xi(\zeta_k^-)), q(\zeta_k^-)), \\ &\xi_i(\zeta_k) = g_i(\alpha(\xi(\zeta_k^-)), q(\zeta_k^-)) + \Delta_i[\alpha_{i-1}(\xi_{i-1}(\zeta_k)) \\ &-g_{i-1}(\alpha(\xi(\zeta_k^-)), q(\zeta_k^-))], i = 2, \ldots, n, \end{aligned} \quad (11)$$

(L2) *for each $x \in \mathscr{X}$, $u \in \mathscr{Q}$, $\xi \in \mathbb{R}^n$, $v \in \mathscr{V}$, and $y, w \in \mathscr{Y}$ the limits*

$$\begin{aligned} &\partial P_{x,\xi,y}(u, v, w) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}\mathbf{P}_{x,\xi,y}\{(x(\tau), \xi(\tau), y(\tau)) \\ &= (u, v, w), \tau < \varepsilon\}, \\ &(u, v, w) \neq (x, \xi, y), \\ &\partial P_{x,\xi,y}(x, \xi, y) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon}[\mathbf{P}_{x,\xi,y}\{(x(\tau), \xi(\tau), y(\tau)) \\ &= (x, \xi, y), \tau < \varepsilon\} - 1] \end{aligned} \quad (12)$$

exist bounded uniformly on $\xi \in \mathbb{R}^n$ and $v \in \mathscr{V}$, where $\mathbf{P}_{x,\xi,y}$ is the conditional probability given (x, ξ, y) defined according to (25) for the process $(x(t), \xi(t), y(t))$,

we have

$$\mathbf{P}_{x,\xi,y}\{\lim_{t \to \infty}\|x(t) - \xi(t)\| = 0\} = 1. \quad (13)$$

for all $x, \xi \in \mathcal{X}$ and $y \in \mathscr{Y}$.

Remark *2.4:* Here $\partial P_{x,\xi,y}(u, v, w)$ for $(x, \xi, y) \neq (u, v, w)$ can be interpreted as the instantaneous transition rate of the process $(x(t), \xi(t), y(t))$ into (u, v, w) at the first exit time τ out of $(x, \xi, y) \neq (u, v, w)$. Also, $-\partial P_{x,\xi,y}(x, \xi, y)$ *can be interpreted as the instantaneous transition rate of the process* $(x(t), \xi(t), y(t))$ *at the first exit time* τ *from* (x, ξ, y) *into some other state* $(u, v, w) \neq (x, \xi, y)$. It can be shown that $\sum_{u \in \mathscr{Q}, y, w \in \mathscr{Y}} \partial P_{x,\xi,y}(u, v, w) = 0$, which corresponds to what is commonly known as a *locally regular jump process* ([3]). By definition of $\mathbf{P}_{x,\xi,y}$ (see (25)) we have that

$$\begin{aligned} &\mathbf{P}_{x,\xi,y}\{(x(\tau), \xi(\tau), y(\tau)) = (u, v, w), \tau < \varepsilon\} \\ &= \mathbf{P}_{x,\xi}^{(y)}\{(x(\tau), \xi(\tau)) = (u, v), \tau < \varepsilon\} \end{aligned} \quad (14)$$

Therefore, (L2) is a condition on the instantaneous transition rate of the process $(x(t), \xi(t))$ for fixed values of $y(t)$ (i.e. the frozen process).□

Remark *2.5:* Note the "chained" structure of (11), in which the derivative $\alpha_{i-1}^+(\xi_{i-1}(t))$ is chained to $\xi_i^+(t)$, and the use of $\alpha(\xi(t))$ as a "saturated" estimate of $x(t)$. This "saturated" estimate was used in [7]. The novelty in our paper is to use "saturated" estimates also to determine the reset maps and to reset the estimates as well at switching times.

While at each switching time the state x_i is reset to the value q_i, the estimate ξ_i is reset to the value $\hat{q}_i + \Delta_i(\alpha_{i-1}(\xi_{i-1}) - \hat{q}_{i-1})$, where $\hat{q}_i, \hat{q}_{i-1}$ are estimated values of q_i and q_{i-1}.

A. A motivating example

Before going further, we want to give an example for which our observer design can be applied. Spontaneous synchronization is fundamental for convergence of synchrony in pulsed-coupled oscillators. When coupled to others an oscillator is receptive to the pulses of its neighbours. Coupling between nodes is assumed instantaneous and at some (random) times an oscillator reset its phase and all nodes in the neighbourhood adjust their phase. As model of one oscillator we consider the Van der Pol oscillator (see [5])

$$\begin{aligned} x_1^+(t) &= x_2(t), \\ x_2^+(t) &= -x_1(t) + x_2(t)(1 - x_1^2(t)), \end{aligned} \tag{15}$$

Assume that the measured output process is $\mu(t) = x_1(t)$ and that the phase $x_1(t)$ is reset at random times $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \cdots \zeta_n \uparrow \infty$. At these times the oscillator is said to "fire" and resets itself and the other oscillators in its neighbourhood. The phase of the i-th oscillator is reset either to 0 or $x_1(\zeta_k) = -\frac{\alpha}{\beta}$ according if the i-th or the j-th ($j \neq i$) oscillator in its neighbourhood fires. Here $\alpha = \exp b\nu$, where b is the dissipation factor and ν the amplitude increment, and $\beta = \frac{\exp b\nu - 1}{\exp b - 1}$. The numbers α and β determine the coupling between oscillators.

Assume that the random times $\zeta_1 < \cdots < \zeta_n < \cdots$ are distributed as in a Markov chain and that the expectation of the period between two consecutive times is never below a certain positive number. The state trajectories $x(t)$ of (15) are well defined and bounded for all times, therefore (H1) is satisfied. Also assumption (H2) is satisfied (see remark 2.2). Let $\mathscr{Q} = \{(0,0), (-\frac{\alpha}{\beta}, 0)\}$ and denote by $q(t)$ the jump Markov process with values in $\mathscr{Q}$ and such that $x(\zeta_k) = q(\zeta_k) = g(x(\zeta_k^-), q(\zeta_k^-))$. Therefore, assumption (H3) is satisfied.

Let the estimation process $\xi(t) = (\xi_1(t), \xi_2(t))$ be defined according to (L1) as follows

$$\begin{aligned} &\xi_1^+(t) = \alpha_2(\xi_2(t)) + \Delta_1[x_1(t) - \xi_1(t)] \\ &\xi_2^+(t) = -\alpha_1(\xi_1(t)) + \alpha_2(\xi_2(t))[1 - \alpha_1^2(\xi_2(t))] \\ &+\Delta_2(\alpha_1^+(\xi_1(t)) - \xi_2(t)), \\ &\xi_1(\zeta_k) = g_1(\alpha(\xi(\zeta_k^-)), q(\zeta_k^-)), \\ &\xi_2(\zeta_k) = g_2(\alpha(\xi(\zeta_k^-)), q(\zeta_k^-)) + \Delta_2[\alpha_1(\xi_1(\zeta_k)) \\ &-g_1(\alpha(\xi(\zeta_k^-)), q(\zeta_k^-))], \end{aligned} \tag{16}$$

where $\Delta_i, \phi_i > 0$, $i = 1, 2$. Here, we select $\phi_j = c^{\delta_j}$, $j = 1, 2$, $\Delta_1 = \Delta_2^3 \left(\frac{\Phi_1}{\Phi_2}\right)^2$ and $\Delta_2 = c^{2\delta_3}$, with any δ_j, $j = 1, 2, 3$, satisfying (H2)-(H3), and sufficiently large $c > 1$.

Let $\mathscr{V}$ be the set defined as follows: for each $q \in \mathscr{Q}$ the point ξ such that

$$\begin{aligned} \xi_1 &= q_1, \\ \xi_2 &= q_2 + \Delta_2[\alpha_1(q_1) - q_1] \end{aligned} \tag{17}$$

is also in $\mathscr{V}$. With each $q \in \mathscr{Q}$ we can associate a process $\{\hat{\mathscr{F}}, \hat{\mathscr{N}}, \mathbf{P}_{x,\xi}^{(q)}\}$ on some phase space $\{\hat{\mathscr{X}}, \hat{\mathfrak{B}}\}$ with conditional probability given (x, ξ) equal to $\mathbf{P}_{x,\xi}^{(q)}$. This process describes $(x(t), \xi(t))$ for fixed values of $q(t)$. We assume that the following transition probabilities are given for each $\varepsilon > 0$, $u \in \mathscr{Q}$ and $v \in \mathscr{V}$

$$\mathbf{P}_{x,\xi}^{(q)}\{(x(\tau), \xi(\tau)) = (u, v), \tau < \varepsilon\}, \qquad (u, v) \neq (x, \xi),$$

where τ represents the first exit time from the point (x, ξ) into the point (u, v), and

$$1 - \mathbf{P}_{x,\xi}^{(q)}\{(x(\tau), \xi(\tau)) = (x, \xi), \tau < \varepsilon\}.$$

Also, we assume that their right-hand derivative with respect to ε (i.e. their instantaneous rates) exist bounded uniformly on $\xi \in \mathbb{R}^n$ and $v \in \mathscr{V}$. Therefore, also assumption (L2) is satisfied and theorem 2.1 yields the convergence to zero a.s. of the estimation error for (15)-(16).

III. Outline of the proof of theorem 2.1

The proof of theorem 2.1 is the result of the following observations. We consider the system (1)-(11) as the result of the switching among the members of a family of systems with given $y(s^-) = (r^-, q^-) \in \mathscr{Y}$ and $y(t) = (r, q)$ for all $t \geq s$. We will see that, given $y(s^-) = (r^-, q^-) \in \mathscr{Y}$ and $y(t) = (r, q)$ for all $t \geq s$, (11) is an observer for (1). By switching among the members of the family of observers (11) with given $y(s^-) = (r^-, q^-) \in \mathscr{Y}$ and $y(t) = (r, q)$ for all $t \geq s$ we prove the claimed asymptotic property (13). For proving that given $y(s^-) = (r^-, q^-) \in \mathscr{Y}$ and $y(t) = (r, q)$ for all $t \geq s$ (11) is an observer for (1) we introduce a novel framework for observer design. This framework is based on the notion of *output immersion*. This notion does not correspond to any new observability notion, rather it establishes a symplifying tool for observer design.

Definition *3.1: A system $\Sigma(x, \xi, \mu) : x^+ = f(x), \xi^+ = g(\xi, \mu), \mu = h(x)$, $x \in \mathscr{X} \subset \mathbb{R}^n$, $\xi \in \mathscr{W} \subset \mathbb{R}^n$, $\mu \in \mathscr{M} \subset \mathbb{R}^p$ is said to be output immersed into $\hat{\Sigma}(\hat{x}, \hat{\xi}, \hat{\mu})$: $\hat{x}^+ = \hat{f}(\hat{x}), \hat{\xi}^+ = \hat{g}(\hat{x}, \hat{\xi}, \hat{\mu})$, $\hat{x} \in \hat{\mathscr{X}} \subset \mathbb{R}^n$, $\hat{\xi} \in \hat{\mathscr{W}} \subset \mathbb{R}^n$, $\hat{\mu} \in \hat{\mathscr{M}} \subset \mathbb{R}^{\hat{p}}$, if there exist a continuous mapping $N : \mathscr{X} \times \mathscr{W} \times \mathscr{M} \to \hat{\mathscr{M}}$ and a diffeomorphism $X : \mathscr{X} \times \mathscr{W} \to \hat{\mathscr{X}} \times \hat{\mathscr{W}}$ such that $(\hat{x}, \hat{x}) = X(x, x)$, $x \in \mathscr{X}$, $\hat{x} \in \hat{\mathscr{X}}$, and*

$$\begin{aligned} &\frac{\partial X(x, \xi)}{\partial x} f(x) + \frac{\partial X(x, \xi)}{\partial \xi} g(\xi, \mu) \\ &= \hat{g}(x, X(x, \xi), N(x, \xi, \mu)) \end{aligned} \tag{18}$$

for all $x \in \mathscr{X}$, $\xi \in \mathscr{W}$ and $\mu \in \mathscr{M}$.

Roughly speaking, a system Σ is output immersed into $\hat{\Sigma}$ if the outputs of Σ are mapped into the outputs of $\hat{\Sigma}$ under some state transformation X. One important consequence of the condition $(\hat{x}, \hat{x}) = X(x, x)$, $x \in \mathscr{X}$, $\hat{x} \in \hat{\mathscr{X}}$, is that if the immersed dynamics $\hat{\xi}^+ = \hat{g}(\hat{x}, \hat{\xi}, \hat{\mu})$ is an asympotic observer for the immersed dynamics $\hat{x}^+ = \hat{f}(\hat{x})$ then $\xi^+ = \hat{g}(\xi, \mu)$ is an asympotic observer for $x^+ = f(x)$. This fact is important as far as the observer design for $\hat{x}^+ = \hat{f}(\hat{x})$ is easier than that for $x^+ = f(x)$, as it will be seen in the next sections.

The following result states that, given $y(s^-) = (r^-, q^-) \in \mathscr{Y}$, $x(s^-) = x \in \mathscr{X}$, $\xi(s^-) = \xi \in \mathscr{W}$ and $y(t) = (r, q)$ for all $t \geq s$, (1)–(11) can be immersed into a system, which can be splitted into n one-dimensional decoupled dynamics, each with its own one-dimensional observer. For any vector w we denote by w_i the i-th component of w.

Theorem *3.1: The system (1)-(11), given $y(s^-) = (r^-, q^-) \in \mathscr{Y}$, $x(s^-) = x \in \mathscr{X}$, $\xi(s^-) = \xi \in \mathscr{W}$ and $y(t) = (r, q)$ for all $t \geq s$, can be output immersed into the system*

$$
\begin{aligned}
&x_i^+(t) = \hat{f}_i(x(t), r), i = 1, \ldots, n,\\
&\hat{\mu}_i(t) = x_i(t) + \hat{h}_i(x(t), \hat{\xi}(t), r), i = 1, \ldots, n,\\
&x_i(s) = g_i(x^-, q^-), t \geq s,\\
&\hat{\xi}_i^+(t) = \hat{f}_i(\alpha(\hat{\xi}(t)), r) + \Delta_i(\hat{\mu}_i(t) - \hat{\xi}_i(t)), i = 1, \ldots, n,\\
&\hat{\xi}_i(s) = g_i(\alpha(\hat{\xi}^-), q^-) + \hat{k}_i(x^-, \hat{\xi}^-, q^-), t \geq s,
\end{aligned}
\tag{19}
$$

where

$$
\begin{aligned}
&\hat{f}_i(x, r) = x_{i+1} + f_i(x, r),\ i = 1, \ldots, n-1,\\
&\hat{f}_n(x, r) = f_n(x, r),\\
&\hat{h}_1(x, \hat{\xi}, r) = \frac{\hat{f}_1(\alpha(\Gamma(x, \hat{\xi}))), r) - \hat{f}_1(\alpha(\hat{\xi}), r)}{\Delta_1},\\
&\hat{h}_i(x, \hat{\xi}, r) := f_{i-1}(x, r) - f_{i-1}(\alpha(\Gamma(x, \hat{\xi})), r)\\
&+\Delta_i(x_{i-1} - \alpha_{i-1}(\Gamma_{i-1}(x, \hat{\xi}_1, \ldots, \hat{\xi}_{i-1})))\\
&\quad + \frac{\hat{f}_i(\alpha(\Gamma(x, \hat{\xi})), r) - \hat{f}_i(\alpha(\hat{\xi}), r)}{\Delta_i},\ i = 2, \ldots, n, \quad (20)\\
&\hat{k}_1(x, \hat{\xi}, q) := g_1(\alpha(\Gamma(x, \hat{\xi})), q) - g_1(\alpha(\hat{\xi}), q),\\
&\hat{k}_i(x, \hat{\xi}, q) := g_i(\alpha(\Gamma(x, \hat{\xi})), q) - g_i(\alpha(\hat{\xi}), q)\\
&+\Delta_i[g_{i-1}(x, q) - g_{i-1}(\alpha(\Gamma(x, \hat{\xi})), q)],\ i = 2, \ldots, n,
\end{aligned}
$$

and $\Gamma(x, \hat{\xi})$ is defined recursively as

$$
\begin{aligned}
&\Gamma_1(x, \hat{\xi}_1) = \hat{\xi}_1,\ \Gamma_i(x, \hat{\xi}_1, \ldots, \hat{\xi}_i) = \hat{\xi}_i - \Delta_i(x_{i-1}\\
&-\alpha_{i-1}(\Gamma_{i-1}(x, \hat{\xi}_1, \ldots, \hat{\xi}_{i-1}))),\ i = 2, \ldots, n,\\
&\Gamma(x, \hat{\xi}) = (\Gamma_1(x, \hat{\xi}_1), \ldots, \Gamma_n(x, \hat{\xi}_1, \ldots, \hat{\xi}_n)).
\end{aligned}
$$

The peculiarity of this output immersion is that (19) can be split into n pairs of one-dimensional systems

$$
\begin{aligned}
&x_i^+(t) = \hat{f}_i(x(t), r), x_i(s) = g_i(x^-, y^-),\\
&\hat{\mu}_i(t) = x_i(t) + \hat{h}_i(x(t), \hat{\xi}(t), r) \qquad (21)\\
&\hat{\xi}_i^+(t) = \hat{f}_i(\alpha(\hat{\xi}(t)), r) + \Delta_i(\hat{\mu}_i(t) - \hat{\xi}_i(t)),\\
&\hat{\xi}_i(s) = g_i(\alpha(\hat{\xi}^-), q^-) + \hat{k}_i(x^-, \hat{\xi}^-, q^-) \qquad (22)
\end{aligned}
$$

for $t \geq s$. Therefore an observer for (19) can be designed by selecting only the parameters Δ_i and ϕ_i, $i = 1, \ldots, n$, for each (21)-(22) on account of the fact that (22) has by construction the form of an observer for (21).

In conclusion the proof of theorem 2.1 is structured according the following steps:

- (*Output Immersion*). For each fixed value $y \in \mathscr{Y}$ of $y(t)$ perform the claimed output immersion of (1)-(11) into (19).
- (*Splitting*). Split (19) into n systems (21)-(22). Also, calculate the value of $\mathbf{A}^y W_i(x_i, \hat{\xi}_i)$, where $\mathbf{A}^y$ is the generating operator of the process $(x(t), \hat{\xi}(t), y(t))$ for fixed $y(t) = y$ and $W_i(x_i, \hat{\xi}_i)$ is some at least twice continuously differentiable function, positive definite around $x = \hat{\xi}$.
- (*Parameter selection*). Select the parameters $\Delta_1, \ldots, \Delta_n$ and $\phi_1, \ldots, \phi_n$ in such a way that the value of $\mathfrak{U}[\sum_{i=1}^n W_i(x_i, \hat{\xi}_i)]$ is negative definite around $x = \hat{\xi}$, where $\mathfrak{U}$ is the generating operator of the process $(x(t), \hat{\xi}(t), y(t))$. This determines by the immersion process the values of $\Delta_1, \ldots, \Delta_n$ and $\phi_1, \ldots, \phi_n$ in (19).
- (*Probability chaining*). Chain over consecutive dwell times the stochastic properties of (1)-(11), obtained from $\mathfrak{U}[\sum_{i=1}^n W_i(x_i, \hat{\xi}_i)]$ via the Dynkin's formula over a single dwell time by using the transition probability of the process $(x(t), \hat{\xi}(t), y(t))$ defined according to (25).

IV. The Probabilistic Setting

In this section we introduce a theoretical framework in which processes $(z(t), y(t))$ as those defined in (1)-(11) can be rigourously described and studied. Also, by combining stochastic kernels associated to each 'frozen' process $(z(t), y)$ a probability measure $\mathbf{P}_{x,y}(A)$ is defined on the cylindrical sets A generated by the values of $(z(t), y(t))$ and determines by extension a transition probability on the probability space. Finally, the characteristic operator for $(z(t), y(t))$ will be calculated.

1) A transition probability for the process $(z(t), y(t))$: Let $y(t)$ be a piecewise constant Markov process given on some probability space $\{\mathscr{Y}, \mathfrak{L}, \mathbf{P}_y\}$, where $\{\mathscr{Y}, \mathfrak{L}\}$ is its phase space and $\mathbf{P}_y$ its conditional probability given y. The process $y(t)$ has right-continuous trajectories with Markov times $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \cdots$, $\zeta_n \uparrow \infty$ such that $y(t) = y_k$ if $\zeta_k \leq t < \zeta_{k+1}$. The random time ζ_j represents the time at which $y(t)$ changes its state y_k into y_{k+1}. The measurable space $\{\mathscr{Y}, \mathfrak{L}\}$ is endowed with the discrete topology and $\mathfrak{L}$ contains all the singletons.

Also, let a strong (homogeneous) Markov process $\{\mathscr{F}, \mathscr{N}, \mathbf{P}_x^{(y)}\}$ be associated on some phase space $\{\mathscr{X}, \mathfrak{B}\}$ with each $y \in \mathscr{Y}$, with $\mathbf{P}_x^{(y)}$ its conditional probability given x. We may assume that $\mathscr{Y} \subset \mathscr{X}$. The space of trajectories $\mathscr{F}$ and the σ-algebra $\mathscr{N}$ are the same for all $y \in \mathscr{Y}$. The trajectories of the process $\{\mathscr{F}, \mathscr{N}, \mathbf{P}_x^{(y)}\}$ are assumed to be continuous and will be denoted by $x(t, y)$. By $z(t) = x(t, y(t))$ we denote the trajectories of the process which results from the "interference" of $y(t)$ with $x(t, y)$, $y \in \mathscr{Y}$.

Let $\mathscr{F}^*$ be the set of functions $x^*(t)$ defined on sums of intervals $\bigcup[t_k, t_{k+1})$, $0 = t_0 < t_1 < \cdots < t_n < \cdots$, $t_n \to \infty$ such that for all k there exists a function $x_k^*(t) \in \mathscr{F}$ which satisfies $x^*(t + t_k) = x_k(t)$ for $0 \leq t < t_{k+1} - t_k$ and $\mathscr{N}^*$ is the σ-algebra generated by the cylinders on $\mathscr{F}^*$. Also, let $\mathscr{F}_{\mathscr{Y}}$ be the set of piecewise constant functions $y(t)$ with values in $\mathscr{Y}$. Let $y(t) = y_k$ if $\zeta_k \leq t < \zeta_{k+1}$, $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \cdots$, $\zeta_n \uparrow \infty$.

We define a probability measure on $\mathscr{N}^* \times \mathfrak{L}$, i.e. a probability measure for $(z(t), y(t))$, as follows. If

$$A = \bigcap_{k=0}^{n} \theta_{\zeta_k} C_{r_{k0},\dots,r_{kh_k}}(A_k) \bigcap_{k=1}^{n+1} \theta_{\zeta_k} C_0(B_k) \quad (23)$$

where $\theta_t A$ is the image of the set A under the mapping θ_t, $C_t(B_k) = \{x^*(\cdot) : x(t) \in B_k\}$ and $C_{r_{k0},\dots,r_{k,h_k}}(A_k)$ is a cylindrical set on $\mathscr{F}^*$ generated by the values of $x^*(t)$ on $0 \le t < \zeta_{k+1} - \zeta_k$, i.e. $C_{r_{k0},\dots,r_{kh_k}}(A_k) = \{x^*(\cdot) : (x^*(r_{k0}), \dots, x^*(r_{kh_k})) \in A_k\}$, with $0 \le r_{k0} < r_{k1} < \dots < r_{k,h_k} < \zeta_{k+1} - \zeta_k$, then

$$\begin{aligned}
&\mathbf{P}_x^{(y(\cdot))}(A) = \textstyle\int_{B_1} \mathbf{P}_x^{(y_0)}\{C_{r_{00},\dots,r_{0h_0}}(A_0) \cap C_{\zeta_1}(dx_1)\} \times \cdots \\
&\times \textstyle\int_{B_2} \mathbf{P}_{x_1}^{(y_1)}\{C_{r_{10},\dots,r_{1h_0}}(A_1) \cap C_{\zeta_2-\zeta_1}(dx_2)\} \times \cdots \\
&\times \textstyle\int_{B_n} \mathbf{P}_{x_{n-1}}^{(y_{n-1})}\{C_{r_{n-1,n-1},\dots,r_{n-1,h_{n-1}}}(A_{n-1}) \\
&\cap C_{\zeta_n-\zeta_{n-1}}(dx_n)\} \times \cdots \\
&\times \mathbf{P}_{x_n}^{(y_n)}\{C_{r_{nn},\dots,r_{nh_n}}(A_n) \cap C_{\zeta_{n+1}-\zeta_n}(B_{n+1})\} \qquad (24)
\end{aligned}$$

Some measurability properties of the $\mathbf{P}_x^{(y(\cdot))}(A)$ may be easily proved.

Lemma *4.1: $\mathbf{P}_x^{(y(\cdot))}(A)$ is for each $A \in \mathscr{N}^* \times \mathfrak{L}$ a $\mathfrak{B} \times \mathscr{N}^{\mathscr{Y}}$-measurable function with respect to the arguments $(x, y(\cdot))$, where $\mathscr{N}^{\mathscr{Y}}$ is the σ-algebra generated by the cylinders in $\mathscr{F}^{\mathscr{Y}}$.*

Next, we define

$$\mathbf{P}_{x,y}(A) = \mathbf{E}_y \mathbf{P}_x^{(y(\cdot))}(A) \qquad (25)$$

where $\mathbf{E}_y$ is the expectation calculated with probability $\mathbf{P}_y$, which corresponds to the process $y(t)$, and $\mathbf{E}_y \mathbf{P}_x^{(y(\cdot))}(A)$ makes sense since $\mathbf{P}_x^{(y(\cdot))}(A)$ is $(\mathfrak{B} \times \mathscr{N}^{\mathscr{Y}})$-measurable with respect to the arguments $(x, y(\cdot))$ on account of lemma 4.1. Therefore, also $\mathbf{P}_{x,y}(A)$ is $(\mathfrak{B} \times \mathscr{Y})$-measurable with respect to (x, y).

Since $\mathbf{P}_{x,y}(A)$ defines by virtue of (25) a probability measure on the cylindrical sets generated by the values of $(z(t), y(t))$, with $z(t) = x(t, y_k)$ whenever $y(t) = y_k$ for $\zeta_k \le t < \zeta_{k+1}$, therefore by standard arguments it can be uniquely extended to a probability measure on $(\mathscr{N}^* \times \mathfrak{L})$, i.e. a probability measure for the process $(z(t), y(t))$. It can be shown that the process $(z(t), y(t))$ has a transition probability

$$\begin{aligned}
&\mathbf{P}\{(z(t+s), y(t+s)) \in B \times C | z(t), y(t)\} \\
&= \mathbf{E}_{y(t)} \mathbf{P}_{z(t)}^{(y(\cdot))}\{C_s(B)\} \qquad (26)
\end{aligned}$$

This can be accomplished by checking that (26) satisfies the Chapman-Kolmogorov equation.

2) Characteristic operator of the process $(z(t), y(t))$: We restrict here to the case of compact phase space $\mathscr{X}$ and finite set $\mathscr{Y}$. Let $\mathscr{D}_{\mathfrak{U},(z,y)}$ be the set of bounded Borel functions $V(z, y)$, $(z, y) \in \mathscr{X} \times \mathscr{Y}$, such that the limit

$$\mathfrak{U}V(z, y) = \lim_{n\to\infty} \frac{\mathbf{E}_{z,y} V(z(\tau_n), y(\tau_n)) - V(z, y)}{\mathbf{E}_{z,y}\tau_n} \quad (27)$$

exists where τ_n is the instant of the first exit from U_n, $\{U_n\}$ being an arbitrary sequence of neighborhoods of the point (z, y), $U_n \downarrow (z, y)$. The quantity $\mathfrak{U}V(z, y)$ is the value on the function f at (z, y). The set $\mathscr{D}_{\mathfrak{U}} = \bigcap_{(z,y)\in\mathscr{X}\times\mathscr{Y}} \mathscr{D}_{\mathfrak{U},(z,y)}$ is the domain of definition of the operator $\mathfrak{U}$ and for $f \in \mathscr{D}_{\mathfrak{U}}$ the function $\mathscr{D}_{\mathfrak{U}}V(z, y)$ is defined by (27). On the other hand, let $\mathscr{D}_{\mathbf{A}}$ be the set of bounded Borel functions such that the limit

$$\mathbf{A}V(z, y) = \lim_{h\downarrow 0} \frac{\mathbf{E}_{z,y} V(z(h), y(h)) - V(z, y)}{h} \qquad (28)$$

It can be shown that $\mathscr{D}_{\mathbf{A}} \subset \mathscr{D}_{\mathfrak{U}}$ and $\mathbf{A} = \mathfrak{U}$ on $\mathscr{D}_{\mathbf{A}}$. Moreover, $\mathfrak{U}$ and $\mathbf{A}$ are the *characteristic operator* and, respectively, the *generating operator* of the process $(z(t), y(t))$ ([3]).

By definition τ_n is the instant of the first exit from U_n, $\{U_n\}$ being an arbitrary sequence of neighborhoods of the point (z, y), $U_n \downarrow (z, y)$ and note that in our setting $U_n = U_n' \times \{y\}$, U_n' being a neighborhood of x in $\mathscr{X}$ and $\{y\}$ the singleton in $\mathscr{Y}$ containing y. Therefore, $\tau_n = \min\{\tau_n', \tau\}$, where τ_n' is the exit moment of the process $z(t)$ from the neighborhood $U_n' \times \mathscr{Y}$ and τ is the moment of first exit of $y(t)$ out of the point $y(0)$. We are going to calculate the value of the characteristic operator $\mathfrak{U}V(z, y)$ of the process $(z(t), y(t))$ for all bounded and at least twice continuously differentiable functions $V(z, y)$ (in [4] only the value of the generating operator of $(z(t), y(t))$ has been calculated).

Lemma *4.2: If the limits*

$$\begin{aligned}
&\partial P_{z,y}(u, w) = \lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon} \mathbf{P}_{z,y}\{(z(\tau), y(\tau)) \\
&= (u, w), \tau < \varepsilon\}, \qquad (u, w) \neq (z, y), \\
&\partial P_{z,y}(z, y) = \lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon} [\mathbf{P}_{z,y}\{(z(\tau), y(\tau)) \\
&= (z, y), \tau < \varepsilon\} - 1] \qquad (29)
\end{aligned}$$

exist finite for all $u, w \in \mathscr{Y}$,

$\mathfrak{U}V(z, y) = \mathbf{A}^y V(z, y) + \sum_{u,w\in\mathscr{Y}} V(u, w) \partial P_{z,y}(u, w).$
In the case of (1)

$$\mathbf{A}^y V(z, y) = \frac{\partial V}{\partial z}(z, y)(Az + f(z, r) \qquad (30)$$

with $y = (r, q)$.

References

[1] S. Battilotti, A. De Santis, Dwell time controllers for stochastic systems with switching Markov chain, *Automatica*, **41**, pp. 923-934, 2005.

[2] M.D.Di Benedetto, S.Di Gennaro, A. D'Innocenzo, Error detection within a specific time horizon and application to air traffic management, *Proceedings of the 43rd IEEE Conference on Decision and Control*, pp.7472-7477 , December 12-15, Sevilla, Spain, December 2005.

[3] I. I. Gihman, A. V. Skorohod, *The theory of Stochastic Processes*, vol. II, Springer Verlag, Berlin Heidelberg New York, 1975.

[4] J.P. Hespanha, A model for stochastic hybrid systems with application to communication networks, *Nonlinear Analysis*, **62**, pp. 1353-1383, 2005.

[5] M. Krstic, I. Kanellakopoulos, P. Kokotovic, *Nonlinear and Adaptive Control Design*, John Wiley & Sons, Inc., 1995.

[6] R.S. Lipster, A.N. Shiryayev, *Statistics of random processes I: general theory*, Springer Verlag, , Berlin Heidelberg New York, 1977.

[7] B. Yang, W. Lin, Semiglobal ouput feedback stabilization of non-uniformly observable and non-smoothly stabilizable systems, *Proc. Joint 44th Conf. Dec. and Contr. and European Contr. Conf.*, pp. 4207-4212, 2005, Seville, Spain.