

System Identification with Binary Observations by Stochastic Approximation and Active Learning

Balázs Csanád Csáji^{†‡} and Erik Weyer[†]

Abstract—We investigate the problem of estimating a constant based on noisy observations via a *binary* sensor. This problem is well-studied for the case when the noise characteristics are *known*, for example, the noise is i.i.d. and we have access to its cumulative distribution function (CDF). Here, we try to reduce the assumptions on the noise to a minimum and, for example, assume only that the noise is symmetrically distributed about zero in each time step, but otherwise the CDF is *unknown*. We neither assume that the noise variables are independent nor that they are stationary. They may also not have densities. We do assume, however, that the *threshold* of the binary sensor *can be controlled*. Based on the setting that the threshold can be set to any value or only to some predefined ones, we suggest solutions based on *stochastic approximation* (SA) and *active learning* (AL). In the former case, we provide a *strongly consistent* estimator, while in the latter case we give a *probably approximately correct* (PAC) algorithm. Finally, we present numerical experiments to support the results.

I. INTRODUCTION

Estimating a parameter of a system based on *noisy* observations is one of the basic problems of statistics and system identification [7], [9]. In some situations, however, the output of the system is only available through *quantization* due to, for example, limited sensor capacities or signal coding. Quantized variants of system identification problems are well-studied [1], [2], [5], [14], but, with some exceptions [4], [15], only for the case when the characteristics of the noise are *known*, for example, the noise is i.i.d., its cumulative distribution function (CDF) is invertible and *known* to the user and its density function is continuously differentiable. However, making strong assumptions on the noise is undesirable as the noise characteristics are often unknown.

In this paper we try to reduce the assumed knowledge on the noise to a minimum and assume basically that we know the CDFs of the noise terms in just *one point*. For example, we may only know that the noise is always symmetrically distributed about zero. We do not assume knowledge of the particular noise distributions, nor that the noise variables are independent or uncorrelated. They may also not have densities. Furthermore, the noise terms may be nonstationary and their conditional CDFs may change in each time step. However, we make the assumption that *the threshold of the binary sensor is under our control*, either totally or partially.

[†]Department of Electrical and Electronic Engineering, Melbourne School of Engineering, The University of Melbourne, 240 Grattan Street, Parkville, Melbourne, Victoria, 3010, Australia, email: bcsaji@unimelb.edu.au and ewey@unimelb.edu.au

[‡]Computer and Automation Research Institute (SZTAKI), Hungarian Academy of Sciences, Kende utca 13–17, Budapest, Hungary, H-1111

In the first case, we have total control over the threshold: we can set it to any possible value. We will show that we can formulate this binary identification problem as a *stochastic root finding* problem [6], [12] and by applying *stochastic approximation* (SA) algorithms we can obtain a *strongly consistent* estimator, i.e., the parameter estimates converge to the true parameter value with probability one.

We also present a convergence result for the case when the thresholds are restricted to an interval and the estimates are projected back to this interval after each iteration.

Finally, we investigate the case when we have only partial control over the threshold: we can only set it to some finitely many predefined values. We will show that, if the noise is i.i.d. and symmetric about zero, this problem can be cast into the framework of *multi-armed bandit* systems [10] and by applying an *active learning* (AL) algorithm we can obtain a *probably approximately correct* (PAC) solution.

II. PRELIMINARIES

In this section we introduce the basic problem setting and briefly present a standard solution method. During this section we will assume, for simplicity, that the noise is i.i.d.

A. Problem Setting

Let us consider the problem of identifying a constant in noise based on quantized, *binary* observations,

$$X_t \triangleq \theta^* + N_t, \quad (1)$$

$$Y_t \triangleq \mathbb{I}(X_t \leq c), \quad (2)$$

where $\theta^* \in \mathbb{R}$ is a constant, $(N_t)_{t \in T}$, where $T = \{0, 1, \dots\}$, is an i.i.d. noise sequence and \mathbb{I} denotes the indicator function, namely, its value is 1 if its argument is true and 0 otherwise. The process $(X_t)_{t \in T}$ is observed via a *binary* sensor with threshold c . The user has access to the observations $(Y_t)_{t \in T}$ and threshold c , which is assumed to be fixed in this first setting. The aim is to estimate θ^* .

B. Standard Solution

It is known that if the cumulative distribution function, F , of the noise, $(N_t)_t$, is known and its density function $f(x) \neq 0$, for all $x \in \mathbb{R}$, as well as it is continuously differentiable, the constant θ^* can be estimated as follows [14].

First, if $p \triangleq \mathbb{P}(Y_t = 1) = F(c - \theta^*)$ and F is invertible, which follows from the assumptions, then θ^* satisfies

$$\theta^* = c - F^{-1}(p). \quad (3)$$

Therefore, if \hat{P} is an estimate of p , then a natural estimate of θ^* based on \hat{P} is $\hat{\theta} \triangleq c - F^{-1}(\hat{P})$. Note that an estimate

of p can be obtained by simply taking the sample average of the observations up to time $t - 1$, Y_0, \dots, Y_{t-1} , that is

$$\widehat{P}_t \triangleq \frac{1}{t} \sum_{i=0}^{t-1} Y_i. \quad (4)$$

Since \widehat{P}_t converges w.p.1 to the true probability p by the *strong law of large numbers*, we also have that

$$\widehat{\theta}_t \triangleq c - F^{-1}(\widehat{P}_t) \quad (5)$$

converges almost surely to θ^* . Therefore, in this setting we can obtain a *strongly consistent* estimator sequence.

Additionally, it can be shown that $\widehat{\theta}_t$ also converges to the parameter θ^* in the *mean-square sense* [14], that is

$$\lim_{t \rightarrow \infty} \mathbb{E} \left[\sqrt{t}(\widehat{\theta}_t - \theta^*) \right]^2 = \frac{F(c - \theta^*)(1 - F(c - \theta^*))}{f^2(c - \theta^*)}, \quad (6)$$

moreover, it can be proved, as well, that the term

$$\frac{F(c - \theta^*)(1 - F(c - \theta^*))}{t f^2(c - \theta^*)}, \quad (7)$$

is the *Cramér-Rao lower bound*, consequently, this method is *asymptotically efficient*. Though, the results above are proved for the case of an i.i.d. noise process, they could be slightly generalized to the case of dependent noise, as long as the dependence is asymptotically diminishing [14].

This basic setting is fundamental for system identification with quantized observations, since more complicated settings, such as FIR systems, rational systems, Wiener and Hammerstein systems, can often be reduced to a finite set of identification problems of this type. Moreover, strong consistency and asymptotic efficiency of identification of a constant carry over to the more complicated settings [14].

C. Adjustable Threshold

A practical disadvantage of the above described approach is that it makes *strong assumptions on the noise*, for example, it requires the complete knowledge of the CDF which must be invertible. In practice, however, we often only have very limited knowledge of the noise. Here we investigate the problem when the assumptions on the noise are considerably reduced, for example, we may only know that it is always symmetrically distributed about zero, but we can set the threshold of the binary sensor for every observation,

$$Y_t \triangleq \mathbb{I}(X_t \leq \theta_t), \quad (8)$$

where the threshold θ_t is under the control of the user, and it can depend on all previous observations, Y_0, \dots, Y_{t-1} , past threshold values, $\theta_0, \dots, \theta_{t-1}$ and extra randomization.

We will analyze two versions of the problem. First, we will assume that the threshold of the sensor can be set to arbitrary values and give a strongly consistent estimator based on stochastic approximation. Then, we will assume that θ_t can only take one of finitely many pre-defined values and apply a multi-armed bandit formulation with an active learning method to get a probably approximately correct solution.

III. STOCHASTIC APPROXIMATION

First, we assume that we can set the threshold θ_t to *any* value and measure the corresponding binary output Y_t .

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying *probability space*, where Ω is the sample space, \mathcal{F} is the σ -algebra of events and \mathbb{P} is the measure. We denote the σ -algebra generated by the past of the process up to time $t \in T$ by $\mathcal{F}_t \subseteq \mathcal{F}$. More precisely,

$$\mathcal{F}_t \triangleq \sigma \{N_0, \dots, N_t, \theta_0, \dots, \theta_t\}, \quad (9)$$

where $(\mathcal{F}_t)_{t \in T}$ is a *filtration*, viz., a nondecreasing sequence of σ -fields. Note that random variables $(Y_k)_{k \leq t}$ are \mathcal{F}_t measurable and we only need to include $(\theta_t)_{t > 0}$ in (9) if want to apply some additional randomization for the thresholds.

The *conditional* distribution function of the noise term at time $t \in T$ is denoted by $F_t(\cdot)$, which is therefore

$$F_t(x) \triangleq \mathbb{P}(N_t \leq x | \mathcal{F}_{t-1}). \quad (10)$$

Note that, in our case, $F_t(x)$ is a *random variable*, since we applied *conditional probability* w.r.t. a σ -algebra.

A. Basic Assumptions

Initially, we will apply the following two assumptions:

Assumption 1: $\forall t \in T : F_t(0) = 1/2$ (\mathbb{P} -a.s.).

Assumption 2: $\exists \delta : \mathbb{R} \rightarrow \mathbb{R} : \forall x \neq 0 : \delta(x) \cdot x > 0$ and $\delta(0) = 0$, such that we have (\mathbb{P} -a.s.)

$$\begin{aligned} \forall t \in T : \forall x > 0 : F_t(x) - \delta(x) &\geq 1/2, \\ \forall t \in T : \forall x < 0 : F_t(x) - \delta(x) &\leq 1/2. \end{aligned}$$

We can always choose δ to be *monotone nondecreasing*, and henceforth we will assume, w.l.o.g., that this holds.

Assumption 1 is very mild, since it only requires that the (conditional) *median* of each noise term is zero, which is, e.g., the case if they are symmetric about zero. Many standard distributions (with zero mean), e.g., Gaussian, Laplace, Cauchy-Lorentz, Bernoulli, Binomial, Student's t, uniform and even several multimodal distributions have this property.

Assumption 2 guarantees that for all time t , zero is almost surely the only x that satisfies $F_t(x) = 1/2$. The introduction of δ rules out, for example, that a subsequence of $(F_t(x))_{t \in T}$, converges (with nonzero probability) to $1/2$ for an $x \neq 0$.

Note that every sequence of i.i.d. random variables with continuous and symmetric distribution satisfies both of these assumptions. Moreover, the assumptions are also satisfied for a sequence of independent random variables which have zero medians and there is a neighbourhood of the origin on which every variable has a density uniformly bounded away from zero. Furthermore, in general, independence is not necessary.

B. Unconstrained Algorithm

We can observe that Y_t can be interpreted as the noisy observation of $F_t(\theta_t - \theta^*)$, the CDF of $\theta_t - \theta^*$ at time t ,

$$Y_t = \mathbb{E} [Y_t | \mathcal{F}_{t-1}] + W_t = \quad (11)$$

$$= \mathbb{P}(\theta^* + N_t \leq \theta_t | \mathcal{F}_{t-1}) + W_t = \quad (12)$$

$$= F_t(\theta_t - \theta^*) + W_t, \quad (13)$$

where $W_t = Y_t - F_t(\theta_t - \theta^*)$ is a zero-mean, bounded noise with $|W_t| \leq 1$. Now, let us define function g_t (for all t) by

$$g_t(x) \triangleq F_t(x - \theta^*) - 1/2, \quad (14)$$

which reformulates our estimation problem as a root finding problem, since we know that for all $t \in T$, $g_t(\theta^*) = 0$.

Let us define a new function $g(x) \triangleq \delta(x - \theta^*)$, where δ is the function introduced in Assumption 2. Then, g is obviously bounded and its only root is also located at θ^* .

Since the function g is monotone nondecreasing, if we could measure its values, which corresponds to the case without noise, the root could be found by simple *bisection*. Alternatively, if g was known and continuously differentiable, the problem could also be solved by *Newton's method*. However, in our case, we do not have access to functions g and $(g_t)_{t \in T}$, only noisy observations are available,

$$G_t(\theta_t) \triangleq g_t(\theta_t) + W_t = Y_t - 1/2, \quad (15)$$

where the noise term, W_t , is the same as before.

Nevertheless, the above problem can be interpreted as *stochastic root finding* and can be solved by a *Robbins-Monro type* stochastic approximation algorithm, which is a Newton-like iterative method. In our case, we can define a stochastic approximation scheme for setting the thresholds, as

$$\theta_{t+1} \triangleq \theta_t - \alpha_t G_t(\theta_t) = \theta_t + \alpha_t [1/2 - Y_t], \quad (16)$$

where $\theta_0 \in \mathbb{R}$ is an arbitrary constant or, if it is random, we assume $\mathbb{E}[|\theta_0|] < \infty$, and α_t are (possibly randomized) step-size parameters satisfying the assumptions below (\mathbb{P} -a.s.):

Assumption 3: $\sum_{t=0}^{\infty} \alpha_t = \infty$

Assumption 4: $\sum_{t=0}^{\infty} \alpha_t^2 < \infty$

Assumption 5: $\forall t \geq 0 : \alpha_t \geq 0$

Assumption 4 could be weakened [8] to $\lim_{t \rightarrow \infty} \alpha_t = 0$ and an archetypical choice of step-sizes is $\alpha_t = 1/t$.

The convergence of iteration (16) would easily follow if the noise terms, $(W_t)_{t \in T}$, in (13) were mutually independent. This is not the case, since, for example, threshold θ_t depends on the past observations $(Y_i)_{i < t}$ and on the initial θ_0 .

However, they still have a *martingale difference* property, which is enough for proving the almost sure convergence of the algorithm [8]. To see this, note that (16) can be rewritten

$$\theta_{t+1} = \theta_t - \alpha_t g_t(\theta_t) + \alpha_t [g_t(\theta_t) + 1/2 - Y_t], \quad (17)$$

where the “deterministic” part, $\theta_t - \alpha_t g_t(\theta_t)$, mimics Newton’s recursion more closely, while the “stochastic” part, $\alpha_t [g_t(\theta_t) + 1/2 - Y_t]$, introduces noise to the iteration [11]. If we could show that the stochastic part (the “noise”) is a martingale difference sequence with bounded variance, then they would “average to zero” and therefore have no effect on the asymptotic behavior of the algorithm [8].

Lemma 1: *The terms $(W_t)_{t \in T}$ form a martingale difference sequence with conditional variances bounded by 1.*

Proof. First, we will compute their conditional expectation

$$\mathbb{E}[W_t | \mathcal{F}_{t-1}] = \mathbb{E}[Y_t - 1/2 - g_t(\theta_t) | \mathcal{F}_{t-1}] = 0, \quad (18)$$

This holds indeed, since, as we have seen in (13), Y_t can be treated as a noisy observation of $F_t(\theta_t - \theta^*)$ and, hence, $\mathbb{E}[Y_t | \mathcal{F}_{t-1}] = \mathbb{E}[g_t(\theta_t) | \mathcal{F}_{t-1}] + 1/2$ that, using the linearity of conditional expectation, immediately proves (18). Thus, the noise on g_t is a martingale difference sequence.

It is also easy to see that, because of the quantization, the new noise terms have bounded variances, more precisely

$$\mathbb{E}[W_t^2 | \mathcal{F}_{t-1}] = \mathbb{E}[(g_t(\theta_t) + 1/2 - Y_t)^2 | \mathcal{F}_{t-1}] \leq 1, \quad (19)$$

since, we have that $|1/2 - Y_t| = 1/2$ and $|g_t(\theta_t)| \leq 1/2$. \square

Note that the original noise process $(N_t)_{t \in T}$ may not be a martingale difference sequence (MDS), moreover, being an MDS is *neither sufficient nor necessary* to satisfy Assumptions 1 and 2. As a special case, a continuous MDS which is also a *mediangale*, i.e., its terms have zero conditional medians, and have densities that are uniformly bounded away from 0 on a neighbourhood of 0 satisfies the assumptions.

The main result of this section can be stated as

Theorem 1: *Under Assumptions 1-5, the sequence generated by iteration (16) converges almost surely to the true parameter θ^* from any initial parameter estimate $\theta_0 \in \mathbb{R}$.*

Note that if g_t is time-invariant, namely, $(N_t)_{t=0}^{\infty}$ is i.i.d., the almost sure convergence of iteration (16) to θ^* follows from the Robbins-Monro theorem [12], [8] given below.

Theorem 2: *Under Assumptions 3-5 and the ones below,*

$$\theta_{t+1} = \theta_t - \alpha_k H_t(\theta_t) \quad (20)$$

converges with probability one to constant θ^ from any θ_0 .*

C1: $H_t(\theta_t) = h(\theta_t) + D_t(\theta_t)$, $h(\cdot)$ is deterministic

C2: $h(\theta^*) = 0$, $\forall x \neq \theta^* : h(x)(x - \theta^*) > 0$,

C3: $\exists c : \forall x : |h(x)| \leq c(|x - \theta^*| + 1)$,

C4: $\forall \delta_1, \delta_2$ satisfying $0 < \delta_1 < \delta_2 < \infty$, we have

$$\inf_{\delta_1 \leq |x - \theta^*| \leq \delta_2} |h(x)| > 0, \quad (21)$$

C5: $\mathbb{E}[D_{t+1} | \mathcal{F}_t] = 0$, $\mathbb{E}[D_{t+1}^2 | \mathcal{F}_t] \leq \sigma^2 < \infty$

Theorem 2 can be applied to prove Theorem 1, in case $(N_t)_{t=0}^{\infty}$ is i.i.d. and hence $\forall t \in T : g_t = g$, as follows.

The step-size conditions of Theorem 2 are the same as of Theorem 1. Condition C1 holds because of the definition of $G_t(\cdot)$ and our additional assumption that $g_t = g$, for all t . Condition C2 is satisfied because of Assumptions 1 and 2. The growth-rate condition, C3, is trivially satisfied, since g is bounded. Condition C4 holds because g is monotone nondecreasing and θ^* is its only root. Finally, we know from Lemma 1 that the noise sequence $(W_t)_{t \in T}$ is an MDS and all of their conditional variances are bounded by 1, which assure condition C5. Thus, $(\theta_t)_{t \in T}$ converges (\mathbb{P} -a.s.) to θ^* .

In the general case, however, each function g_t is *random* and *may change* at every time t , consequently, Theorem 2 is no longer applicable. In this case, Theorem 1 can be proved by the Robbins-Siegmund theorem, a.k.a. the *nonnegative almost supermartingale* convergence theorem [7].

Theorem 3: *Let $(V_t)_{t=0}^{\infty}$ be a sequence of nonnegative random variables adapted to a filtration $(\mathcal{F}_t)_{t=0}^{\infty}$. If (a.s.)*

$$\mathbb{E}[V_{t+1} | \mathcal{F}_t] \leq (1 + \mu_t) V_t - \lambda_t + \gamma_t, \quad (22)$$

for all t , where $\lambda_t \geq 0$, $\gamma_t \geq 0$; $\mathbb{E}[V_0] < \infty$ and (a.s.)

$$\sum_{t=0}^{\infty} |\mu_t| + \gamma_t < \infty, \quad (23)$$

then sequence $(V_t)_{t=0}^{\infty}$ converges almost surely to a finite random variable and $\sum_{t=0}^{\infty} \lambda_t < \infty$ almost surely.

Proof of Theorem 1. We are going to apply Theorem 3 to $V_t \triangleq (\theta_t - \theta^*)^2$. First, let us check whether (22) holds

$$\begin{aligned} \mathbb{E}[(\theta_{t+1} - \theta^*)^2 | \mathcal{F}_t] &= \mathbb{E}[\theta_{t+1}^2 + \theta^{*2} - 2\theta_{t+1}\theta^* | \mathcal{F}_t] = \\ &= \mathbb{E}[(\theta_t - \alpha_t G_t(\theta_t))^2 + \theta^{*2} - 2(\theta_t - \alpha_t G_t(\theta_t))\theta^* | \mathcal{F}_t] = \\ &= \mathbb{E}[\theta_t^2 + \alpha_t^2 G_t^2(\theta_t) - 2\theta_t \alpha_t G_t(\theta_t) + \theta^{*2} - 2\theta_t \theta^* - \\ &\quad - 2\alpha_t G_t(\theta_t)\theta^* | \mathcal{F}_t] = \theta_t^2 + \theta^{*2} - 2\theta_t \theta^* + \\ &\quad + \alpha_t^2 g_t^2(\theta_t) - 2\theta_t \alpha_t g_t(\theta_t) - 2\alpha_t g_t(\theta_t)\theta^* = \\ &= (\theta_t - \theta^*)^2 + \alpha_t^2 g_t^2(\theta_t) - 2\alpha_t g_t(\theta_t)(\theta^* - \theta_t) \end{aligned}$$

Then by setting $\mu_t \triangleq 0$, $\lambda_t \triangleq 2\alpha_t g_t(\theta_t)(\theta_t - \theta^*)$ and $\gamma_t \triangleq \alpha_t^2 g_t^2(\theta_t)$ the conditions of Theorem 3 are satisfied, because, e.g., γ_t is nonnegative and square summable, since g_t^2 is bounded and Assumption 4. Its also easy to see that because of Assumptions 1, 2 and 5, λ_t is (\mathbb{P} -a.s.) nonnegative.

Consequently, we know that $(\theta_t - \theta^*)^2$ converges (\mathbb{P} -a.s.) to a finite random variable and, moreover, (\mathbb{P} -a.s.)

$$\sum_{t=0}^{\infty} \alpha_t g_t(\theta_t)(\theta_t - \theta^*) < \infty, \quad (24)$$

We want to prove that $(\theta_t - \theta^*)^2$ converges (\mathbb{P} -a.s.) to 0. We prove this by *contradiction* and assume indirectly that $(\theta_t - \theta^*)^2$ converges (\mathbb{P} -a.s.) to another finite random variable Z that is not equal zero with some positive probability. Then,

$$|\theta_t(\omega) - \theta^*| \rightarrow Z(\omega) > 0 \quad \text{as } t \rightarrow \infty \quad (25)$$

for all $\omega \in A \subseteq \Omega$, and we have $\mathbb{P}(A) > 0$. It means that

$$\exists \varepsilon : 0 < \varepsilon < Z(\omega) : \exists t_0 : \forall t \geq t_0 : |\theta_t(\omega) - \theta^*| > \varepsilon, \quad (26)$$

for all $\omega \in A$; note that ε , t_0 and α_t may also depend on ω , but for simplicity, we will not denote these dependencies explicitly. Let us introduce a new random variable B as

$$B(\omega) \triangleq \min \{|g(\theta^* + Z(\omega) - \varepsilon)|, |g(\theta^* - Z(\omega) + \varepsilon)|\}$$

Using (26) and the construction of g and B , we have

$$\forall t \geq t_0 : 0 < B(\omega) \leq |g(\theta_t(\omega))| \leq |g_t(\theta_t(\omega))|, \quad (27)$$

for all $\omega \in A_0 \subseteq A \subseteq \Omega$, such that $\mathbb{P}(A_0) = \mathbb{P}(A) > 0$.

Using the results above and Assumption 3, we have

$$\sum_{t=t_0}^{\infty} \alpha_t |g(\theta_t(\omega))| |\tilde{\theta}_t(\omega)| \geq B(\omega) \varepsilon \sum_{t=t_0}^{\infty} \alpha_t = \infty, \quad (28)$$

for all $\omega \in A_0$, where $\tilde{\theta}_t(\omega) \triangleq \theta_t(\omega) - \theta^*$. But, we also have

$$\sum_{t=t_0}^{\infty} \alpha_t |g(\theta_t(\omega))| |\tilde{\theta}_t(\omega)| \leq \sum_{t=t_0}^{\infty} \alpha_t g_t(\theta_t(\omega)) \tilde{\theta}_t(\omega) < \infty,$$

for all $\omega \in A_1 \subseteq \Omega$, and we know that $\mathbb{P}(A_1) = 1$, which is a *contradiction*, thus, $|\theta_t - \theta^*|$ converges (\mathbb{P} -a.s.) to 0. \square

C. Relaxed Assumptions and Generalizations

There are some immediate generalizations of the theory. We can easily modify the results, e.g., to the case where the CDF of the noise in 0 at time t is known, $F_t(0) = q_t \neq 0$, but not necessary 1/2. In that case, the SA takes the form

$$\theta_{t+1} = \theta_t + \alpha_t [q_t - Y_t], \quad (29)$$

and a straightforward modification of the results (e.g., the bound in Lemma 1 changes to 4) shows the (\mathbb{P} -a.s.) convergence of this method, as well. Obviously, we can also allow the violation of Assumptions 1 and 2 for *finitely many* terms.

The *binary* identification of a *finite impulse response* (FIR) system can be traced back to finitely many instances of problem (1) by using a *periodic* input signal [14], thus, the presented SA scheme can also be applied for that problem.

D. Constrained Algorithm

In many applications the threshold cannot be set to an arbitrary real-number because of the limitations of the sensor. In this subsection, we will assume that the threshold can only take values from an *interval* $\Theta = [a, b]$. Then, we apply a *truncated* stochastic approximation scheme as follows

$$\theta_{t+1} = \Pi_{\Theta}(\theta_t + \alpha_t [1/2 - Y_t]), \quad (30)$$

where $\Pi_{\Theta} : \mathbb{R} \rightarrow \Theta$ is the *projection* mapping defined as follows: $\forall x \in [a, b] : \Pi_{\Theta}(x) = x$, $\forall x < a : \Pi_{\Theta}(x) = a$ and $\forall x > b : \Pi_{\Theta}(x) = b$. We will apply another assumption:

Assumption 6: The true parameter $\theta^* \in (a, b)$.

Then we can claim the (\mathbb{P} -a.s.) convergence of (30) as

Theorem 4: Under Assumptions 1-6, the sequence generated by iteration (30) converges almost surely to the true parameter θ^* from any initial parameter estimate $\theta_0 \in \Theta$.

Proof. We will show that the (\mathbb{P} -a.s.) convergence of (30) can be traced back to (\mathbb{P} -a.s.) convergence of (16). We will construct an *unconstrained* algorithm that behaves the same way as (30) in case we remove those elements from the trajectories that are outside of Θ . The new algorithm also starts from $\theta_0 \in \Theta$ and proceeds according to (16), but, if it leaves Θ , we introduce a “correction” step as follows. Assume that $\theta_t + \alpha_t [1/2 - Y_t] > b$, then we introduce a new time index $t < t' < t+1$ and let $\theta_{t'} \triangleq \theta_t + \alpha_t [1/2 - Y_t]$. Then, we use a noise at t' that is uniform in $[\theta^* - b, b - \theta^*]$, thus, $F_{t'}(\theta_{t'} - \theta^*) = 1$ (the noise can depend on the available information) and apply step-size $\alpha_{t'} \triangleq 2(\theta_{t'} - b)$. Then, this step corrects the error and we will (\mathbb{P} -a.s.) have $\theta_{t+1} = \theta_{t'} + \alpha_{t'} [1/2 - Y_{t'}] = b$. The case when $\theta_{t'} < a$ is similar.

The new algorithm works on the index set T' and if we consider only those elements that have indexes in T , the behavior of the constructed algorithm is the same as the behavior of the projected one. In the worst-case, we have a correction step between each t and $t+1$, but $(\alpha_t)_{t \in T'}$ still satisfies Assumptions 3-5, since for all t' : $0 < \alpha_{t'} \leq \alpha_t$. Additionally, the corrective noise terms satisfy Assumption 1 and 2, consequently, the new algorithm converges (\mathbb{P} -a.s.) to θ^* , hence, the projected original one also does. \square

IV. ACTIVE LEARNING

Now, let us consider the problem which arises if the threshold θ of the binary sensor can only be set to one of *finitely many* values from a given set $\Theta_0 = \{\theta_1, \dots, \theta_m\}$. In this section we will work with the following assumptions:

Assumption 7: The noise $(N_t)_{t \in T}$ is i.i.d. and its CDF is symmetric about zero: $F(x) = 1 - F(-x)$, for all $x \in \mathbb{R}$.

Assumption 8: For all $i \neq j$: $F(\theta_i - \theta^*) \neq F(\theta_j - \theta^*)$.

The nature of this problem is a bit different than the one we investigated in the previous section. For example, in general, unless we assume the full knowledge of the CDF, we cannot expect to identify the parameter θ^* asymptotically, since a consistent estimator may not exist. We can only hope to find a threshold from Θ_0 that is the “closest” to the parameter θ^* . This relaxed task is, however, too easy if we can make unlimited observations, since we could simply measure $Y_t(\theta)$ periodically for each $\theta \in \Theta_0$ and maintain the sample mean, $\hat{P}_t(\theta)$, cf. equation (4), of all $\theta \in \Theta_0$. Then, as before, $\hat{P}_t(\theta) \rightarrow F(\theta - \theta^*)$ almost surely as $t \rightarrow \infty$ and the best threshold is simply the one with $F(\theta - \theta^*)$ closest to $1/2$.

A. Multi-Armed Bandit Setting

As we saw, allowing unlimited number of measurements makes the problem easy. Here we take a different route and try to identify the best candidate threshold as soon as possible, using only *finite* number of observations. We will show that this can be reformulated as a multi-armed bandit problem and a *probably approximately correct* (PAC) solution can be given with an active learning method.

An m -armed bandit problem can be defined by m distributions ν_1, \dots, ν_m . In each step the decision-maker can select a distributions and sample it. This procedure is called taking an *action*. The obtained value is referred to as the *immediate-reward* and, in the classical setup, the aim is to maximize the expected total rewards or minimize the regret, which is the difference between the accumulated rewards and the total payoff based on the complete knowledge of the distributions [13]. In the PAC setting of *active learning* [10], however, the goal is to find a *stopping rule* for the identification of an (ε, δ) -optimal arm. That is a distribution ν_i with

$$\mathbb{P}(|\mu_i - \mu^*| \leq \varepsilon) \geq 1 - \delta, \quad (31)$$

where μ_i is the expected value of ν_i and $\mu^* = \max_j \mu_j$.

The binary identification problem with finitely many thresholds can be reformulated as an active learning problem as follows. We define an *action*, a_θ , for each threshold $\theta \in \Theta_0$. The result of performing action a_θ is that we take two consecutive measurements with θ . Let us denote the resulting observations by $Y_0(\theta)$ and $Y_1(\theta)$. Then, we define the *immediate-reward* of action a_θ , denoted by $R(\theta)$, as

$$R(\theta) \triangleq \begin{cases} 1 & \text{if } Y_0(\theta) \neq Y_1(\theta), \\ 0 & \text{otherwise.} \end{cases} \quad (32)$$

The intuition behind this definition is that if threshold θ is “close” to θ^* , then the observations are 0 and 1 with “almost” equal probability. We have the following result.

Lemma 2: Under Assumptions 7 and 8, the best action a_θ in the above defined multi-armed bandit problem corresponds to a threshold $\theta \in \Theta_0$ whose distance from θ^* is minimal.

Proof. The expected immediate-reward of action a_θ is

$$\begin{aligned} \mathbb{E}[R(\theta)] &= \mathbb{E}[Y_0(1 - Y_1) + (1 - Y_0)Y_1 | \theta] = \\ &= \mathbb{E}[\mathbb{I}(X_0 \leq \theta)\mathbb{I}(X_1 > \theta) + \mathbb{I}(X_0 > \theta)\mathbb{I}(X_1 \leq \theta) | \theta] = \\ &= \mathbb{P}(\theta^* + N_0 \leq \theta | \theta)(1 - \mathbb{P}(\theta^* + N_1 \leq \theta | \theta)) + \\ &\quad + (1 - \mathbb{P}(\theta^* + N_0 \leq \theta | \theta))\mathbb{P}(\theta^* + N_1 \leq \theta | \theta) = \\ &= 2(F(\theta - \theta^*) - F(\theta - \theta^*)^2) \end{aligned}$$

Now, let $z \triangleq F(\theta - \theta^*)$, then $2(z - z^2)$ is a parabola which has its maximum at $1/2$, when $\theta = \theta^*$. Using that function F is nondecreasing and its range is $[0, 1]$ as well as the two assumptions, namely that F is symmetric and different thresholds have different expected rewards, we have that if $|\theta_0 - \theta^*| < |\theta_1 - \theta^*|$ for $\theta_0, \theta_1 \in \Theta_0$, then we also have that $\mathbb{E}[R(\theta_1)] < \mathbb{E}[R(\theta_0)] \leq \mathbb{E}[R(\theta^*)] = 1/2$ and if $|\theta_0 - \theta^*| = |\theta_1 - \theta^*|$, then $\mathbb{E}[R(\theta_1)] = \mathbb{E}[R(\theta_0)]$. \square

B. Probably Approximately Correct Algorithm

An efficient algorithm to find an (ε, δ) -optimal PAC solution of the active learning problem is the method of *median elimination* [3]. The idea of the algorithm is to eliminate the “worst” half of the actions iteratively, until only one arm remains. An action is in the “worst” half if its estimated mean reward is less than the median of the estimated mean rewards. Naturally, we cannot expect to identify the optimal arm in this way, but we are satisfied with a PAC solution.

In order to find a guaranteed (ε, δ) -optimal solution, the median elimination algorithm requires $\Theta((m/\varepsilon^2) \log(1/\delta))$ trials, where $f(n) \in \Theta(h(n))$ if and only if $\exists k_1, k_2, n_0$, with $k_1 h(n) \leq f(n) \leq k_2 h(n)$, for all $n \geq n_0$. It can be proven that for the general multi-armed bandit active learning problem this bound is *asymptotically tight*, i.e., the bound on the number of trials is both sufficient and necessary [10].

Given the bandit formulation of the binary identification problem, this algorithm can be applied to get a PAC estimate of the threshold closest to θ^* , while noting that sampling an arm corresponds to making two observations with the sensor.

THE MEDIAN ELIMINATION ALGORITHM

Input: ε , maximum tolerated error, δ , confidence probability

1. $t := 0; k := 0; S_0 := \Theta_0; Q_0 := \underline{0};$
2. $\varepsilon_0 := \varepsilon/4; \delta_0 := \delta/2;$
3. *Repeat*
4. *Repeat* $[1/(\varepsilon_k/2)^2 \log(3/\delta_k)]$ *times*
5. *For each* $\theta \in S_k$ *do*
6. *Apply action* a_θ *and observe reward* $R_t(\theta);$
7. $Q_{t+1}(\theta) := (R_t(\theta) + t Q_t(\theta))/(t + 1);$
8. *End loop*
9. $t := t + 1;$
10. *End loop*
11. $S_{k+1} := \{\theta \in S_k : Q_t(\theta) \geq \text{Median}(Q_t)\};$
12. $\varepsilon_{k+1} := 3/4 \varepsilon_k; \delta_{k+1} := \delta_k/2; k := k + 1;$
13. *Until* $|S_k| = 1$

Output: θ , the only remaining threshold in S_k

V. EXPERIMENTAL RESULTS

The projected stochastic approximation based solution was tested with different noises. In this experiment the noise was i.i.d. with 3 distributions: Gaussian $(0, 1)$, Laplacian $(0, 2)$ and Uniform $(-10, 10)$. The parameters were as follows: $\alpha_t = 2/t$, $\theta_0 = 7$ and $\theta^* = \sqrt{2}$. The iterations were projected back to $\Theta = [1, 7]$ according to (30). In Figure 1 the average trajectories, the standard deviation areas and a particular trajectory realization for each case are presented. It was made by averaging 100 trajectories of 10^3 iterations. In agreement with the theory, the estimates converge to the parameter θ^* and the closer the CDF of the noise is to $1/2$ in an interval around the origin, the slower the convergence.

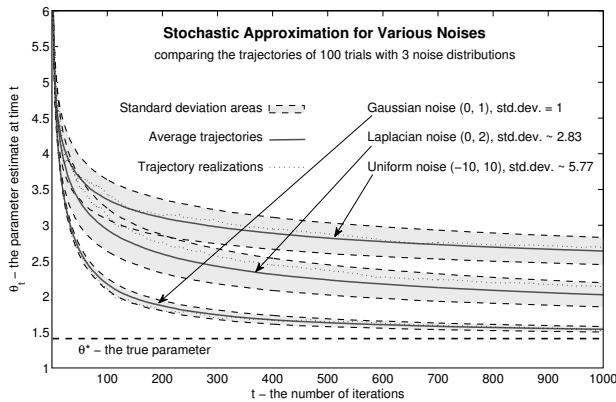


Fig. 1. Stochastic approximation trajectories of binary identification for three kinds of noise sequences: Gaussian, Laplacian and uniform.

The active learning approach to find a PAC solution for the finite binary identification problem was tested with an i.i.d. standard normal noise sequence and a 32-armed bandit. Figure 2 shows the arm elimination process when the true parameter θ^* was 1 and the thresholds formed an equidistant grid in $[0, 2]$. The unit of the figure is 10^5 trials of all non-eliminated arms (each such trial needs 2 binary observations). In this test a $(0.1, 0.05)$ -optimal PAC solution was the aim. The grey bars correspond to the thresholds and indicate the presence or absence of them in the maintained set, S_k . At the end, only two thresholds remained and the algorithm found the best solution (illustrated by the dark grey part).

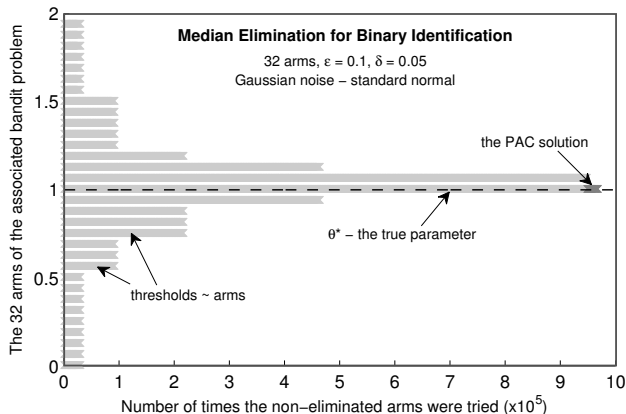


Fig. 2. Identifying a PAC threshold with median elimination.

VI. CONCLUDING REMARKS

The problem of estimating a constant parameter based on noisy observations from a *binary* sensor has been analyzed. There are standard strongly consistent and asymptotically efficient solutions to this problem, but only under strong assumptions on the noise, for example, that it is i.i.d. and its cumulative distribution function is *known* and invertible. We investigated a modified version of the original problem, where the *threshold of the sensor is under our control*, but where the statistical assumptions on the noise are significantly reduced, e.g., their distributions can be *unknown* apart from one point, the noise terms can be dependent, they can be nonstationary and they may not have densities.

We suggested two types of solutions depending on whether the threshold can be set to any value or only to a finite set of pre-defined values. In the first case, we showed that we can apply a *stochastic approximation* scheme and we proved that this method results in a *strongly consistent* estimator. In the second case, we proved, under the additional assumption that the noise is i.i.d. and symmetric about zero, that this problem can be formulated as a *multi-armed bandit* problem. We gave a *probably approximately correct* solution with the help of a successive elimination type active learning algorithm.

The theory immediately generalizes to FIR systems and a future direction is to extend it to other dynamical systems.

REFERENCES

- [1] M. Casini, A. Garulli, and A. Vicino. Input design in worst-case system identification using binary sensors. *IEEE Transactions on Automatic Control*, 56:1186–1991, 2011.
- [2] E. Colinet and J. Juillard. A weighted least-squares approach to parameter estimation problems based on binary measurements. *IEEE Transactions on Automatic Control*, 55:148–152, 2010.
- [3] E. Even-Dar, S. Mannor, and Y. Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research (JMLR)*, 7:1079–1105, 2006.
- [4] L. Gerencsér and Zs. Vágó. Stochastic approximation for function minimization under quantization error. In *Proceedings of the 38th IEEE CDC, Phoenix, Arizona, USA*, pages 2373–2377, 1999.
- [5] B. I. Godoy, G. C. Goodwin, J. C. Agüero, D. Marelli, and T. Wigren. On identification of FIR systems having quantized output data. *Automatica*, 47:1905–1915, 2011.
- [6] C. A. Goodsell and D. L. Hanson. Almost sure convergence for the Robbins-Monro process. *Annals of Probability*, 4(6):890–901, 1976.
- [7] Graham C. Goodwin and Kwai Sang Sin. *Adaptive Filtering: Prediction and Control*. Prentice-Hall, 1984.
- [8] Harold J. Kushner and G. George Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2nd edition, 2003.
- [9] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River, N J, 2nd edition, 1999.
- [10] Shie Mannor and John N. Tsitsiklis. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research (JMLR)*, 5:623–648, 2004.
- [11] Raghu Pasupathy and Sujin Kim. The stochastic root-finding problem: overview, solutions, and open questions. *ACM Transactions on Modeling and Computer Simulation*, 21, 2011.
- [12] James C. Spall. *Introduction to Stochastic Search and Optimization*. Wiley-Interscience, 2003.
- [13] Csaba Szepesvári. *Algorithms for Reinforcement Learning*. Morgan & Claypool Publisher, 2010.
- [14] Le Yi Wang, G. George Yin, Ji-Feng Zhang, and Yanlong Zhao. *System Identification with Quantized Observations*. Birkhäuser, 2010.
- [15] E. Weyer, S. Ko, and M. C. Campi. Finite sample properties of system identification with quantized output data. In *Procs. of the 48th IEEE CDC, Shanghai, China, December 16-18*, pages 1532–1537, 2009.