Undermodelling Detection with Sign-Perturbed Sums [⋆]

Algo Carè*, Marco C. Campi[†] Balázs Cs. Csáji** Erik Weyer[‡]**

*Centrum Wiskunde & Informatica (CWI), Science Park 123, Amsterdam, The Netherlands; (email: algocare@gmail.com)
[†]Department of Information Engineering, University of Brescia, Via Branze 38, 25123 Brescia, Italy; (email: marco.campi@unibs.it)
**Institute for Computer Science and Control (SZTAKI), Hungarian Academy of Sciences (MTA), Kende utca 13-17, H-1111, Budapest, Hungary; (email: balazs.csaji@sztaki.mta.hu)
[‡]Department of Electrical and Electronic Engineering, Melbourne School of Engineering, The University of Melbourne, Parkville, Melbourne, Victoria, 3010, Australia; (email: ewey@unimelb.edu.au)

Abstract: Sign-Perturbed Sums (SPS) is a finite sample system identification method that can build exact confidence regions for the unknown parameters of linear systems under mild statistical assumptions. Theoretical studies of the SPS method have assumed so far that the order of the system model is known to the user. In this paper we discuss the implications of this assumption for the applicability of the SPS method, and we propose an extension that, under mild assumptions, i) still delivers guaranteed confidence regions when the model order is correct, and ii) it is guaranteed to detect, in the long run, if the model order is wrong.

Keywords: system identification, confidence regions, finite sample results, least squares, parameter estimation, distribution-free results

1. INTRODUCTION

Estimating parameters of partially unknown systems based on observations corrupted by noise is a fundamental problem in system identification, signal processing, machine learning and statistics, (Lehmann and Casella, 1998; Ljung, 1999; Hastie et al., 2009). Standard solutions such as the Least Squares (LS) method or, more generally, prediction error methods provide *point estimates*. In many situations, for example, when the safety, stability or quality of a process has to be guaranteed, a point estimate should be accompanied with a *confidence region* that certifies the accuracy of the estimate. The *Sign-Perturbed Sums* (SPS) method (Csáji et al., 2012b, 2014, 2015; Kolumbán et al., 2015) constructs confidence regions which have an *exact* coverage probability of the true system parameter based only on a *finite* sample of observations and under very *mild* statistical assumptions on the noise terms.

Consider an ARX system:

$$Y_t + a_1^* Y_{t-1} + \cdots a_{n_a}^* Y_{t-n_a} \triangleq b_1^* U_{t-1} + \cdots b_{n_b}^* U_{t-n_b} + N_t, \quad (1)$$

where Y_t is the output, N_t is the noise, and U_t is a measured input, at time t. This system can be written in linear regression form as follows $Y_t = \varphi_t^\top \theta^* + N_t$, where the regressor vectors φ_t is defined as $\varphi_t = [-Y_{t-1}, \ldots, -Y_{t-n_a}, U_{t-1}, \ldots, U_{t-n_b}]^\top$, and $\theta^* = [a_1^*, \ldots, a_{n_a}^*, b_1^*, \ldots, b_{n_b}^*]^\top$ is referred to as the true parameter.

Following (Csáji et al., 2015), we assume that the measured inputs $\{U_t\}$ are *deterministic*, but all the results here presented can be immediately generalised to random inputs when they are independent of the noise. The SPS algorithm builds exact confidence regions for the unknown parameters under the assumption that the noise sequence $\{N_t\}$ is independent and symmetric (not necessarily identically distributed). [1] Given n observations $Y_1, \ldots, Y_n$ and $\varphi_1, \ldots, \varphi_n$, SPS constructs confidence regions that contain the LS estimate $\hat{\theta}_n$, which is defined as the minimiser of the sum of the squared prediction errors, that is

$$\hat{\theta}_n \triangleq \underset{\theta \in \mathbb{R}^{n_a+n_b}}{\arg\min} \sum_{t=1}^{n} (Y_t - \varphi_t^\top \theta)^2. \quad (2)$$

In the construction of the SPS regions, a crucial role is played by the fact that the system can be inverted, and the symmetric noise sequence, $\{N_t\}$, can be correctly recovered when the true parameter θ^* is correctly guessed. As a consequence of this fact, in order for the standard SPS method to be rigorously guaranteed by the theory, the knowledge of the "true" model order (or at least an upper bound of it) must be available.

[1] For a discussion of the robustness of SPS with respect to violations of the symmetry assumptions, see (Carè et al., 2016).

Model order selection is a standard topic in system identification and various tools are available to the user, see, e.g., (Ljung, 1999; Stoica and Selen, 2004; Pillonetto et al., 2013). However, it is still a challenging problem, and it is realistic to assume that the selection procedure might end up with a model order $(\hat{n}_a, \hat{n}_b)$ smaller than the "true" one, i.e., with $\hat{n}_a < n_a$ and/or $\hat{n}_b < n_b$. So far theoretical studies of SPS did not consider this possibility, while a simulation experiment on the effects of undermodelling was carried out in (Csáji et al., 2015).

Finite sample methods for the case when the input signal can be designed are available and can be used to obtain guaranteed confidence regions when the user is interested in estimating *a subset* of the parameters, see (Campi et al., 2009). In principle, these techniques can be applied also when the true model order is unknown but, arguably, higher than the selected one. However, they are not aimed at obtaining regions around the LS estimate, and they can be applied only if the input satisfies certain statistical properties, such as symmetry.

Aim of the paper
In this paper, we move a step towards SPS methods that can be used in the presence of imperfect knowledge of the true model order, or even of the true system structure. In particular, we propose an approach to modify the existing SPS method in such a way that:

- If the system is not undermodelled, the algorithm builds exact confidence regions for the true model parameter θ^*. This property holds true under the same assumptions as for the original SPS.

- On the other hand, if the system is undermodelled, the algorithm detects undermodelling as soon as a sufficient amount of data is available. This property holds true under some mild, additional assumptions.

Structure of the paper
In the following Section 2 the standard SPS method and its main properties are briefly reviewed. The issues of standard SPS in the presence of undermodelling are discussed in Section 3 and provide a motivation for a new algorithm, which we call UD-SPS (SPS with Undermodelling Detection). UD-SPS is presented in Section 4.

The results of this paper deal with an FIR-oriented SPS method, which is an archtypical case that allows us to point out the main ideas without technical complications. However, in Section 5 we argue that our ideas are applicable to more general models. The properties of the new UD-SPS algorithm are illustrated on some simulation examples in Section 6, while conclusions are drawn in Section 7.

2. REVIEW OF THE STANDARD SPS ALGORITHM

The SPS algorithm in its standard form (Csáji et al., 2015, 2014) is summarised in this section.

We will assume that $n_a = 0$, that is, we restrict ourselves to the FIR case where the regressors $\{\varphi_t\}$ do not depend on the noise $\{N_t\}$.[2] Recall that the LS estimate, see formula (2), can be obtained by solving the normal equation,

---

[2] Extensions to ARX and more general systems are available (Csáji et al., 2012b,a; Kolumbán et al., 2015).

$$\sum_{t=1}^{n} \varphi_t (Y_t - \varphi_t^\top \theta) = 0, \qquad (3)$$

which, when $\sum_{t=1}^{n} \varphi_t \varphi_t^\top$ is invertible (this will be always assumed throughout this paper), has the analytic solution

$$\hat{\theta}_n = \left(\sum_{t=1}^{n} \varphi_t \varphi_t^\top \right)^{-1} \left(\sum_{t=1}^{n} \varphi_t Y_t \right).$$

The fundamental step in SPS consists of generating $m-1$ *sign-perturbed sums* by randomly perturbing the sign of the prediction error in the normal equations (3), that is, for $i = 1, \ldots, m-1$, we define

$$H_i(\theta) \triangleq \sum_{t=1}^{n} \varphi_t \alpha_{i,t} (Y_t - \varphi_t^\top \theta),$$

where $\{\alpha_{i,t}\}$ are random signs, i.e., i.i.d. random variables that take on the values ± 1 with probability $1/2$ each. For a given θ, the *reference sum* is instead defined as

$$H_0(\theta) \triangleq \sum_{t=1}^{n} \varphi_t (Y_t - \varphi_t^\top \theta).$$

It is shown in (Csáji et al., 2015, 2014) that a suitable linear transformation of these sums ensures better properties, and therefore we apply the following functions,

$$S_i(\theta) \triangleq \frac{1}{n} R_n^{-\frac{1}{2}} H_i(\theta), \ i = 0, \ldots, m-1,$$

where $R_n = \frac{1}{n} \sum_{t=1}^{n} \varphi_t \varphi_t^\top$ and "$-\frac{1}{2}$" denotes the inverse of its square root.

Denote by $\mathcal{R}(\theta)$ the *rank* of $\|S_0(\theta)\|$ in the ordering of $\|S_0(\theta)\|, \|S_i(\theta)\|, i = 1, \ldots, m-1$, e.g., $\mathcal{R}(\theta) = 1$ means that $\|S_0(\theta)\|$ is the smallest one, and so on. In case of ties, the rank is broken by randomisation, see (Csáji et al., 2015) for details. The SPS region is formally defined as

$$\widehat{\Theta}_n \triangleq \left\{ \theta : \mathcal{R}(\theta) \leq m \left(1 - \frac{q}{m} \right) \right\},$$

and the following theorem holds true, (Csáji et al., 2015).

Theorem 1. (Exact confidence of SPS). *If $N_1, \ldots, N_n$ is a sequence of independent random variables distributed symmetrically about zero, then it holds that*

$$\Pr\{\theta^* \in \widehat{\Theta}_n\} = 1 - \frac{q}{m}. \qquad *$$

Moreover, under some mild additional assumptions, SPS is strongly consistent, that is, for every $\varepsilon > 0$, $\widehat{\Theta}_n$ is almost surely contained in an ε-ball around the true parameter θ^* for sufficiently large n (Csáji et al., 2014).

3. SPS IN THE PRESENCE OF UNDERMODELLING

In this section, we discuss the behaviour of the SPS algorithm when the model chosen by the user does not correspond to the true data-generation mechanism.

3.1 The user-chosen model

We assume that the user has decided to use the FIR model

$$\widehat{Y}_t(\theta) = \varphi_t^\top \theta \qquad (4)$$

for predicting Y_t, where $\varphi_t = [U_{t-1}, \ldots, U_{t-\hat{n}_b}]$ and $\theta = [b_1, \ldots, b_{\hat{n}_b}]$ is a generic parameter. This assumption is

held throughout the paper, where all the applications of the SPS algorithm use (4) as model structure. However, it can be relaxed as discussed in Section 5.

3.2 The true data-generation mechanism

If data $\{Y_t\}$ are *really* generated according to a FIR system of order $n_b = \hat{n}_b$, that is

$$Y_t = b_1^* U_{t-1} + b_2^* U_{t-2} + \cdots + b_{\hat{n}_b}^* U_{t-\hat{n}_b} + N_t, \quad (5)$$

where $\{N_t\}$ is independent and symmetric, then the standard SPS is guaranteed to deliver a confidence region for θ that contains the true parameter, $\theta^* = [b_1^*, \ldots, b_{\hat{n}_b}^*]$, with user-chosen probability. Moreover, under mild assumptions, $\theta \to \theta^*$ as $n \to \infty$, and the SPS regions shrinks around θ^*, see Section 2 and references therein.

It is interesting to consider, instead, the situation where data $\{Y_t\}$ are generated by a higher order FIR model or even a more general linear system. Assume therefore that Y_t can be written as follows

$$Y_t = \varphi_t^\top \theta^* + E_t + N_t, \quad (6)$$

where $\{N_t\}$ is the usual, independent symmetric noise; the (linear) effect of $U_{t-1}, \ldots, U_{t-\hat{n}_b}$ is correctly described by the term $\varphi_t^\top \theta^*$; E_t is an extra, non-zero component that can depend *linearly* on *all* the past inputs $U_{t-\hat{n}_b-1}, U_{t-\hat{n}_b-2}, \ldots$ and on *all* the past noise samples $N_{t-1}, N_{t-2} \ldots$. For example, this situation includes the case where the true system is ARX as in (1). In the special case where the true system is an FIR of order $n_b > \hat{n}_b$, E_t depends linearly on $U_{t-\hat{n}_b-1}, \ldots, U_{t-n_b}$ only.

3.3 The effect of undermodelling on standard SPS

Under (6), the conditions of Theorem 1 are not met in general by the "noise" $\{N_t + E_t\}$, so the confidence region built by SPS is not guaranteed. The next theorem characterises the asymptotic behaviour of the SPS region in the case of undermodelling. Defining $\hat{\theta}_\infty$ as the limit of the LS estimate $\hat{\theta}_n$, the theorem states that, under some mild assumptions, for every $\varepsilon > 0$, the SPS region $\widehat{\Theta}_n$ will remain inside an ε-ball centred around $\hat{\theta}_\infty$ for all n large enough. Since $\hat{\theta}_\infty \neq \theta^*$ in general, the theorem implies that the SPS region will *not* include θ^* in the long run.

Theorem 2. (Asymptotic behaviour with undermodelling). Assume that (6) is the true data-generating mechanism. Define $R_n = \frac{1}{n} \sum_{t=1}^n \varphi_t \varphi_t^\top$, and assume that $\det(R_n) \neq 0$ and there exists a finite limit matrix $\bar{R}$, $\bar{R} \succ 0$ such that

$$\lim_{n\to\infty} R_n = \bar{R} \succ 0. \quad (7)$$

Assume also that there is a finite real vector $\bar{E}$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^n \varphi_t \mathbb{E}[E_t] = \bar{E}, \quad (8)$$

and, moreover,

$$\sum_{t=1}^\infty \frac{\|\varphi_t\|^4}{t^2} < \infty, \quad \sum_{t=1}^\infty \frac{\mathbb{E}[\|N_t\|^2]^2}{t^2} < \infty, \quad \sum_{t=1}^\infty \frac{\mathbb{E}[\|E_t\|^2]^2}{t^2} < \infty. \quad (9)$$

Then,

$$\hat{\theta}_\infty \triangleq \lim_{n\to\infty} \hat{\theta}_n = \theta^* + \bar{R}^{-1}\bar{E} \quad \text{(w.p.1)}, \quad (10)$$

and, for all $\varepsilon > 0$,

$$\Pr\left[\bigcup_{\bar{n}=1}^\infty \bigcap_{n=\bar{n}}^\infty \left\{ \widehat{\Theta}_n \subseteq B_\varepsilon(\hat{\theta}_\infty) \right\} \right] = 1,$$

where $B_\varepsilon(\theta)$ denotes an ε-ball centred around θ.   ∗

Technically, the theorem states that the event that $\widehat{\Theta}_n \subseteq B_\varepsilon(\hat{\theta}_\infty)$ for all n larger than some (realisation-dependent) value $\bar{n}$ is a probability 1 tail-event. Thus, by (10), if $\bar{E}$ is nonzero, that is, if the sequence of the residuals $\{E_t\}$ is correlated with the sequence of the regressors $\{\varphi_t\}$ (in the sense of (8)), the region built by SPS by using the model (4) shrinks around the "wrong" value $\hat{\theta}_\infty \neq \theta^*$.

By relying only on the standard SPS algorithm, there is no way for the user to recognise that the assumptions are not satisfied and the SPS region is going to exclude the true parameter systematically. This is the motivation for the work of this paper and for the SPS algorithm with undermodelling detection, presented in the next section.

4. UD-SPS: A MODIFIED SPS METHOD

We now define the UD-SPS algorithm and discuss the main ideas behind it. Explaining the connection between UD-SPS and the standard SPS makes it easy to prove that UD-SPS inherits the most important properties of standard SPS when the system is correctly specified. Finally, we study the undermodelling-detection property of UD-SPS.

4.1 Definition of UD-SPS

The UD-SPS algorithm is obtained from the standard SPS algorithm by substituting the functions $S_0(\theta), \ldots, S_{m-1}(\theta)$ with the following ones

$$Q_0(\theta) \triangleq \begin{bmatrix} R_n & B_n \\ B_n^\top & D_n \end{bmatrix}^{-\frac{1}{2}} \frac{1}{n} \sum_{t=1}^n \begin{bmatrix} \varphi_t \\ \psi_t \end{bmatrix} (Y_t - \varphi_t^\top \theta),$$

$$Q_i(\theta) \triangleq \begin{bmatrix} R_n & B_n \\ B_n^\top & D_n \end{bmatrix}^{-\frac{1}{2}} \frac{1}{n} \sum_{t=1}^n \alpha_{i,t} \begin{bmatrix} \varphi_t \\ \psi_t \end{bmatrix} (Y_t - \varphi_t^\top \theta),$$

$$\text{for} \quad i = 1, \ldots, m-1, \quad (11)$$

where ψ_t is a vector that includes s extra input values preceding the $\hat{n}_b$ that are included in φ_t, i.e.,

$$\psi_t \triangleq [U_{t-\hat{n}_b-1}, \cdots, U_{t-\hat{n}_b-s}]^\top,$$

and

$$B_n \triangleq \frac{1}{n} \sum_{t=1}^n \varphi_t \psi_t^\top, \quad D_n \triangleq \frac{1}{n} \sum_{t=1}^n \psi_t \psi_t^\top.$$

So, while the prediction error $(Y_t - \varphi_t^\top \theta)$ in (11) is the usual prediction error for the user-chosen model class, the regressor vector and the shaping matrix are larger than in the standard SPS algorithm. s is a parameter that can be chosen by the user and, clearly, it need *not* be equal to the difference between the true order of the system, which is unknown, and $\hat{n}_b$. The region built by UD-SPS will be denoted by $\widehat{\Theta}_n^\circ$.

4.2 The idea of UD-SPS

The key idea is stated in the following Fact 1.

Fact 1. The UD-SPS region $\widehat{\Theta}_n^\circ$ for estimating $\theta^* \in \mathbb{R}^{\hat{n}_b}$ can be interpreted as the restriction to a $\hat{n}_b$-dimensional space of a full-fledged standard SPS region, say $\widehat{\Theta}_n'$, that lives in the domain $\{\theta' \in \mathbb{R}^{\hat{n}_b+s}\}$, which is the $\hat{n}_b$-dimensional identification space *augmented* with s extra components. More precisely, $\widehat{\Theta}_n^\circ$ can be identified with the first $\hat{n}_b$ components of the set $\widehat{\Theta}_n' \cap (\mathbb{R}^{\hat{n}_b} \times \{0\}^s)$.   *

In order to see that Fact 1 is true, consider the functions $S_0'(\theta'), S_1'(\theta'), \ldots, S_{m-1}'(\theta')$ of $\theta' \in \mathbb{R}^{\hat{n}_b+s}$ defined as

$$S_0'(\theta') = R_n'^{-\frac{1}{2}} \frac{1}{n} \sum_{t=1}^n \varphi_t'(Y_t - \varphi_t'^\top \theta'),$$

$$S_i'(\theta') = R_n'^{-\frac{1}{2}} \frac{1}{n} \sum_{t=1}^n \alpha_{i,t} \varphi_t'(Y_t - \varphi_t'^\top \theta'),$$

$$\text{for} \qquad i = 1, \ldots, m-1, \tag{12}$$

where

$$R_n' = \begin{bmatrix} R_n & B_n \\ B_n^\top & D_n \end{bmatrix} \tag{13}$$

and

$$\varphi_t'^\top = \begin{bmatrix} \varphi_t \\ \psi_t \end{bmatrix}^\top = [U_{t-1}, \ldots, U_{t-\hat{n}_b}, U_{t-\hat{n}_b-1}, U_{t-\hat{n}_b-s}]. \tag{14}$$

These are the standard S_i-functions based on which the *standard* SPS region, say $\widehat{\Theta}_n'$, can be built in the augmented space $\mathbb{R}^{\hat{n}_b+s}$ for the user-chosen model $\varphi_t'^\top \theta'$. Comparing (12) with (11), it can be immediately observed that functions (12) take the same values of functions (11) whenever θ' is restricted to $\mathbb{R}^{\hat{n}_b} \times \{0\}^s$, i.e.,

$$S_i'(\theta')|_{\theta'=[\theta^\top, \mathbf{0}^\top]} = Q_i(\theta). \tag{15}$$

Computational feasibility of the UD-SPS algorithm
There is no significant difference in the computational complexity between UD-SPS and the standard SPS: on the one hand, it is easy to check whether a certain value of θ is inside or outside an SPS region, while on the other hand computing and handling a complete and explicit representation of the region becomes unpractical in a high dimensional parameter space. For SPS, computationally feasible approximations techniques have been studied that rely on interval analysis, e.g. (Kieffer and Walter, 2013), or on semidefinite programming (SDP), (Csáji et al., 2015). We focus here on the latter option, which allows us to compute outer ellipsoidal approximations of SPS regions. Denote by $\widetilde{\Theta}_n'$ the outer-approximating ellipsoid of the SPS region $\widehat{\Theta}_n'$ in the augmented space $\mathbb{R}^{\hat{n}_b+s}$ (Fact 1). In (Csáji et al., 2015), $\widetilde{\Theta}_n'$ is defined as the set $\{\theta' \in \mathbb{R}^{\hat{n}_b+s} : \|S_0'(\theta')\|^2 \leq \gamma^*\}$, where γ^* can be computed from the solutions of some suitable (convex) SDP problems. In virtue of Fact 1, the restriction of this ellipsoid to the domain $\mathbb{R}^{\hat{n}_b} \times \{0\}^s$, denoted by $\widetilde{\Theta}_n^\circ$, can be written as

$$\widetilde{\Theta}_n^\circ = \{\theta \in \mathbb{R}^{\hat{n}_b} : \|Q_0(\theta)\|^2 \leq \gamma^*\}, \tag{16}$$

and contains $\widehat{\Theta}_n^\circ$.

Remark 1. With small modifications, it is possible to find a smaller outer ellipsoidal approximation for $\widehat{\Theta}_n^\circ$ than $\widetilde{\Theta}_n^\circ$ as defined above. In fact, the optimisation space of the SDP problems can be restricted to the domain $\mathbb{R}^{\hat{n}_b} \times$

$\{0\}^s$ with no harm. In general, from this restriction a smaller, but still valid, γ^* to be used in (16) can be obtained. Moreover, in this way, only $\hat{n}_b$ decision variables are involved in the optimisation problem instead of $\hat{n}_b + s$.

UD-SPS and the LS estimate   It can be shown that the LS estimate, $\hat{\theta}_n$, is always included in the outer-approximation ellipsoid, $\widetilde{\Theta}_n^\circ$, whenever $\widetilde{\Theta}_n^\circ$ is not empty:

Theorem 3. If $\widetilde{\Theta}_n^\circ \neq \emptyset$, then $\hat{\theta}_n \in \widetilde{\Theta}_n^\circ$.   *

4.3 UD-SPS with correct system specification

Now, we study the case when the system and its order are correctly specified, i.e., the true system is (5). In this case, the most important properties of standard SPS carry over to UD-SPS, by applying the standard SPS results (see Section 2) in the θ' space and then restricting the result to the first $\hat{n}_b$ coordinates (Fact 1). In particular, if (5) is the true system, $\widehat{\Theta}_n'$ is guaranteed to contain the "true" parameter $\theta'^* \triangleq [\theta^{*\top} 0 \cdots 0]^\top$ with probability $1 - \frac{q}{m}$, so the following theorem is obtained.

Theorem 4. (Exact confidence of UD-SPS). If the FIR system is correctly specified, i.e., (5) holds true, then

$$\Pr\{\theta^* \in \widehat{\Theta}_n^\circ\} = 1 - \frac{q}{m}. \qquad *$$

Moreover, UD-SPS is strongly consistent.

Theorem 5. (Strong consistency of UD-SPS). Assume that (7)-(9) and the following statements hold true

$$\bar{B} \triangleq \lim_{n \to \infty} B_n < \infty, \quad \bar{D} \triangleq \lim_{n \to \infty} D_n < \infty, \tag{17}$$

with

$$\begin{bmatrix} \bar{R} & \bar{B} \\ \bar{B}^\top & \bar{D} \end{bmatrix} \succ 0; \tag{18}$$

and

$$\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^n \psi_t \mathbb{E}[E_t] < \infty, \quad \sum_{t=1}^\infty \frac{\|\psi_t\|^4}{t^2} < \infty. \tag{19}$$

If the system is correctly specified, i.e., (5) holds true, then, for all $\varepsilon > 0$, we have that

$$\Pr\left[\bigcup_{\bar{n}=1}^\infty \bigcap_{n=\bar{n}}^\infty \left\{\widehat{\Theta}_n^\circ \subseteq B_\varepsilon(\theta^*)\right\}\right] = 1,$$

where $B_\varepsilon(\theta)$ denotes an ε-ball centred around θ.   *

The strong consistency of the outer approximation $\widetilde{\Theta}_n^\circ$ follows by the strong consistency of $\widetilde{\Theta}_n'$ in the augmented space, see (Csáji et al., 2015, 2014), and Fact 1.

4.4 UD-SPS in the presence of undermodelling

Consider now the case where the true data-generating mechanism is system (6). In this case, the region $\widehat{\Theta}_n^\circ$ is not guaranteed. However, the following theorem guarantees that, for n large enough, the region is empty.

Theorem 6. (Undermodelling detection). Assume that (6) is the true system, and relations (7)-(9) and (17)-(19) hold true. With the notation

$$\bar{R}' \triangleq \lim_{n \to \infty} \begin{bmatrix} R_n & B_n \\ B_n^\top & D_n \end{bmatrix}, \quad \bar{E}' \triangleq \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^n \begin{bmatrix} \varphi_t \\ \psi_t \end{bmatrix} \mathbb{E}[E_t],$$

if

$$\bar{R}'^{-1}\bar{E}' \notin \mathbb{R}^{\hat{n}_b} \times \{0\}^s, \qquad (20)$$

then

$$\Pr\left[\bigcup_{\bar{n}=1}^{\infty}\bigcap_{n=\bar{n}}^{\infty}\left\{\widehat{\Theta}_n^{\circ} = \emptyset\right\}\right] = 1. \qquad (21)$$

*

The main statement of the theorem is (21), which is formulated using the tail-event notation as in Theorems 2 and 5. Equation (21) means that, with probability 1, there is a (realisation-dependent) value of $\bar{n}$ such that the region $\widehat{\Theta}_n^{\circ}$ is empty for every $n \geq \bar{n}$. Condition (20) is a technical detectability condition, which, in practice, is expected to be met, unless E_t does not depend on the input or there are contrived correlation patterns in the input sequence. [3]

In concluding, after a certain amount of data $\bar{n}$ has been observed, UD-SPS warns the user when the method is working beyond its domain of applicability by building empty confidence regions. The number $\bar{n}$ depends on the system and the degree of misspecification, as we will see in the example in Section 6.

5. UD-SPS FOR MORE GENERAL SYSTEMS

The SPS algorithm has been generalised to ARX systems (Csáji et al., 2012b) and general linear systems (Csáji et al., 2012a; Kolumbán et al., 2015). It is possible to extend the asymptotic results that hold true for the FIR case to the ARX case. Relying on these extensions, the main arguments in Section 4, which rely only on the strong consistency of the SPS method, carry over to the ARX setting.

6. NUMERICAL EXPERIMENTS

Consider the following ARX(1,1) generating system

$$Y_t = a^* Y_{t-1} + b^* U_{t-1} + N_t,$$

with zero initial conditions, where $a^* = 0.7$ and $b^* = 1$ are the true system parameters and $\{N_t\}$ is a sequence of i.i.d. Laplacian random variables with zero mean and variance 0.1. The input signal is generated as $U_t = 0.75U_{t-1} + V_t$, where $\{V_t\}$ is a sequence of i.i.d. Gaussian random variables with zero mean and variance 1. The user-chosen predictor is

$$\widehat{Y}_t(\theta) = \varphi_t^{\top}\theta = b\, U_{t-1},$$

that is, the autoregressive part is missing, $\theta = [\,b\,]$ is the model parameter, and $\varphi_t = [\,U_{t-1}\,]$ is the input-dependent regressor at time t.

We choose $s = 1$ in the UD-SPS algorithm, that is, $\psi_t = [U_{t-2}]$ in (11). We construct the outer approximation ellipsoid $\widehat{\Theta}_n^{\circ}$ for the 95 % confidence UD-SPS region $\widehat{\Theta}_n^{\circ}$ by using the algorithm of Section 4.3, see definition (16).

---

[3] A notable situation where the detectability condition fails is when the true system is a FIR system with $n_b > \hat{n}_b$, the input is an uncorrelated sequence and none of the inputs $U_{t-\hat{n}_b-1}, \dots, U_{t-n_b}$ corresponding to a nonzero coefficient among b_k^*, $k = \hat{n}_b+1, \dots, n_b$ is included among the s extra components in the augmented regressor φ_t'. However, in this case, the LS estimate $\hat{\theta}_n$ will not be biased, i.e., $\hat{\theta}_n \to \theta^*$. Moreover, if the input can be thought of as the realisation of an independent and symmetric process, the region for $\theta^* \in \mathbb{R}^{\hat{n}_b}$ will be still guaranteed for finite samples (Theorem 4).
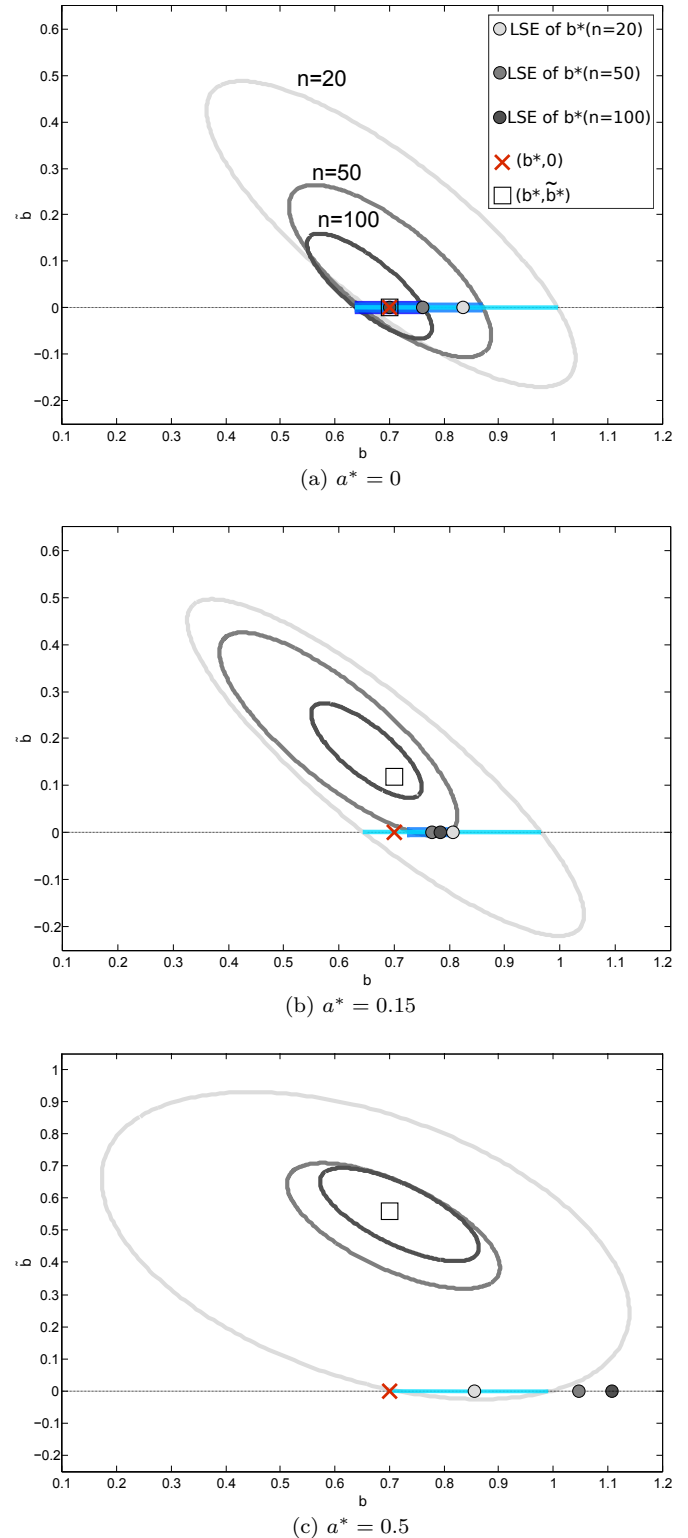

Fig. 1. In each of these pictures, the SPS ellipsoids $\widetilde{\Theta}_n'$ in the 2-dimensional (augmented) space for $n = 20$ (light gray), $n = 50$ (gray), $n = 100$ (dark gray) are shown. The corresponding outer-approximation of the UD-SPS region, $\widehat{\Theta}_n^{\circ}$, is obtained by intersecting the ellipsoid $\widetilde{\Theta}_n'$ with the b-axis (horizontal line). The "true" parameter b^* is also represented in the 2-dimensional space as $(b^*, 0)$, together with the convergence point $(b^*, \widetilde{b}^*)$ of the SPS region in the 2-dimensional space. The circles on the b-axis denote the LS estimates of b^*, namely $\hat{\theta}_{20}$ (light gray), $\hat{\theta}_{50}$ (gray), and $\hat{\theta}_{100}$ (dark gray).

Although this approximation algorithm can be refined in line with Remark 1, it is here used for illustrative purposes as it reflects in an intuitive manner the relation between UD-SPS and standard SPS (i.e., the fact that UD-SPS region can be obtained by restricting a higher-dimensional SPS region). Note that, in this case, $\widetilde{\Theta}_n^o$ is an interval.

Although the normal use of the method is in the 1-dimensional space where the model parameter $\theta = [\, b \,]$ takes value, in Fig. 1 the augmented 2-dimensional space is represented for explanatory purposes. In Fig. 1, the b-axis corresponds to the unknown parameter that we want to estimate and the $\tilde{b}$-axis is the extra coordinate accounting for the extra input in ψ_t. In this space, the SPS 2-dimensional ellipse $\widetilde{\Theta}_n'$ can be built according to the standard algorithm in (Csáji et al., 2015). According to Section 4.3, the interval $\widetilde{\Theta}_n^o$ can then be interpreted as the intersection of $\widetilde{\Theta}_n'$ with the b-axis.

Note that, as expected, whenever the 1-dimensional intersection of the SPS ellipsoid with the b-axis is non-empty, the LS estimate $\hat{\theta}_n$ is included in $\widetilde{\Theta}_n^o$. When $a^* = 0$ (Fig. 1a), the augmented SPS region shrinks around the true parameter $(b^*, 0)$ as n increases, and the corresponding UD-SPS interval becomes smaller and smaller around b^*. However, when $a^* \neq 0$, the system is misspecified and, as n increases, the augmented SPS region shrinks around a parameter value that does not lie on the b axis, so that $\widetilde{\Theta}_n^o$ becomes empty and undermodelling is detected. The limit point of SPS in the augmented space, denoted by $(b^*, \tilde{b}^*)$, can be computed according to formula (10). Undermodelling is detected when $\widetilde{\Theta}_n^o$ is empty. This happens when n is 100 in Fig.1b ($a^* = 0.15$) and is 50 in Fig.1c ($a^* = 0.5$).

Table 1.

n	UD-SPS ellipsoid coverage $(\theta^* \in \widetilde{\Theta}_n^o)$	Detection with UD-SPS ellipsoid $(\widetilde{\Theta}_n^o = \emptyset)$	Standard SPS ellipsoid coverage $(\theta^* \in \widetilde{\Theta}_n)$
$a^* = 0$			
20	99.8%	0.2%	98.7%
50	99.0%	0%	97.5%
100	98.6%	0.4%	97.7%
$a^* = 0.15$			
20	84.5%	2.4%	77.0%
50	29.5%	31.4%	41.5%
100	3.6%	72.2%	11.1%
$a^* = 0.5$			
20	13.3%	63.5%	37.1%
50	0%	99.9%	0.4%
100	0%	100%	0.1%

In Table 1, the results of 1000 Monte Carlo simulations are shown for the same three values of a^* and n that are used in Fig.1. The empirical coverage of $\widetilde{\Theta}_n^o$ is compared with the empirical coverage of the standard SPS outer interval $\widetilde{\Theta}_n \subseteq \mathbb{R}$, computed as in (Csáji et al., 2015), and the frequency with which $\widetilde{\Theta}_n^o$ is empty is also shown. In the cases of misspecification ($a^* = 0.15$ and $a^* = 0.5$), the frequency with which $\widetilde{\Theta}_n^o$ is empty estimates the probability that undermodelling is detected; in the case of correct system model ($a^* = 0$), the same frequency can be interpreted as an estimate of the probability of false detection, which turned out to be very small.

7. CONCLUSIONS

In this paper we have studied the behaviour of SPS, a guaranteed finite-sample system identification method, in presence of undermodelled dynamics. In this case, the confidence regions generated by the standard SPS algorithm are not rigorously guaranteed, nor do they provide any warning that the algorithm is working outside of its applicability domain.

We have proposed an extension of SPS, the UD-SPS algorithm, which is able to detect that the algorithm is working outside of its applicability domain. We have shown that UD-SPS provides guaranteed confidence regions if the model order is correctly specified, otherwise it almost surely detects undermodelling in the long run. Finally, we demonstrated UD-SPS through numerical experiments.

REFERENCES

Campi, M.C., Ko, S., and Weyer, E. (2009). Non-asymptotic confidence regions for model parameters in the presence of unmodelled dynamics. *Automatica*, 45, 2175–2186.

Carè, A., Csáji, B.Cs., and Campi, M.C. (2016). Sign-Perturbed Sums (SPS) with asymmetric noise: Robustness analysis and robustification techniques. In *Procs. of the 55th IEEE Conf. on Decision and Control*, 262–267.

Csáji, B.Cs., Campi, M.C., and Weyer, E. (2012a). A method for constructing exact finite-sample confidence regions for general linear systems. In *Procs. of the 51st IEEE Conference on Decision and Control*, 7321–7326.

Csáji, B.Cs., Campi, M.C., and Weyer, E. (2012b). Non-asymptotic confidence regions for the least-squares estimate. In *Procs. of the 16th IFAC Symposium on System Identification*, 227–232.

Csáji, B.Cs., Campi, M.C., and Weyer, E. (2014). Strong consistency of the Sign-Perturbed Sums method. In *Procs. of the 53th IEEE Conference on Decision and Control*, 3352–3357.

Csáji, B.Cs., Campi, M.C., and Weyer, E. (2015). Sign-Perturbed Sums: A new system identification approach for constructing exact non-asymptotic confidence regions in linear regression models. *IEEE Transactions on Signal Processing*, 63(1), 169–181.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning*. Springer, 2nd edition.

Kieffer, M. and Walter, E. (2013). Guaranteed characterization of exact non-asymptotic confidence regions as defined by LSCR and SPS. *Automatica*, 49, 507–512.

Kolumbán, S., Vajk, I., and Schoukens, J. (2015). Perturbed datasets methods for hypothesis testing and structure of corresponding confidence sets. *Automatica*, 51, 326–331.

Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*. Springer, 2nd edition.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River, 2nd edition.

Pillonetto, G., Chen, T., and Ljung, L. (2013). Kernel-based model order selection for identification and prediction of linear dynamic systems. In *Procs. of the 52nd IEEE Conference on Decision and Control*, 5174–5179.

Stoica, P. and Selen, Y. (2004). Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4), 36–47.