A New Class of Hybrid Dynamical Systems: State Estimators with Bit-Rate Constraints

Nikolai G. Dokuchaev and Andrey V. Savkin

School of Electrical Engineering and Telecommunications, University of New South Wales,
Sydney, 2052, Australia

email: a.savkin@unsw.edu.au

ABSTRACT: This paper introduces a new class of hybrid dynamical systems. We consider a state estimation problem involving bit-rate communication capacity constraints for a discrete-time partially observed system. The observation must be coded and transmitted via a digital communication channel with a limited capacity. A recursive coder-estimator is proposed and investigated. An upper bound for the average estimation error is derived, and convergence properties are analyzed.

KEYWORDS: hybrid systems, state estimation, communication bit-rate constraints, Kalman filter

AMS (MOS) subject classification: 62M10, 93E11, 94A40

1. INTRODUCTION

In classical filtering theory (see e.g. Anderson and Moore [1]), the standard assumption is that all data transmission required by the algorithm can be performed with infinite precision. However, in some new models, it is common to encounter situations where observation and control signals are sent via a communication channel with a limited capacity. This problem may arise when a large number of mobile units need to be controlled remotely by a single decision maker. Since the radio spectrum is limited, communication constraint are a real concern. In Stiwell and Bishop [7], the problem of design of large-scale control systems for platoons of underwater vehicles highlights the need for control strategies that address reduced communications, since communication bandwidth is severely limited underwater. Another class of examples is offered by complex networked sensor systems containing a very large number of low power sensors. Furthermore, nowadays, it is becoming more common to use networks in systems, especially in those that are large-scale and physically distributed. All these new engineering applications motivate development of a new chapter of control and state estimation theory in which control and communication issues are combined together, and all the limitations of the communication channels are taken into account. Communications requirements,
especially regarding bandwidth limits, are often challenging obstacles to control systems design. In these problems, classical Kalman estimation theory cannot be applied since the estimator only observes the transmitted sequence of finite-valued symbols. In fact, we need to design a hybrid dynamical system which consists of two subsystems. The first subsystem, that is called Coder, receives real-valued measurements and converts them into a finite-valued symbolic sequence which is sent over the limited capacity communication channel. The second subsystem (Decoder) receives this symbolic sequence and converts it into a real-valued state estimate. In other words, such state estimators with bit-rate constraints form an important subclass of so-called hybrid dynamical systems. In general, hybrid systems are those that combine continuous and discrete event dynamics and involve both real and symbolic variables; e.g., see Matveev and Savkin [3].

A natural question to ask is how much communication capacity is needed to achieve a specified estimation accuracy. The problem studied in this paper was introduced by Wong and Brockett [8], where some algorithms and models were proposed and investigated for the case of bounded random disturbances. A case of decreasing Gaussian disturbances was studied by Nair and Evans [4], where the idea to code the Kalman state estimate was proposed. However, the main results of these papers were restricted to the case of scalar systems.

In this paper, we investigate a state estimation problem involving constraints on bit-rate communication capacity for a discrete-time partially observed system of an arbitrary order with non-decreasing Gaussian disturbances. It is assumed that the observation must be coded and transmitted via a digital communication channel with a limited capacity. A recursive estimation algorithm is proposed and investigated for the case when the system may be unstable. In this case, any large deviation of disturbances implies increasing all the following values of the state vector. We show that our algorithm provides state estimation with a bounded average error. Moreover, we obtain sufficient conditions of a convergence of the average error to zero as the digital communication channel capacity increases. As in the paper of Nair and Evans [4], our recursive coder-estimator includes the Kalman state estimator. It should be pointed out, that the proposed state estimation method is different from those described in literature; it is computationally non-expansive and easy to implement in real time. The most restrictive feature is that the algorithm is not adaptive to reducing of noise to zero, i.e., there is a given minimal level of tracking error which remains fixed even if the noise dissipates with time; we are not able to extend our proofs for a modification of the algorithm without this feature. (It can be added that we provided some numerical experiments with such modifications, but it appears that they lost stability; moreover, we have not found any proofs for other algorithms in literature for the case considered). However, the algorithm is adaptive to changing (i.e., increasing or decreasing) of the noise level above some minimal level which corresponds the minimal level of error. The obtained results can be extended to the case of uncertain linear systems (see e.g. Savkin and Petersen
The remainder of this paper proceeds as follows. In Section 2, we introduce the class of systems under consideration and state the problem of estimation via limited capacity communication channels. Section 3 contains some well-known properties of the Kalman state estimator. In Section 4, we formulate the state estimation problem with communication constraints for a fully-observed system. Section 5 presents our recursive coding-estimation scheme. The main results of the paper are given in Section 6. Section 7 presents an illustrative example. Section 8 contains brief conclusions. The proofs of all the results of Section 6 are given in Appendix.

2. PROBLEM STATEMENT

Consider the following discrete-time linear system

\[
\begin{align*}
X_{t+1} &= A_{t+1} X_t + B_t W_t, \\
Y_t &= H_t X_t + D_t W_t,
\end{align*}
\]  

(2.1)

where \( X_t \in \mathbb{R}^n \) is the state, \( W_t \in \mathbb{R}^d \) is the random disturbance input, \( Y_t \in \mathbb{R}^m \) is the measured output, \( t = 0, 1, 2, \ldots \).

We assume that the vectors \( X_0 \) and \( W_t \) are Gaussian, \( \mathbb{E}[W_t] = 0 \), and \( W_t \) does not depend on \( W_0, W_1, \ldots, W_{t-1} \) and \( X_0 \).

Suppose estimates of the current state are required at a distant location, and are to be transmitted via a digital communication channel such that only \( M \) bits of data may be sent at each time \( t \). We consider a system which consists of the coder, the transmission channel, and the decoder. Using an observation of \( Y_1, \ldots, Y_t \), the coder produces a \( M \)-bit word \( h_t \) which is transmitted via the channel and then received by the decoder; the decoder produces an estimate \( \hat{X}_t \) which depends only on \( h_1, \ldots, h_t \).

Let \( A \) be the set \( A = \{ h \} \) of words \( h = (h^{(1)}, \ldots, h^{(M)}) \), such that \( h^{(i)} \in \{0, 1\} \). The set \( A \) consists of \( 2^M \) elements.

Let \( h_t \in A \) be the signal which is produced by the coder, \( \hat{X}_t \) be an estimate of \( X_t \) which is produced by the decoder.

Introduce the following vector and matrix norms:

\[
\| x \| \overset{\Delta}{=} \max_{i=1, \ldots, n} |x^{(i)}| \quad \text{for} \quad x \in \mathbb{R}^n, \quad \| A \| \overset{\Delta}{=} \max_{i=1, \ldots, n} \sum_{j=1}^{n} |A^{(i,j)}| \quad \text{for} \quad A \in \mathbb{R}^{n \times n}.
\]

We consider this problem as a problem of choosing the deterministic measurable functions \( \Phi_t : (\mathbb{R}^m)^t \rightarrow A \) and \( F_t : (A)^t \rightarrow \mathbb{R}^n \), \( t = 1, 2, \ldots \) such that

\[
\begin{align*}
h_t &= \Phi_t (Y_1, Y_2, \ldots, Y_t) \in A, \\
\hat{X}_t &= F_t (h_1, h_2, \ldots, h_t) \in \mathbb{R}^n,
\end{align*}
\]
and the following estimate holds:

\[ \mathbb{E}|\hat{X}_t - X_t| \leq \text{const} \quad (\forall t > 0). \]

Here \(|\cdot|\) denotes the standard Euclidean norm. The main difficulty of this estimation problem is in a case of non-stable systems, when any large deviation of disturbances implies increasing all the following values \(|X_t|\).

It is well known that under some standard assumptions on \(A_t, B_t, H_t, D_t\), there exists so-called Kalman estimate \(X_t^{KE}\) of \(X_t\) which minimizes the average error \(\mathbb{E}|X_t^{KE} - X_t|^2\).

In this paper we propose an estimation algorithm which involves the Kalman estimation. The Kalman estimate is supposed to be computed, coded, transmitted via the channel and then decoded. The block diagram of our state estimation system is shown in Figure 2.1.

Figure 2.1: Block diagram of the estimator.

3. BASIC PROPERTIES OF THE KALMAN ESTIMATE

It is well known that the Kalman estimate \(X_t^{KE}\) satisfies the following equations:

\[
\begin{align*}
X_t^{KE} &= A_t X_{t-1}^{KE} + V_t, \\
X_0^{KE} &= \mathbb{E}X_0,
\end{align*}
\]

where

\[ V_{t+1} = P_t (Y_t - H_t X_t^{KE}). \]

The matrix \(P_t\) is calculated recursively from the corresponding Riccati equation (see, e.g., Anderson and Moore [1]) and is uniformly bounded under some standard assumptions on the system. The estimation error

\[ \Delta_t \doteq X_t - X_t^{KE} \]

is independent on \(X_t^{KE}\). The vectors \(X_t^{KE}, \Delta_t\) are Gaussian, \(\mathbb{E}\Delta_t \equiv 0\), \(\mathbb{E}|\Delta_t|^2 \leq \text{const}\) (see, e.g., Anderson and Moore [1]).

Let us discuss basic properties of \(V_t\). We have that

\[ V_{t+1} = P_t (Y_t + H_t \Delta_t - H_t X_t) = P_t (D_t W_t + H_t \Delta_t). \]

Hence the vectors \(V_{t+1}\) are Gaussian, \(\mathbb{E}V_t \equiv 0\), \(\mathbb{E}|V_t|^2 \leq \text{const} \quad (\forall t > 0)\).
Furthermore,
\[
\Delta_{t+1} = A_{t+1} \Delta_t + B_t W_t - V_{t+1} \\
= A_{t+1} \Delta_t + B_t W_t + P_t H_t X_{t+1}^{KE} - P_t Y_t \\
= A_{t+1} \Delta_t + B_t W_t + P_t H_t (X_t - \Delta_t) - P_t Y_t.
\]

Hence
\[
\Delta_{t+1} = (A_{t+1} - P_t H_t) \Delta_t + B_t W_t - P_t D W_t.
\]

By the assumptions on \( W_t \), we have that \( \mathbb{E}(X_{s}^{KE})^t \Delta_t = \mathbb{E}(X_{s}^{KE})^t \Delta_s = 0 \) for \( s \leq t \). Hence \( X_{s}^{KE} \) is independent on \( \Delta_t \), and \( \mathbb{E}(X_{s}^{KE})^t P_t H_t \Delta_t = 0 \). Moreover,
\[
\mathbb{E}(X_{s}^{KE})^t P_t D_t W_t = 0,
\]
\[
\mathbb{E}(X_{s}^{KE})^t V_{t+1} = \mathbb{E}(X_{s}^{KE})^t P_t H_t \Delta_t + \mathbb{E}(X_{s}^{KE})^t P_t D_t W_t = 0.
\]

Then the vector \( V_t \) is independent on \( X_{t,1}^{KE}, X_{t,2}^{KE}, \ldots, X_{t,-1}^{KE} \). As stated earlier, the vectors \( V_{t+1} \) are Gaussian, \( \mathbb{E} V_t \equiv 0, \mathbb{E} |V_t|^2 \leq \text{const} \quad (\forall t > 0) \).

It may be concluded that the initial problem of estimation \( X_t \) may be stated as follows: *Estimate the state of the fully-observed system (3.1) under bit-rate constraints in the case of Gaussian disturbances \( V_t \) which do not depend on the previous states and have bounded variance*.

### 4. STATE ESTIMATION FOR THE FULLY-OBSERVED SYSTEM

Consider the process
\[
\begin{cases}
  x_t = A_t x_{t-1} + b_t + v_t, \\
  x_0 = \pi_0.
\end{cases}
\]

(4.1)

Here \( \pi_0 \) is a deterministic vector, \( v_t \) are random disturbances, \( t \geq 0, x_t, v_t, b_t \in \mathbb{R}^n \) and \( A_t \in \mathbb{R}^{n \times n} \). We assume that \( A_t, b_t \) and \( \pi_0 \) are known.

In this paper, \( [\alpha] \) denotes the integer part of a real number \( \alpha > 0 \), such that
\[
[\alpha] = \max\{z \in \mathbb{Z} : z \leq \alpha\}.
\]

Let
\[
\mathcal{G} \triangleq 2^{M-1}, \quad \nu \triangleq \lfloor \mathcal{G}^{1/n} \rfloor, \quad \theta \triangleq \nu^n, \quad L \triangleq \sup_{t \geq 0} \|A_t\|.
\]

We suppose that the following assumptions hold:

**Assumption 4.1** \( \nu \geq 1, \nu > L \).

**Assumption 4.2** *Vectors \( v_t \) are Gaussian, \( v_t \) is independent on \( x_1, x_2, \ldots, x_{t-1} \),*
\[
\mathbb{E} v_t = 0, \quad \mathbb{E} \|v_t\|^2 \leq \delta^2 \quad (\forall t \geq 0),
\]

*where \( \delta > 0 \) is a given constant.*
Also, we assume without a loss of generality that $L > 1$.

5. ESTIMATION ALGORITHM

The estimate $\hat{x}_t$ of the process $x_t$ will be found as a solution of the following equations:

$$
\begin{cases}
\hat{x}_t = A_t \hat{x}_{t-1} + b_t + c_t, & t > 0, \\
\hat{x}_0 = \pi_0,
\end{cases}
$$

Here

$$c_t = C_t(h_1, ..., h_l) \in \mathbb{R}^n, \quad h_t = \Phi_t(x_1, ..., x_t) \in \mathcal{A}
$$

The words $b_t$ are to be calculated by the decoder. The vectors $c_t$ are to be calculated by the decoder. In (5.2), $\mathcal{A}$ is the set of words introduced in Section 2, $C_t : \mathcal{A}^l \rightarrow \mathbb{R}^n$, $\Phi_t : \mathbb{R}^m \rightarrow \mathcal{A}$ are deterministic measurable functions. We consider the problem as a problem of choosing the functions $C_t()$, $\Phi_t()$ such that $E\|x_t - \hat{x}_t\|^r \leq \text{const} \ (\forall t > 0)$ for a given constant $r \geq 1$.

We assume below that the set $\mathcal{A}$ is the set of pairs $\mathcal{A} = \{[(\gamma, s)], \gamma = 0$ or $\gamma = 1, s \in \{1, ..., \bar{s}\}$. Note that the set $\mathcal{A}$ consists of $2\bar{s} = 2^M$ elements and $\theta \leq \bar{s}$.

Furthermore, let numbers $l > 0$, $a > 1$ and an integer $R \geq 1$ be given parameters.

For any $\lambda > 0$, set $D(\lambda) \triangleq \{x \in \mathbb{R}^n : \|x\| \leq \lambda\}$.

Consider a discrete subset $\tilde{D}(\lambda) = \{y_j(\lambda)\}_{j=1}^{\theta} \subset D(\lambda)$ such that for any $x \in D(\lambda)$ there exists a vector $y \in \tilde{D}(\lambda)$ such that $\|x - y\| \leq \lambda^{\nu-1}$. It can be easily seen that such a subset $\tilde{D}(\lambda)$ does exist.

For any $\lambda > 0$, introduce the following maps $S_1(x, \lambda) : \mathbb{R}^n \rightarrow \{1, ..., \theta\}$, $S_2(x, \lambda) : \mathbb{R}^n \rightarrow \{1, ..., \bar{s}\}$ and $F(x, \lambda) : \mathbb{R}^n \rightarrow \tilde{D}(\lambda)$:

$$
S_1(x, \lambda) = \min \left\{ \arg \min_{j \in \{1, ..., \theta\}} \|y_j(\lambda) - x\| \right\},
$$

$$
S_2(x, \lambda) = \max \left\{ k \in \{1, ..., \bar{s}\} : x \notin D(\lambda R^{k-1}) \right\},
$$

$$
F(x, \lambda) = y_j(\lambda), \text{ where } j = S_1(x, \lambda).
$$

Note that if $x \in D(\lambda)$ and $S(x, \lambda) = j$ then $\|y_j(\lambda) - x\| \leq \|y_i(\lambda) - x\| \ (\forall i = 1, ..., \theta)$ and $\|y_j(\lambda) - x\| \leq \lambda^{\nu-1}$. If $x \in D(\lambda)$ then $\|F(x, \lambda) - x\| \leq \lambda^{\nu-1}$.

Introduce the following vectors:

$$
\tilde{z}_t \triangleq x_t - A_t \tilde{x}_{t-1} - b_t, \quad t \geq 1.
$$

Let $l_0 \triangleq 1$. Then the following sequence of $h_t$, $l_t, \gamma_t$ is to be computed:

(i) The coder produces a word $b_t = (\gamma_t, s_t)$ and a number $l_t$, where

$$
\gamma_t = \begin{cases}
0 & \text{if } \tilde{z}_t \in D(l_{t-1}) \\
1 & \text{if } \tilde{z}_t \notin D(l_{t-1})
\end{cases}, \quad s_t = \begin{cases}
S_1(\tilde{z}_t, l_{t-1}) & \text{if } \gamma_t = 0 \\
S_2(\tilde{z}_t, l_{t-1}) & \text{if } \gamma_t = 1,
\end{cases}
$$

\[\text{Algorithm 5.1: Estimation Algorithm}\]
\[ l_t = \begin{cases} 
\frac{l_{t-1}}{a} & \text{if } \gamma_t = 0, l_{t-1} > l/a \\
l_{t-1} & \text{if } \gamma_t = 0, l_{t-1} = l/a \\
a^{\mu \gamma_t}l_{t-1} & \text{if } \gamma_t = 1.
\end{cases} \quad (5.4) \]

(ii) The word \( h_t \) is transmitted via the channel.

(iii) The decoder computes \( l_t \) by the rule (5.4), and then it calculates

\[ c_t = \begin{cases} 
F(h_t, l_{t-1}) & \text{if } \gamma_t = 0 \\
0 & \text{if } \gamma_t = 1.
\end{cases} \quad (5.5) \]

(iv) Finally, the decoder computes \( \hat{x}_t \) by the formula (5.1).

6. THE MAIN RESULTS

In this section we show how to choose parameters \( l, a, R \) of the state estimation algorithm from Section 5 to guarantee that the average estimation error is bounded or converges to zero.

Introduce the process of the estimation error

\[ z_t = x_t - \hat{x}_t. \]

**Theorem 6.1** Consider the system (4.1) and the estimation algorithm described in the Section 5 with given parameters \( a > 1 \) and \( R \geq 1 \). Suppose that

\[ L \in \left( \frac{\nu}{a}, a \tilde{\rho} \right). \quad (6.1) \]

Then for any \( r \geq 1 \) there exists a parameter \( l > 0 \) and a constant \( C_s > 0 \) such that

\[ \mathbb{E}\|z_{rT}\| \leq C_s \quad \forall T > 0. \quad (6.2) \]

Starting from now, we assume that \( r > 1 \) is fixed. To formulate sufficient conditions on \( l \) to guarantee (6.2) we need to introduce the following constants: Let \( r \geq 1 \) be a given number, and

\[ \mu \overset{\Delta}{=} \min\{ i \in \{2, 4, 6, \ldots\} : i > r \}, \quad (6.3) \]

\[ h \overset{\Delta}{=} \frac{\nu - La}{\nu}, \quad G_i \overset{\Delta}{=} \frac{\sqrt{2\pi(i/2)!}2^{i/2}}{\sqrt{i}} \quad \text{for } i = 2, 4, \ldots. \quad (6.4) \]

**Theorem 6.2** Suppose the assumptions (6.1) hold and the parameter \( l > 0 \) satisfies the inequality

\[ l^{\mu} \geq \frac{h^{\mu}G_\mu}{h^{\mu}(1-a^{-r})(a^{\mu-r}-1)}. \quad (6.5) \]

Then there exists a constant \( C_s \) such that the inequality (6.2) holds.
To give an upper estimate of $C_s$, we need to introduce the following constants:

$$\beta \Delta a^{b^p}, \quad C_0 \Delta \frac{\beta^2}{(L - 1)^2}, \quad (6.6)$$

Furthermore, for any $\rho \geq 1$, introduce the following constants:

$$q_\rho \Delta \min \left\{ i \in \{2, 4, 6, \ldots \} : \quad i > \rho, \quad \frac{\beta^p L^i}{\beta^p L^i} < 1 \right\}, \quad Q_\rho \Delta \frac{\beta^p L^i}{\beta^p L^i}, \quad (6.7)$$

$$\bar{\sigma}(\rho) \Delta \sup_{t \geq 0} (E\|v_t\|^p)^{1/p}, \quad f(\rho) \Delta \beta^p \left( 1 + \frac{G_{\rho L^i} \beta^p L^i Q_\rho}{\nu(\beta - L)^{\rho} - 1 - Q_\rho} \right), \quad (6.8)$$

**Theorem 6.3** Under the assumptions of Theorem 6.2, the inequality (6.2) holds with

$$C_s = \frac{V}{\nu^p} + \left( 3f(r)^{1/r} + 2f(r)^{1/r} \frac{\bar{\sigma}(r)}{L - 1} \right), \quad (6.9)$$

where $\rho > 1$ is an arbitrary number, $\rho' \Delta (\rho - 1)^{-1}$.

The following theorem gives sufficient condition for a convergence for the case of increasing channel capacity $M$.

**Theorem 6.4** Suppose a time $T > 0$ is fixed, and $t^{-1} + \nu^{-1} \to 0$, where $\nu = \left\lfloor \frac{M-1}{n} \right\rfloor$. Then $E\|z_t\|^r \to 0$ uniformly on $t \leq T$.

The proofs of all the results of this section are given in Appendix.

### 7. ILLUSTRATIVE EXAMPLE

To illustrate the results of this paper, we consider a deconvolution problem similar to those considered in Chen and Chen [2]. The block diagram of deconvolution system is shown in Figure 7.1.

![Diagram](image_url)

**Figure 7.1**: Deconvolution system.

Combining the signal model and the channel model, we obtain the following system:

$$
\begin{bmatrix}
X_{t+1}^{(1)} \\
X_{t+1}^{(2)} \\
X_{t+1}^{(3)}
\end{bmatrix} =
\begin{bmatrix}
1.98 & -1 & 0 \\
1 & 0 & 0 \\
0.4 & 0 & 0.2
\end{bmatrix}
\begin{bmatrix}
X_t^{(1)} \\
X_t^{(2)} \\
X_t^{(3)}
\end{bmatrix} +
\begin{bmatrix}
0.707 \\
0 \\
0
\end{bmatrix}
n_t^{(1)},
$$

$$Y_t = X_t^{(3)} + n_t^{(1)}.$$
In this state space description, $X^{(1)}_t$ and $X^{(2)}_t$ are the state variables of the signal model. In addition, $u_t = X^{(3)}_t$ is the state variable of the channel model, $Y_t$ is the measured signal. To apply our results to this deconvolution problem, we consider a corresponding system of the form (2.1). In this case, the matrices $A_t$, $B_t$, $K_t$, $D_t$ are given by

$$A_t = \begin{bmatrix} 1.98 & -1 & 0 \\ 1 & 0 & 0 \\ 0.4 & 0 & 0.2 \end{bmatrix}, \quad B_t = \begin{bmatrix} 0.707 \\ 0 \\ 0 \end{bmatrix},$$

$$K_t = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad D_t = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. $$

Also, the constants $n$, $d$, $L$ are given by $n = 3$, $d = 2$, $L = 2.98$.

To illustrate the performance of our coder-estimator, we consider the Gaussian noise signal $W_t = (n^{(1)}_t, n^{(2)}_t)$ with $\mathbb{E}W_t = 0$, $\mathbb{E}W_tW'_t = 0.8 \cdot I$. Also, we take initial condition $X_0 = (1, -1, 1)'$. We consider the system in the cases of a communication channel with capacity $M = 8$ and $M = 10$ bits. We apply the estimation algorithm from Section 5 with the parameters $l = 10.2$, $R = 20$, and $a = 1.3$ for $M = 8$, $a = 1.53$ for $M = 10$. Figure 7.2 shows the true value $u_t$, the Kalman estimate, and the resulting estimates of the signal $u_t$ for times $t = 150, 200$ in the cases of communication channel with capacity $M = 8$ and $M = 10$ bits. Figure 7.3 shows the true value $u_t$ and the resulting estimate of the signal $u_t$ for $t = 1, 200$ in the case of the communication channel with capacity $M = 10$ bits.

8. CONCLUSIONS

This paper describes a new class of hybrid dynamical systems. It considers a state estimation problem involving bit-rate communication capacity constraints for a discrete-time partially observed system. The observation must be coded and transmitted via a digital communication channel with a limited capacity. Classical estimation theory cannot be applied since the estimator only observes the transmitted sequence of finite-valued symbols. A recursive estimation algorithm is proposed and investigated. We show that our algorithm provides state estimation with a bounded average error which converges to zero as the digital communication channel capacity increases. The proposed state estimation method is computationally non-expansive and easy to implement in real-time systems.

9. ACKNOWLEDGEMENT

This work was supported by the Australian Research Council.

REFERENCES

Figure 7.2: Estimates of $u_t$ for $t = 140, 200$; "-" - true value of $u_t$; "- - -" - Kalman estimate; "oox" - estimate for 8 bit channel; "...." - estimate for 10 bit channel


Figure 7.3: Estimates of $u_i$ for $t = 1, 2, \ldots$ - true value of $u_i$; "..." - estimate for 10 bit channel


APPENDIX: PROOFS

Let $l_i$, $c_i$ and $z_i$ be computed by the algorithm, $t = 1, 2, \ldots$

Introduce a sequence $\{s_i, \tau_i, t_i\}_{i=0}^{+\infty}$ of triplets of integer random times, such that the following conditions hold:

(i) $s_0 = \tau_0 = 0$, $1 \leq s_i \leq \tau_i < t_i < s_{i+1}$ ($\forall i \geq 1$);

(ii) if $s_1 = 1$ then $t_0 = 0$; if $s_1 > 1$ then $t_0 = 1$;

(iii) if $t \notin \cup_{i \geq 1} \{s_i, s_i + 1, \ldots, t_i - 1, t_i\}$ and $t > 0$ then $l_t = l/a$;

(iv) if $t \in \cup_{i \geq 1} \{s_i, s_i + 1, \ldots, \tau_i - 1, \tau_i\}$ then $l_t > l/a$, $\gamma_l = 1$;

(v) if $t \in \cup_{i \geq 1} \{\tau_i + 1, \ldots, t_i - 1\}$ then $l_t > l/a$, $\gamma_l = 0$;

(vi) if $t \in \cup_{i \geq 1} \{t_i: t_i < s_{i+1} - 1\}$ then $l_t = l/a$, $\gamma_l = 0$. 
Introduce the following random sequences:

\[ \xi_i \triangleq \tau_i - s_i + 1, \quad \eta_i \triangleq t_i - \tau_i \quad \text{for} \quad i = 0, 1, 2, \ldots, \]

\[ \zeta_0 \triangleq -\eta_0, \quad \zeta_i \triangleq \zeta_{i-1} + \log_a \frac{l_n}{l_{s_i-1}} \quad \text{for} \quad i = 1, 2, \ldots. \quad (A.1) \]

We assume that

\[ z_{-1} = 0, \quad s_{-1} = \tau_{-1} = t_{-1} = 0, \quad \xi_{-1} = 0, \quad \eta_{-1} = 0, \quad \zeta_{-1} = 0. \]

Let \( T > 0 \) be a fixed deterministic integer number. Introduce random integer variables \( m \) and \( k \) such that \( s_m \leq T < s_{m+1} \) and \( k = m - 1 \). Let \( \mathcal{F}' \) be the \( \sigma \)-algebra of random events which is generated by the random values \( \{\zeta_k, \tau_k\} \), and \( \mathcal{F}'' \) be the \( \sigma \)-algebra of random events which is generated by the random values \( \{\zeta_k, \tau_k, \eta_m, s_m, z_{s_m-1}\} \).

We will use the notation \( \text{Ind} \) for the indicator function.

**Proposition A.1** If \( \gamma_l = 0 \) then \( \|z_t\| \leq \frac{l_{t-1}}{\nu} \leq l_t. \)

**Proof.** The equation (5.3) can be rewritten as follows:

\[ \tilde{z}_t = A_t x_{t-1} - A_t \tilde{x}_{t-1} + v_t, \quad (A.2) \]

Hence

\[ \tilde{z}_t = A_t z_{t-1} + v_t. \quad (A.3) \]

We have that \( z_t = \tilde{z}_t - \alpha_t, \)

\[ z_t = A_t z_{t-1} + v_t - \alpha_t, \quad z_0 = 0. \quad (A.4) \]

Then \( \|z_t\| = \|l_{t-1}/\nu\| \), and \( \nu > a \). Hence \( \|z_t\| \leq l_t. \) This completes the proof. \( \square \)

**Proposition A.2** If \( \tau_j \leq t < s_{j+1}, \ j \geq 1, \) then \( l_t \leq l_{a^j \tau_j \wedge (t-\tau_j)}. \) If \( s_j \leq t \leq \tau_j, \ j \geq 1, \) then

\[ l_{a^j \tau_j \wedge (t-\tau_j)} \leq l_t \leq la^j \tau_j \wedge (t-\tau_j), \quad j = 1, 2, \ldots. \]

**Proof.** Proposition A.2 follows immediately from the description of the algorithm. \( \square \)

**Proposition A.3** The random variable \( \text{Ind} \{k \geq 1\} \) is \( \mathcal{F}' \)-measurable and \( \mathcal{F}'' \)-measurable.

**Proof.** We have that \( k < 1 \) if and only if \( \zeta_k \leq 0, \ \tau_k = 0. \) It means that \( \text{Ind} \{k \geq 1\} \) is \( \mathcal{F}' \)-measurable. Furthermore, \( \mathcal{F}' \subset \mathcal{F}'' \), hence \( \text{Ind} \{k \geq 1\} \) is \( \mathcal{F}'' \)-measurable. This completes the proof of Proposition.
Proposition A.4 Let \( \psi = \left( \psi^{(1)}, \psi^{(2)}, \ldots, \psi^{(n)} \right) \) be a Gaussian random \( n \)-dimensional vector such that
\[
|\mathbf{E}\psi^{(i)}| \leq \alpha, \quad \text{Var} \psi^{(i)} \leq c^2 \quad \forall i = 1, \ldots, n,
\]
where \( \alpha \geq 0, \ c > 0 \) are fixed. Then the following estimate holds:
\[
P(\|\psi\| > u) \leq \frac{G_n c^\nu}{u - \alpha} \quad \forall u > \alpha, \quad \forall \nu = 2, 4, 6, \ldots,
\]
where the constants \( G_n \) are defined by (6.4).

Proof. (i) Let \( n = 1, \ \mathbf{E}\psi = 0, \ \mathbf{E}\psi^2 = 1 \). Then
\[
P(\psi > u) = P(\psi < -u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-t^2/2} dt
\]
for all \( q = 1, 2, \ldots \). This completes the proof of Proposition for this case.

(ii) Let \( n = 1, \ \mathbf{E}\psi = 0, \ \mathbf{E}\psi^2 = c^2 \). Let \( \psi/c = \psi_c \). Then
\[
P(\psi > u) = P(\psi < -u) = P\left(\psi_c > \frac{u}{c}\right) = P\left(\psi_c < -\frac{u}{c}\right) \leq \frac{1}{\sqrt{2\pi}} \frac{q^{1/2} c^{2q}}{u^{-2q}}
\]
for all \( q = 1, 2, \ldots \). This completes the proof of Proposition for this case.

(iii) Let \( n = 1, \ \mathbf{E}\psi = \overline{\psi}, \ |\overline{\psi}| \leq \alpha \). Let \( \psi_0 = \psi - \overline{\psi}, \ \mathbf{E}\psi_0^2 = c^2 \). Then \( \mathbf{E}\psi_0 = 0, \) and
\[
P(\|\psi\| > u) = P(\psi_0 < -u - \overline{\psi}) + P(\psi_0 > u - \overline{\psi})
\]
for all \( q = 1, 2, \ldots \). This completes the proof of Proposition for this case.

(iv) Let \( n > 1 \), then
\[
P(\|\psi\| > u) = P(\exists i: \ |\psi^{(i)}| > u) \leq n \max_{i=1,\ldots,n} P(\|\psi^{(i)}\| > u) \leq \sqrt{\frac{n}{\pi}} \frac{q^{1/2} c^{2q}}{u^{-2q}}
\]
for all \( q = 1, 2, \ldots \). This completes the proof of Proposition. \( \square \)

We assume below that a number \( r \geq 1 \) is fixed, and the number \( \mu \) is defined in (6.3).
We use the notations \( q, \ Q \) for the constants \( q_r, \ Q_r \) defined in (6.7) with \( \rho = r \).

Lemma A.1 The following estimate holds:
\[
\text{Ind } \{ k \geq 1 \} \left( a^{r \mu} \mathbf{E} \{ a^{-r\mu} | X^r \} - 1 \right) \leq 0 \quad \text{a.s.}
\]

Proof. For \( j \geq 0 \), introduce random events
\[
\Omega^{(j)} \triangleq \{ \| z_{\tau_k+j} \| \leq l_{\tau_k+j-1}, \quad \| z_{\tau_k+j+1} \| > l_{\tau_k+j+1} \}.
\]
Substituting (5.5) into (A.4), we have that the event \( \Omega^{(j)} \) implies
\[
\| z_{\tau_k+j} \| \leq \frac{l_{\tau_k+j-1}}{\nu}, \quad \| A_{\tau_k+j+1} z_{\tau_k+j} + v_{\tau_k+j+1} \| \geq l_{\tau_k+j}.
\]
Hence
\[ \|A_{\tau_j + j} z_{\tau_j + j}\| \leq \frac{L \tau_j + 1}{\nu} \leq \frac{a \tau_j + j}{\nu}. \]

It follows that the event \( \Omega^{(j)} \) implies
\[ \|v_{\tau_j + j + 1}\| \geq l_{\tau_j + j} \left( 1 - \frac{La}{\nu} \right) = l_{\tau_j + j} h. \quad (A.6) \]

Furthermore,
\[ P(\eta_k = j | F') \leq P(\|z_{\tau_j + j + 1}\| > l_{\tau_j + j} \text{, } \| \bar{z}_{\tau_j + j}\| \leq l_{\tau_j + j - 1} | F') \text{, } j \geq 1. \]

Let \( \omega \triangleq G \mu \delta h^{-\mu} \). We have that
\[ P(\eta_k = j | F') \leq P(\|v_{\tau_j + j + 1}\| \geq l_{\tau_j + j} h | F') . \]

It follows from Proposition A.4 that
\[ P(\eta_k = j | F') \leq \frac{G \delta \mu}{l_{\tau_j + j} h \mu} = \frac{\omega}{l_{\mu}}. \]

By (A.1), we have that \( l_{\tau_j + j} = a^j l_{\tau_j} \). Moreover, from Proposition A.2 we have that \( l_{\tau_j} = a^\xi l a^{-1} \). Hence
\[ P(\eta_k = j | F') \leq \frac{\omega a^j}{l_{\mu} a^{\mu j - \mu \xi}}. \]

Then
\[ \sum_{j=1}^{\xi_k} a^{-r j} P(\eta_k = j | F') \leq \frac{\omega a^{-\mu \xi}}{l_{\mu}} \sum_{j=1}^{\xi_k} a^{j(\mu-r)-ja^{-1}} a^{-r j} \]
\[ \leq \frac{\omega a^{-\mu \xi}}{l_{\mu}} a^{-\mu \xi} a^{j(\mu-r)-ja^{-1}} a^{-r j} \leq \frac{\omega a^{-\mu \xi}}{l_{\mu} a^{\mu-r-1}} a^{-r j}. \]

Hence
\[ \text{Ind} \{ k \geq 1 \} \{ a^{-r \eta_k} | F' \} \]
\[ \leq \text{Ind} \{ k \geq 1 \} \left[ \sum_{j=1}^{\xi_k} a^{-r j} P(\eta_k = j | F') + a^{-r \xi_k - r} P(\eta_k \geq \xi_k + 1 | F') \right] \]
\[ \leq \text{Ind} \{ k \geq 1 \} \left( \frac{\omega a^{-r \xi_k}}{l_{\mu} a^{\mu-r-1}} + a^{-r \xi_k - r} \right) \leq \text{Ind} \{ k \geq 1 \} a^{-r \xi_k} \left( \frac{\omega a^{-1}}{l_{\mu} a^{\mu-r-1}} + a^{-r} \right) \text{ a.s.} \]

The equation (6.5) implies that
\[ \frac{\omega a^{-1}}{l_{\mu} a^{\mu-r-1}} + a^{-r} \leq 1. \]

This completes the proof of Lemma A.1. \( \square \)

Introduce the following matrices:
\[ \overline{A}(t, s) \triangleq A_{t} A_{t-1} \cdots A_{s} \text{ for } s \leq t, \quad \overline{A}(t-1, t) \triangleq I. \quad (A.7) \]

Here \( I \) is the unit matrix. Furthermore, introduce the following random vectors:
\[ \overline{z}_{t, s_m} \triangleq \overline{A}(t, s_m) z_{s_m - 1} + \sum_{j = s_m}^{t} \overline{A}(t, j+1) u_j, \text{ where } t = s_m + 1, s_m + 2, \ldots. \]
**Proposition A.5** For the conditional probabilistic measure $\mathbf{P}(\cdot | \mathcal{F}^n)$, the random vectors $\overline{z}_{t,m}$ are Gaussian, and the following inequalities hold:

$$\|\mathbb{E}\{\overline{z}_{t,m} | \mathcal{F}^n\}\| \leq L^{t-s_m+1} \frac{l_{s_m-1}}{\nu} \text{ a.s.,}$$  \hspace{1cm} (A.8)

$$\mathbf{E}\left\{\|\overline{z}_{t,m} - \mathbf{E}\{z_{t,m} | \mathcal{F}^n\}\|^2 | \mathcal{F}^n\right\} \leq C_0 L^{2(t-s_m+1)} \text{ a.s.,}$$  \hspace{1cm} (A.9)

where $C_0$ is defined by (6.6).

**Proof.** By Assumption 4.1, the vector $v_{t+1}$ does not depend on $x_t, x_{t-1}, \ldots$ for any deterministic $t$. Moreover, (5.2) implies that there exist deterministic measurable functions $\phi_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $\overline{x}_t = \phi_i(x_1, \ldots, x_t)$. We have that $z_t = x_t - \overline{x}_t$. Hence $v_{t+1}$ does not depend on $z_t, z_{t-1}, \ldots$. Furthermore, the vectors $v_{s_m+i}, i = 1, 2, \ldots$, are Gaussian and independent on $z_{s_m-1}$ for the conditional probabilistic measure $\mathbf{P}(\cdot | \mathcal{F}^n)$, and $\mathbf{E}\{z_{t,m} | \mathcal{F}^n\} = \overline{A}(t, s_m) z_{s_m-1}$.

Furthermore,

$$\|\overline{z}_{s_m-1}\| \leq \frac{l_{s_m-1}}{\nu}, \quad \|\overline{A}(t, j)\| \leq L^{t-j+1}.$$  

Hence (A.8) holds. Further, 

$$\mathbf{E}\left\{\|\overline{z}_{t,m} - \mathbf{E}\{z_{t,m} | \mathcal{F}^n\}\|^2 | \mathcal{F}^n\right\} = \mathbf{E}\left\{\left(\sum_{j=s_m}^{t} \overline{A}(t, j+1) v_j \right)^2 | \mathcal{F}^n\right\} \leq \sum_{j=s_m}^{t} \left(\|\overline{A}(t, j+1)\| \mathbf{E}\{\|v_j\|^2 | \mathcal{F}^n\}\right)^{1/2}\left(\|v_j\| \mathbf{E}\{\|v_j\|^2 | \mathcal{F}^n\}\right)^{1/2} \leq \delta^2 \left(\sum_{j=s_m}^{t} L^{t-j-i} \right)^2 \leq \delta^2 \left(\sum_{j=0}^{t-s_m} L^{t-j-s_m} \right)^2 \leq \delta^2 \left(\sum_{j=0}^{t-s_m} L^{-j} \right)^2 \leq \delta^2 \left(\frac{L^2(t-s_m+1)}{(L-1)^2}\right) = C_0 L^{2(t-s_m+1)}.$$  

Hence (A.9) holds. This completes the proof of Proposition.

**Proposition A.6** The following estimate holds:

$$\mathbf{P}(\xi_m \geq i + 1 | \mathcal{F}^n) \leq G_q \frac{L^{i_q} \beta^q C_0^{q/2}}{q^q \beta^q \sqrt{(\beta - L^{-1})^q}} \quad \forall i \geq 1, \quad \forall q = 2, 4, 6, \ldots$$

**Proof.** After substituting (5.5) into (A.4), we obtain that

$$\overline{z}_{s_m+i} = 0, \quad z_{s_m+i} = \overline{z}_{s_m+i} = \overline{z}_{s_m+i} \text{ for } i = 0, \ldots, \tau_m - s_m.$$  \hspace{1cm} (A.10)

Moreover, it can be easily seen, that

$$\mathbf{P}(\xi_m \geq i + 1 | \mathcal{F}^n) \leq \mathbf{P}(\|z_{s_m+i}\| > l_{s_m+i-1} | \mathcal{F}^n), \quad i \geq 1.$$  

Furthermore, if $s_m \leq i \leq \tau_m$ then

$$l_{s_m-1} \geq \frac{l_{s_m-1}}{\alpha}, \quad l_{s_m+i-1} \geq l_{s_m-1} \beta^i,$$

$$l_{s_m+i-1} - \frac{l_{s_m-1}}{\alpha} \geq l_{s_m-1} \left(\frac{\beta^i - \frac{l_{s_m-1}}{\alpha}}{\alpha}\right) \geq \frac{l_{s_m-1}(\beta^i)}{\alpha} \geq l_{s_m-1}(\beta^i - 2).$$
where
\[ \omega_1 \triangleq \beta - L. \]

By Propositions A.4-A.5 and (A.10), it follows that
\[
P(\xi_m \geq i + 1|\mathcal{F}_n) \leq \frac{G_q Li\sigma C_0 q^2}{(l_{i+1} - l_{i-1})/\nu} \leq \frac{G_q \beta \omega \sigma C_0 q^2}{\omega_1 q^2} \quad \forall q = 2, 4, 6, \ldots
\]

This completes the proof of Proposition A.6.

**Lemma A.2** The following estimate holds:
\[
E\left\{\beta^r \xi_m | \mathcal{F}_n\right\} \leq f(r) \quad a.s.,
\]
where \( f(\cdot) \) defined by (6.8).

**Proof of Lemma A.2.** We have that
\[
P(\xi_m \geq 1|\mathcal{F}_n) \leq 1
\]
and
\[
E\left\{\beta^r \xi_m | \mathcal{F}_n\right\} \leq \beta^r + \sum_{i=1}^{+\infty} \beta^{i+1} r P(\xi_m = i + 1|\mathcal{F}_n) \leq \beta^r + \frac{G_q \beta \omega \sigma C_0 q^2}{\nu \omega_1 q^2} \sum_{i=1}^{+\infty} Q_i,
\]
where \( Q = Q_r \). Then
\[
E\left\{\beta^r \xi_m | \mathcal{F}_n\right\} \leq \beta^r + \frac{G_q \beta \omega \sigma C_0 q^2}{\nu \omega_1 q^2} \frac{Q}{1 - Q},
\]
where \( q = q_r \). This completes the proof of Lemma A.2.

**Proof of Theorem 6.1-6.3.** Let \( m \) and \( k \) be the random numbers defined above for the integer \( T > 0 \). Consider the following random events:
\[
\Omega_0 \triangleq \{T < s_1\}, \quad \Omega_1 \triangleq \{s_1 \leq T \leq \tau_1\}, \quad \Omega_2 \triangleq \{\tau_1 < T < s_2\}, \quad \Omega_3 \triangleq \{s_2 \leq T\}.
\]

(A.11)

It is easy to see that
\[
P(\bigcup_{i=0}^{3} \Omega_i) = 1, \quad P(\Omega_i \cap \Omega_j) = 0, \quad \Omega_i \neq \Omega_j, \quad i, j = 0, 1, 2, 3,
\]
\[
\Omega_0 = \{m = 0\}, \quad \Omega_1 \cup \Omega_2 \cup \Omega_3 = \{m = 1\}, \quad \Omega_3 = \{m \geq 2\} = \{k \geq 1\}.
\]

We have that
\[
E\|z_T\|^r \leq E\{\|z_T\|^r|\Omega_0\} P(\Omega_0) + \sum_{i=1}^{3} E1\text{Ind} \{\Omega_i\}\|z_T\|^r.
\]

If \( T < s_1 \) then \( \|z_T\| \leq l/\nu \), hence
\[
E\{\|z_T\|^r|\Omega_0\} \leq \frac{l^r}{\nu^r}. \quad (A.12)
\]

Set
\[
V \triangleq \sum_{j=s_m}^{T} A(T, j + 1) v_j.
\]
If \( s_m \leq T \leq \tau_m \), then \( z_T = \overline{A}(T, s_m) z_{s_m} + \overline{V} \), where \( \overline{A}(T, j) \) is defined by (A.7). By (6.1),

\[
\| \overline{A}(T, s_m) \| \leq L^{T-s_m+1} \leq a^L \xi_m \delta \\
\| V \| \leq \sum_{j=s_m}^{T} L^{-j} \| v_j \| \leq L^{T-s_m+1} \sum_{j=s_m}^{T} L^{s_m-j-1} \| v_j \|. \tag{A.13}
\]

Consider the random event \( \Omega' \triangleq \{ s_m \leq T \leq \tau_m \} \). We have that

\[
E \text{Ind} \{ \Omega' \} \| V \|^{r} \leq E \left( \text{Ind} \{ \Omega' \} L^{T-s_m+1} \sum_{j=s_m}^{T} L^{s_m-j-1} \| v_j \| \right)^{r/p} \leq \left( E \text{Ind} \{ \Omega' \} L^{r \rho (T-s_m)} \right)^{1/p} \geq \left( \sum_{j=s_m}^{T} L^{s_m-j-1} (E \| v_j \|^{r})^{1/r} \right)^{1/p} \leq f(r \rho)^{1/p} \frac{\overline{A}(r \rho)}{r-1} \tag{A.14}
\]

Further, we have that \( \| z_0 \| \leq l/\nu \) for \( t \leq s_1 \) and \( \Omega_1 \subset \Omega' \). Clearly,

\[
\left( E \text{Ind} \{ \Omega_1 \} \| z_T \|^{r} \right)^{1/r} \leq \left( E \text{Ind} \{ \Omega_1 \} \| \overline{A}(T, s_1) z_{s_1} \|^{r} \right)^{1/r} + \left( E \text{Ind} \{ \Omega_1 \} \| V \|^{r} \right)^{1/r}.
\]

Then (A.13) and Lemma A.2 imply that

\[
\left( E \text{Ind} \{ \Omega_1 \} \| z_T \|^{r} \right)^{1/r} \leq \frac{l}{\nu} \left( E \text{Ind} \{ \Omega_1 \} L^{\xi_1} \right)^{1/r} + f(r \rho)^{1/r} \frac{\overline{A}(r \rho)}{r-1} \tag{A.15}
\]

If \( \tau_1 < T \leq \tau_2 \), then \( \| z_T \| \leq l_1 \leq l \beta^{\xi_1} \), hence

\[
E \text{Ind} \{ \Omega_2 \} \| z_T \|^{r} \leq l^{r} E \text{Ind} \{ \Omega_2 \} \beta^{\xi_1} = l^{r} E \text{Ind} \{ \Omega_2 \} \beta^{\xi_m} \leq l^{r} f(r). \tag{A.16}
\]

By the definitions, it follows that

\[
\| z_T \| \leq l a^{\xi_m-1} a^{-\eta_m-1} \delta \beta^{\xi_m} \text{ if } T > \tau_m, \quad m \geq 2,
\]

\[
\| z_T \| \leq \| \overline{A}(T, s_m) z_{s_m} + \overline{V} \| \text{ if } T \leq \tau_m, \quad m \geq 2.
\]

In other words,

\[
(1 - \text{Ind} \{ \Omega' \}) \text{Ind} \{ \Omega_3 \} \| z_T \| \leq (1 - \text{Ind} \{ \Omega' \}) \text{Ind} \{ \Omega_3 \} l a^{\xi_m-1} a^{-\eta_m-1} \beta^{\xi_m},
\]

\[
\text{Ind} \{ \Omega' \} \text{Ind} \{ \Omega_3 \} \| z_T \| \leq \text{Ind} \{ \Omega' \} \text{Ind} \{ \Omega_3 \} \left( \| \overline{A}(T, s_m) z_{s_m} + \overline{V} \| \right).
\]

Note that \( L < \beta = a^L \delta \), and if \( \text{Ind} \{ \Omega_3 \} \neq 0 \) then \( m \geq 2 \). From (A.13), we have that

\[
\text{Ind} \{ \Omega_3 \} \| z_{s_m} \| \leq \text{Ind} \{ \Omega_3 \} l a^{\xi_m-1} a^{-\eta_m-1},
\]

\[
\text{Ind} \{ \Omega_3 \} \| A(T, s_m) z_{s_m} \| \leq \text{Ind} \{ \Omega_3 \} l a^{\xi_m-1} a^{-\eta_m-1} \beta^{\xi_m}.
\]

It means that

\[
\text{Ind} \{ \Omega_3 \} \| z_T \| \leq \text{Ind} \{ \Omega_3 \} \left( l a^{\xi_m-1} a^{-\eta_m-1} \beta^{\xi_m} + \text{Ind} \{ \Omega' \} \| V \| \right).
\]
Hence
\[
\left( E \text{Ind} \{ \Omega_3 \} \| z_T \|^{r} \right)^{1/r} \leq E \left( \text{Ind} \{ \Omega_3 \} l^{r} a^{r \xi_m - 1} a^{-r \eta_m - 1} \beta^{r \xi_m} \right)^{1/r} + E \left( \text{Ind} \{ \Omega_3 \cap \Omega' \} \| V \|^{r} \right)^{1/r}.
\]
In other words,
\[
\left( E \text{Ind} \{ \Omega_3 \} \| z_T \|^{r} \right)^{1/r} \leq K_1^{1/r} + K_2^{1/r},
\]
(A.17)
where
\[
K_1 \triangleq l^{r} E \text{Ind} \{ \Omega_3 \} a^{r \xi_m - 1} a^{-r \eta_m - 1} \beta^{r \xi_m}, \quad K_2 \triangleq E \text{Ind} \{ \Omega_3 \cap \Omega' \} \| V \|^{r}.
\]
By (A.14), it follows that
\[
K_2 \leq E \text{Ind} \{ \Omega' \} \| V \|^{r} \leq f(rp)^{1/ \beta} \frac{\overline{m}(rp)^{r}}{(L - 1)^{r}}.
\]
(A.18)
By Proposition A.3, the random variable \( \text{Ind} \{ \Omega_3 \} = \text{Ind} \{ m \geq 2 \} \) is \( \mathcal{F} \)-measurable and \( \mathcal{F}'' \)-measurable; \( \text{Ind} \{ k \geq 1 \} = \text{Ind} \{ m \geq 2 \} \). By Lemma A.1 and Lemma A.2, it follows that
\[
K_1 \leq E \text{Ind} \{ \Omega_3 \} \| z_T \|^{r} = E E \left\{ l^{r} a^{r \xi_m - 1} a^{-r \eta_m - 1} \beta^{r \xi_m} \text{Ind} \{ \Omega_3 \} | \mathcal{F} \right\}
\]
\[
= E \left( l^{r} \text{Ind} \{ \Omega_3 \} a^{r \xi_m - 1} E \left\{ a^{-r \eta_m - 1} \beta^{r \xi_m} | \mathcal{F} \right\} \right)
\]
\[
= E \left( l^{r} \text{Ind} \{ m \geq 2 \} a^{r \xi_m - 1} E \left\{ a^{-r \eta_m - 1} | \mathcal{F} \right\} \right) \quad \text{(A.19)}
\]
\[
\leq f(r) l^{r} E \left\{ a^{r \xi_m - 1} \text{Ind} \{ k \geq 1 \} | a^{-r \eta_m - 1} \mathcal{F} \right\} \leq l^{r} f(r).
\]
By (A.12)-(A.19), it follows the inequality (6.2) for the constant (6.9). This completes the proof of Theorems 6.1–6.3. \( \square \)

Proof of Theorem 6.4. Let \( \overline{t} > 0 \) be such that (6.5) holds for \( l = \overline{t} \). Introduce random events \( \Omega_{0,t} \triangleq \{ t < s_1 \}, t = 1, \ldots, T \). Let \( \varepsilon_t \triangleq 1 - P(\Omega_{0,t}) \). We have that \( \| z_t \| \leq l/\nu \) for \( t \leq s_1 \). Let \( C_t = C_t(\nu) \) be the constant defined in Theorem 6.3 with \( r = \rho \). By Theorem 6.3, it follows that if \( l \geq \overline{t}, l/\nu \leq 1, \nu > L + 1 \) then
\[
E \| z_t \|^{r} \leq \frac{l^{r}}{\nu} P(\Omega_{0,t}) + E(1 - \text{Ind} \{ \Omega_{0,t} \}) \| z_t \|^{r}
\]
\[
\leq \frac{l^{r}}{\nu} P(\Omega_{0,t}) + (E \| z_t \|^{2r})^{1/2} (E(1 - \text{Ind} \{ \Omega_{0,t} \})^{2})^{1/2}
\]
\[
\leq \frac{l^{r}}{\nu} P(\Omega_{0,t}) + C_t(2r)^{1/2} (1 - P(\Omega_{0,t}))^{1/2} = \frac{l^{r}}{\nu} (1 - \varepsilon_t) + C_t(2r)^{1/2} \varepsilon_t^{1/2}.
\]
Let \( \kappa \triangleq \min\{i \in \{2,4,\ldots\} : i > 2r\} \). Similarly the proof of Lemma A.1, we obtain that
\[
\varepsilon_t \leq P \left( \exists t : \| z_{t+1} \| > l, \| z_t \| \leq l \right) \leq t \max_{i \leq \kappa} P(\| u_i \| > lh) \leq \frac{G^{\kappa} \delta}{l \kappa l^h}.
\]
By (6.9), it follows that there exists a constant \( c_1 > 0 \) such that \( C_t(\nu) \leq c_1(l^{\nu} + 1) \) for all \( l \geq \overline{t}, \nu \geq \max(L + 1, \overline{t}), \rho \in [r, 2r] \). Hence
\[
\varepsilon_t C_t(2r) \leq \frac{G^{\kappa} \delta}{l \kappa l^h} c_1(l^{2r} + 1) \to 0 \quad \text{as} \quad M \to +\infty, \quad l \to +\infty, \quad \frac{L}{\nu} \to 0.
\]
This completes the proof of Theorem 6.4.