Recursive Estimation of ARX Systems Using Binary Sensors with Adjustable Thresholds

Balázs Csanád Csáji [1,2] **Erik Weyer** [1]

(1) Department of Electrical and Electronic Engineering, The University of Melbourne

(2) Computer and Automation Research Institute, Hungarian Academy of Sciences

16th IFAC Symposium on System Identification, Brussels, July 11–13, 2012

Outline

- Problem: identifying an ARX systems via binary sensors
- Previous solutions typically assumed fully known noise characteristics
- They also assumed that the input signal can be chosen by the user
- We try to reduce the assumptions on the noise and the input
- Full knowledge on the distribution is not needed; the input is only observed
- But, the threshold of the binary sensor can be controlled $\sim$ dither signal
- Here, two recursive identification algorithms are proposed
- Algorithm I: FIR approximation; it is proved to be strongly consistent
- Algorithm II: simultaneous state and parameter estimation (simulations)

Structural Overview

PART I. **Problem Setting**
(ARX System via Binary Sensors, Dithering, Assumptions)

PART II. **General Form of the Algorithms**
(Sign-Error, Step-Sizes, Expanding Truncation Bounds)

PART III. **Recursive Identification: Algorithms I and II**
(FIR Approximation, Strong Consistency, Simultaneous Estimation)

PART IV. **Experimental Results**
(Simulation: Algorithms I and II on an ARX(2,2) System)

PART V. **Summary and Concluding Remarks**
(Main Ideas, Contributions and Highlights)

Problem Setting

- We observe an ARX system via a binary sensor:

$$X_t \triangleq \sum_{i=1}^{p} a_i^* X_{t-i} + \sum_{i=1}^{q} b_i^* U_{t-i} + N_t,$$
$$Y_t \triangleq \mathbb{I}(X_t \leq C_t),$$

where X_t — output (hidden state), U_t — input, N_t — noise (at time t)

- The thresholds of the binary sensor, $(C_t)_t$, can be controlled at each t
- Data: the inputs $(U_t)_t$ and the binary outputs $(Y_t)_t$ are observed
- Aim: to identify (estimate) $\theta^* = (a_1^*, \dots, a_p^*, b_1^*, \dots, b_q^*) \in \mathbb{R}^{p+q}$

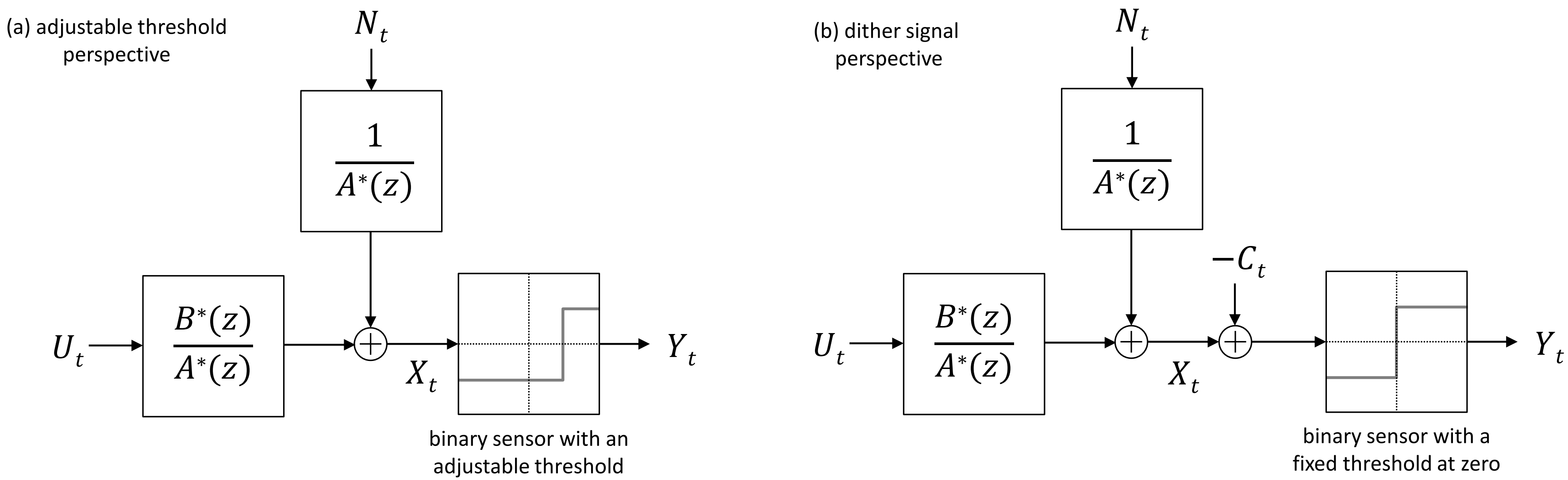
Adjustable Thresholds $\sim$ Dithering

- The binary output can be rewritten as

$$Y_t = \mathbb{I}(\varphi_t^{\mathrm{T}}\theta^* + N_t \leq C_t) = \mathbb{I}(\varphi_t^{\mathrm{T}}\theta^* + N_t - C_t \leq 0),$$

where $\varphi_t = (X_{t-1}, \dots, X_{t-p}, U_{t-1}, \dots, U_{t-q})$ — random regressor

- Choosing the threshold is equivalent to dithering



System Assumptions

- $(N_t)_t$ is i.i.d., continuous, zero mean, zero median, has a finite variance: $\sigma_n^2 \triangleq \mathbb{E}\left[\, N_t^2 \,\right] < \infty$, and has a continuous and positive density at zero
- $(U_t)_t$ is i.i.d., zero mean, $(U_t)_t$ and $(N_t)_t$ are independent, and $0 < \sigma_u^2 < \infty$, where $\sigma_u^2 \triangleq \mathbb{E}\left[\, U_t^2 \,\right]$
- The system is stable, i.e., the roots of $A^*(z)$ lie strictly inside the unit circle; additionally, the transfer function $B^*(z)/A^*(z)$ is irreducible,

$$\begin{aligned} A^*(z) &\triangleq 1 - a_1^* z^{-1} - a_2^* z^{-2} - \cdots - a_p^* z^{-p}, \\ B^*(z) &\triangleq b_1^* z^{-1} + b_2^* z^{-2} + \cdots + b_q^* z^{-q}, \end{aligned}$$

  where z^{-1} is the backward shift operator, $z^{-i} x_t \triangleq x_{t-i}$.
- The orders p and q are known

General Form of the Algorithms

- The general form of both proposed algorithms is

$$\hat{\theta}_{t+1} \;=\; \Pi_{M_{\mu(t)}}\Big[\, \hat{\theta}_t + \alpha_t\, \widehat{\varphi}_t \Big(1 - 2\, \mathbb{I}(X_t \leq \widehat{\varphi}_t^{\mathrm{T}} \hat{\theta}_t) \Big) \,\Big],$$

  where $\widehat{\varphi}_t$ is a regression vector defined differently in the two algorithms, $(\alpha_t)_t$ is a sequence of step-sizes and $\Pi_{M_{\mu(t)}}$ is a sequence of projections

- Assuming that N_t is continuous, we ($\mathbb{P}$-a.s.) have

$$\text{sign}(X_t - \widehat{\varphi}_t^{\mathrm{T}} \hat{\theta}_t) \;=\; 1 - 2\, \mathbb{I}(X_t \leq \widehat{\varphi}_t^{\mathrm{T}} \hat{\theta}_t),$$

- Thus, the above algorithm will behave almost surely as

$$\hat{\theta}_{t+1} \;=\; \Pi_{M_{\mu(t)}}\Big[\, \hat{\theta}_t + \alpha_t\, \widehat{\varphi}_t\, \text{sign}(X_t - \widehat{\varphi}_t^{\mathrm{T}} \hat{\theta}_t) \,\Big],$$

  which is a sign-error type algorithm with expanding truncation bounds

Step-Sizes

- Typical step-size assumption of stochastic approximation algorithms

$$\sum_{t=0}^{\infty} \alpha_t = \infty,$$

$$\sum_{t=0}^{\infty} \alpha_t^2 < \infty,$$

$$\forall\, t \geq 0 : \alpha_t \geq 0.$$

  The second condition can often be weakened to $\lim_{t\to\infty} \alpha_t = 0$

- Here, we will simply assume that

$$\alpha_0 = 1 \qquad \text{and} \qquad \forall\, t > 0 : \alpha_t = 1/t.$$

Expanding Truncation Bounds

- Let $(M_t)_t$ be a sequence of (strictly) monotone increasing positive real numbers with $M_t \to \infty$ as $t \to \infty$,
- Let $\mathbb{I}(\cdot)$ be the indicator function and define $\mu(t)$ and $\Delta\hat{\theta}_i$ as

$$\mu(t) \triangleq \sum_{i=1}^{t-1} \mathbb{I}\big(|\hat{\theta}_i + \Delta\hat{\theta}_i| > M_{\mu(i)}\big),$$

$$\Delta\hat{\theta}_i \triangleq \alpha_i\, \widehat{\varphi}_i \big(1 - 2\,\mathbb{I}(X_i \leq \widehat{\varphi}_i^{\mathrm{T}} \hat{\theta}_i)\big).$$

- Given a positive real M, projection Π_M is

$$\Pi_M(x) \triangleq \begin{cases} x & \text{if } \|x\| \leq M, \\ 0 & \text{otherwise.} \end{cases}$$

Algorithm I: FIR Approximation

- Using impulse responses, $(c_i^*)_{i=1}^{\infty}$ and $(d_i^*)_{i=0}^{\infty}$, we have

$$X_t \;=\; \sum_{i=1}^{\infty} c_i^* U_{t-1} + \sum_{i=0}^{\infty} d_i^* N_{t-i},$$

- Let's approximate our ARX system with an FIR system of order $p+q$

$$\boxed{X_t \;=\; \bar{\varphi}_t^{\mathrm{T}} \bar{\theta}^* + W_t,}$$

$$\bar{\varphi}_t \;\triangleq\; (U_{t-1}, \ldots, U_{t-p-q})^{\mathrm{T}}, \qquad \bar{\theta}^* \;\triangleq\; (c_1^*, \ldots, c_{p+q}^*)^{\mathrm{T}}.$$

- W_t is simply the unmodelled part of the system

$$W_t \;\triangleq\; \sum_{i=p+q+1}^{\infty} c_i^* U_{t-i} + \sum_{i=0}^{\infty} d_i^* N_{t-i}.$$

Algorithm I: FIR Approximation

- If we can estimate $\bar{\theta}^*$, we can also estimate the true parameter vector θ^*
- There is a function f, which we use for post processing, such that

$$\theta^* \; = \; f(\bar{\theta}^*),$$

- Algorithm I is defined by using $\widehat{\varphi}_t \triangleq \bar{\varphi}_t$ in the General Algorithm

Theorem 1 (Strong Consistency of Algorithm I)**.** *Let $(\hat{\theta}_t)_{t=0}^{\infty}$ be the sequence generated by Algorithm I (i.e. $\widehat{\varphi}_t = \bar{\varphi}_t$). Then, under the given assumptions, $f(\hat{\theta}_t)$ converges $(\mathbb{P}$-a.s.$)$ to θ^*, as $t \to \infty$, for any $\hat{\theta}_0 \in \mathbb{R}^{p+q}$.*

- Furthermore, $\sqrt{t}(\hat{\theta}_t - \bar{\theta}^*)$ is approximately normal

Algorithm II: Simultaneous Estimation

- Main idea: to achieve a direct estimate of θ^* by simultaneously maintaining an estimate for the output, $\widehat{X}_t$ and for the parameter, $\hat{\theta}_t$, at time t.
- The sequence of output estimates is defined as

$$\widehat{X}_t \triangleq \begin{cases} \sum_{i=1}^{p} \hat{a}_{t,i} \widehat{X}_{t-1} + \sum_{i=1}^{q} \hat{b}_{t,i} U_{t-i} & \text{if } t \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $(\hat{a}_{t,i})_{i=1}^{p}$ and $(\hat{b}_{t,i})_{i=1}^{q}$ are the estimates of the true parameters.

- Algorith II: is defined by setting the General Algorithm as

$$\begin{aligned} \widehat{\varphi}_t &\triangleq (\widehat{X}_{t-1}, \dots, \widehat{X}_{t-p}, U_{t-1}, \dots, U_{t-q})^{\mathrm{T}}, \\ \hat{\theta}_t &\triangleq (\hat{a}_{t,1}, \dots, \hat{a}_{t,p}, \hat{b}_{t,1}, \dots, \hat{b}_{t,q})^{\mathrm{T}}. \end{aligned}$$

Simulation Experiment: ARX(2, 2)

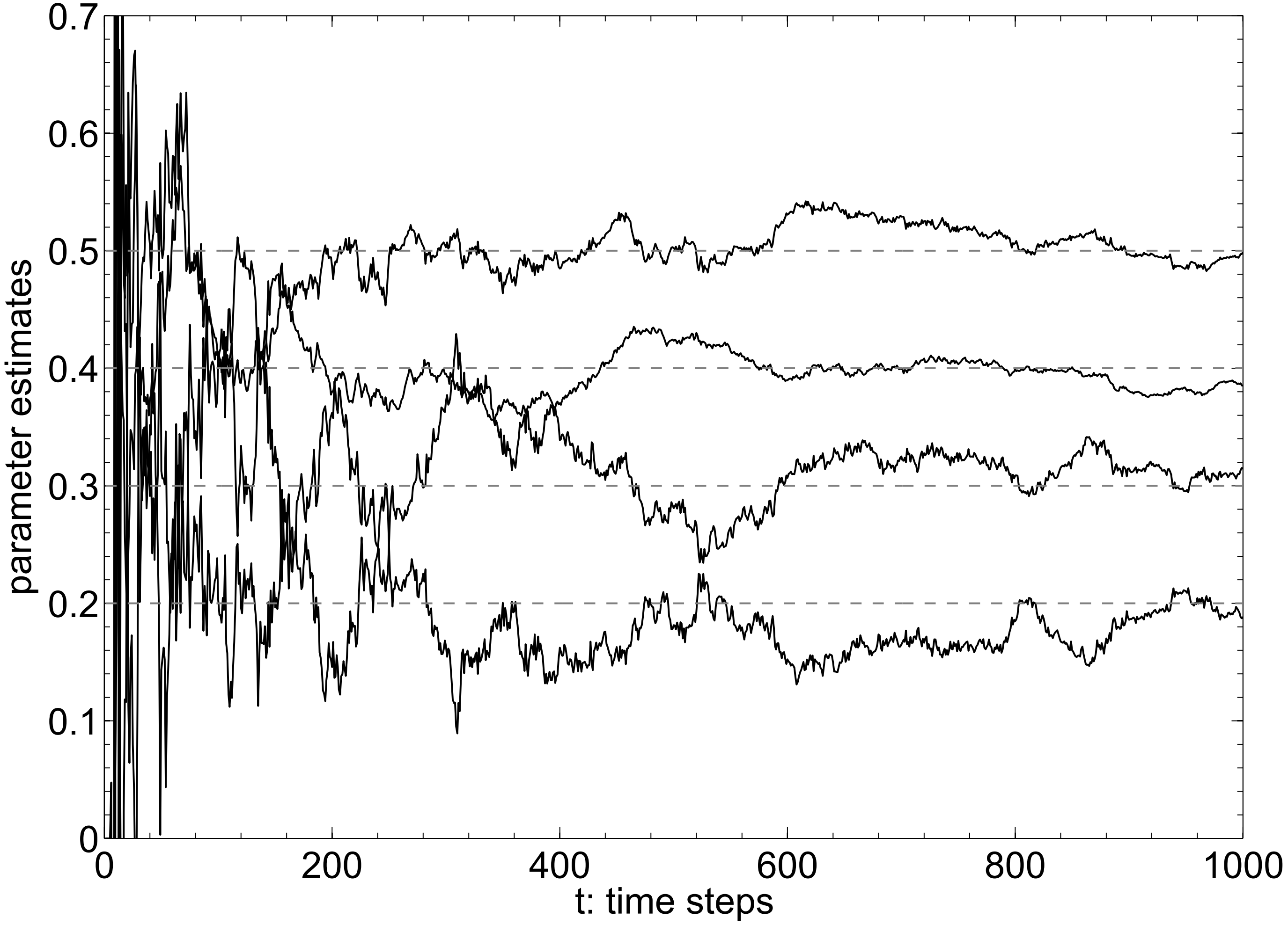


Figure 1: Recursive estimation with Algorithm I

Simulation Experiment: ARX(2, 2)

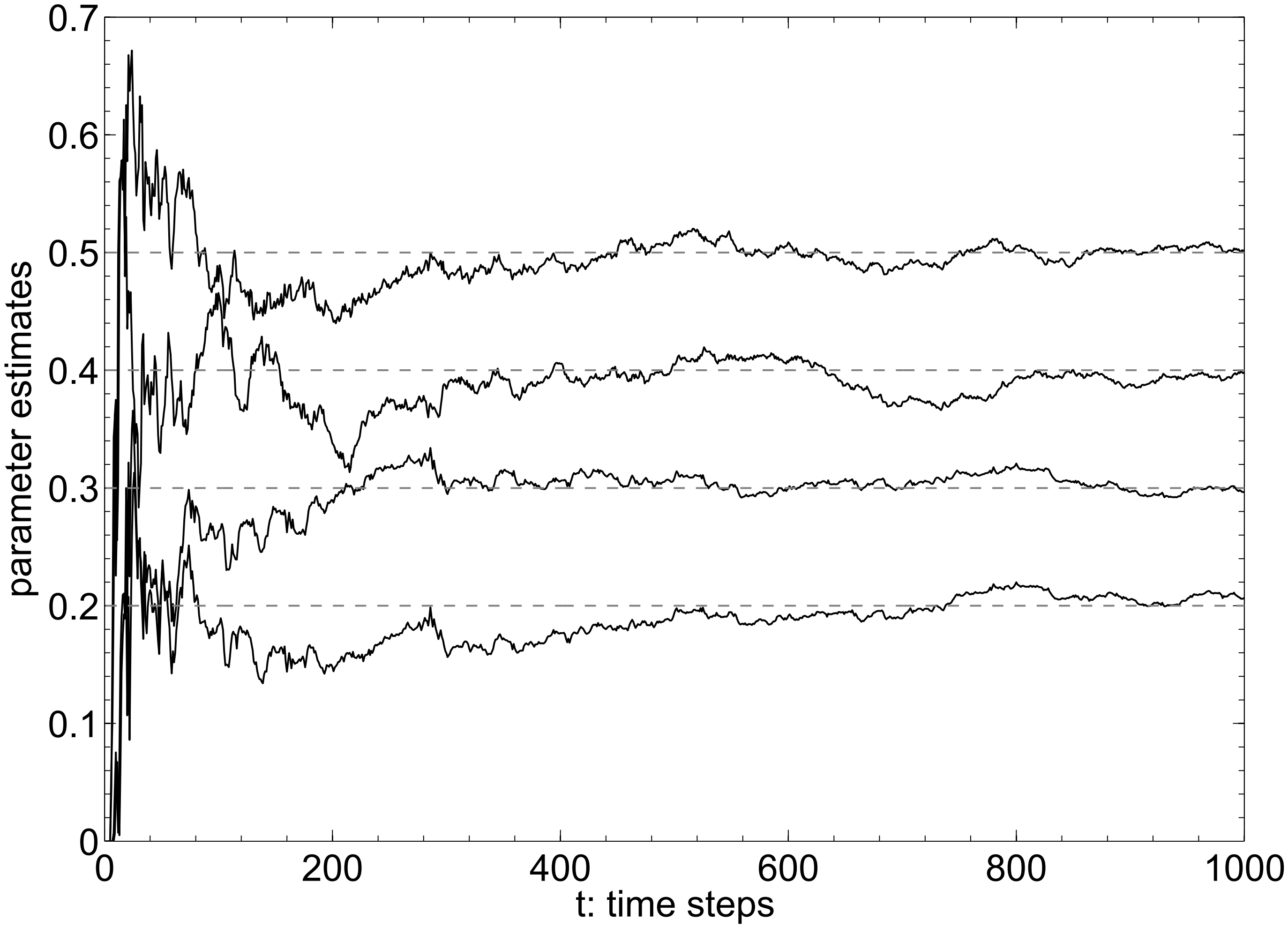


Figure 2: Recursive estimation with Algorithm II

Summary and Concluding Remarks

- Two recursive identification algorithms have been proposed for identifying ARX systems via binary sensors
- These algorithms neither assume the knowledge of the particular noise distributions, nor assume that the input signal can be chosen by the user
- But, they do assume that the threshold of the sensor can be controlled
- This is assumption is equivalent to allowing a dither signal
- Algorithm I: FIR approximation; it was proved to be strongly consistent
- Algorithm II: simultaneous state and parameter estimation (no theorem)
- Experimental results demonstrated that both algorithms efficiently approximated the parameters of an ARX(2,2) system

Thank you for your attention!

bcsaji@unimelb.edu.au