Recursive nuclear norm based subspace identification

Telsang, B.; Navalkar, S. T.; van Wingerden, J. W.

DOI
10.1016/j.ifacol.2017.08.1585

Publication date
2017

Document Version
Final published version

Published in
IFAC-PapersOnLine

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 50-1 (2017) 9490–9495

Recursive Nuclear Norm based Subspace Identification

B. Telsang * **S.T. Navalkar** ** **J.W. van Wingerden** ***
Delft Center for Systems and Control, Delft, The Netherlands

* *e-mail: bhagyashri.telsang17@gmail.com*
** *e-mail: S.T.Navalkar@tudelft.nl*
*** *e-mail: j.w.vanwingerden@tudelft.nl*

Abstract: Nuclear norm based subspace identification methods have recently gained importance due to their ability to find low rank solutions while maintaining accuracy through convex optimization. However, their heavy computational burden typically precludes the use in an online, recursive manner, such as may be required for adaptive control. This paper deals with the formulation of a recursive version of a nuclear norm based subspace identification method with an emphasis on reducing the computational complexity. The developed methodology is analyzed through simulations on Linear Time-Varying (LTV) systems particularly in terms of convergence rate, tracking speed and the accuracy of identification and it is shown to be computationally lighter and effective for such systems, with the considered rate of change of dynamics.

Keywords: Recursive subspace identification, ADMM, Nuclear norm, Initial condition, Warm-start, Stopping criteria

1. INTRODUCTION

System identification plays a crucial role in a specific class of adaptive controllers, where the real-time identification of the underlying system is required (Favoreel, W et al. (1999)). A common approach to system identification in real-time is the class of Prediction-Error Methods (PEM), but these are not easily extendable to MIMO systems. Subspace IDentification (SID) methods mainly emerged as an alternative (Gevers, M (2006)) approach. They typically employ convex cost functions that are amenable to recursive implementations. However, due to the computational complexity associated with high-fidelity SID methods, an online extension for the purpose of adaptive control may not always be tractable in real time.

This issue was recognized in the area of subspace identification soon after the pioneering SID methods (Van Overschee, P and Moor, B (1994), Verhaegen, M and Dewilde, P (1992), Larimore, W. E. (1990)) were published and a number of solutions were proposed (Verhaegen, M and Verdult, V, 2007). For a recursive approach to subspace identification, a subspace tracking algorithm PAST (Yang, B (1995)) from the field of signal processing was introduced in the system identification community for different variants of MOESP (Lovera, M et al. (2000)). Since PAST involves approximations, the estimation results were suboptimal and a recursive solution was proposed in Mercére, G et al. (2004). Although the computational time was reduced, the main drawback of these methods is that they are limited to open-loop identification. This is because of the assumption of zero correlation between noise and the input, which leads to biased estimates when directly applied to closed-loop systems. This issue led to the development of Predictor-Based Subspace IDentification (PBSID) methods. A recursive closed-loop subspace identification method based on an optimized version of PBSID (Chuiso, A (2007)) was proposed in Houtzager, I et al. (2009), which reduced the computational complexity while resulting in unbiased estimates.

Traditional subspace identification methods can be broken down into three distinct steps (Qin, J (2006)): (1) estimation of high-order models, (2) reduction of estimated models to lower dimensional subspace and (3) realization of a state-space system from the lower dimensional subspace. In recent years, a class of SID methods which combine the first two steps of the traditional SID methods have emerged. They do so by embedding the rank minimization criterion (step 2) directly in the identification problem. However, the rank minimization problem, which is in terms of the ℓ_0 norm, is NP-hard. Therefore, a heuristic alternative was developed in Fazel, M et al. (2001), in which it was proved that the nuclear norm – sum of the singular values of a matrix – can be used as a convex envelope of the rank of the matrix. With this heuristic, the problem of rank minimization is relaxed to ℓ_1 norm which has well-established mathematical properties. Employing nuclear norm for rank minimization is attractive primarily because it forms a convex envelope on the rank function, hence rendering the identification problem a convex optimization problem.

Most of the nuclear norm based SID methods (e.g., Verhaegen, M and Hansson, A (2015) and Hansson, A et al. (2012)) formulate the problem in two terms: one term to reduce the model order and the other to obtain an accurate description of the system. For a given number of measurements, it was proved in Candés, E. J and Tao, T (2009) that a low rank solution describing the underlying

10.1016/j.ifacol.2017.08.1585

system can be retrieved by solving such an optimization problem.

Several methods have been proposed to solve an optimization problem based on the nuclear norm: e.g, Liu, Z and Vandenberghe, L (2009), Ayazoglu, M and Sznaier, M (2012), Cai, J. F et al. (2008), Ji, S and Ye, J (2009), Mohan, K and Fazel, M (2012). In this paper, the Alternating Direction Method of Multipliers (ADMM) algorithm developed in Boyd, S et al. (2011) will be employed. Since the nuclear norm operator is not differentiable with respect to its arguments, ADMM is particularly favorable in our case because it does not require the optimization function to be differentiable; sub-differentials of the function can be readily used. Moreover, the computational complexity of ADMM is lighter than Semi-Definite Programming solvers which were employed in for example, Verhaegen, M and Hansson, A (2014).

Although much progress has been made in the direction of recursive subspace identification methods, nuclear norm based SID methods still remain computationally heavy, rendering them to be less useful in the design of an online adaptive controller. While a recursive algorithm that formulates the problem based on the nuclear norm variant of PBSID is developed in Navalkar, S.T and van Wingerden, J.W (2016), the formulation developed here is based on the nuclear norm variant of SID methods. Accordingly, this paper aims to formulate a recursive version of a nuclear norm based SID method, called N2SID developed in Verhaegen, M and Hansson, A (2015), thereby facilitating its use in adaptive control. The non-recursive algorithm is briefly summarized in Section 2 along with the optimization algorithm ADMM that is used to solve it. The contributions made in this paper are three fold. In order to facilitate recursive identification, the conditions for a warm-start are developed in Section 3.1. Secondly, the characteristics of ADMM are exploited in Section 3.2 to redefine the stopping criteria with an aim to improve the speed. The two improvements are combined together to result in a recursive identification method. The effects of these are studied in Section 4 by performing recursive identification on linear time-varying systems. Finally, the paper is ended with some concluding remarks.

2. BACKGROUND

In this section, the N2SID method, presented in Verhaegen, M and Hansson, A (2015) is recapitulated, so that a recursive implementation can be formulated in the next section. Consider that the system to be identified can be realised (instantaneously) as an LTI state space model in the innovation form:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\ y(k) &= Cx(k) + Du(k) + e(k) \end{aligned} \tag{1}$$

where $u(k) \in \mathbb{R}^{m_u}, x(k) \in \mathbb{R}^n, y(k) \in \mathbb{R}^{p_y}$ and $e(k) \in \mathbb{R}^{p_y}$ is the zero mean innovation sequence. Let (1) be compactly expressed in the observer form as:

$$\begin{aligned} x(k+1) &= \bar{A}x(k) + \bar{B}u(k) + Ky(k) \\ y(k) &= Cx(k) + Du(k) + e(k) \end{aligned} \tag{2}$$

with $\bar{A} = (A - KC)$ and $\bar{B} = (B - KD)$.

The system identification problem is to approximate the system matrices for the system described in (1). This is solved using the N2SID method (Verhaegen, M and Hansson, A (2015)) which formulates the problem as:

$$\min_{\hat{Y}_s, T_{u,s}, T_{y,s}} \; ||\hat{Y}_s - T_{u,s}U_s - T_{y,s}Y_s||_\star + \frac{\lambda}{N}\sum_{k=1}^{N} ||y(k) - \hat{y}(k)||_2^2 \tag{3}$$

with $||\ ||_\star$ being the nuclear norm operator and λ being the regularization parameter. Through the tuning of λ, the trade-off between model order reduction and estimation accuracy is quantified. Here, $y(k)$ is the measured output of the system and $\hat{y}(k)$ is the estimated output. s is the number of block rows and N is the number of measurements or samples used in the identification dataset. $U_s \in \mathbb{R}^{m \times q}$, where $m = sm_u$ and $q = (N - s + 1)$, is the Hankel matrix formulation of the input $u(k)$ (Verhaegen, M and Verdult, V (2007)):

$$U_s = \begin{bmatrix} u(1) & u(2) & \dots & u(N-s+1) \\ u(2) & u(3) & & \vdots \\ \vdots & & \ddots & \\ u(s) & u(s+1) & \dots & u(N) \end{bmatrix} \tag{4}$$

Similarly, the Hankel matrices for the measured output $y(k)$ and the estimated output $\hat{y}(k)$ are formulated as $Y_s \in \mathbb{R}^{p \times q}$ and $\hat{Y}_s \in \mathbb{R}^{p \times q}$ respectively, where $p = sp_y$. The variables $T_{u,s}$ and $T_{y,s}$ are Toeplitz matrices containing the system matrices (Verhaegen, M and Hansson, A (2015)) from model (2).

The optimization problem formulated in (3), which is a convex relaxed problem, can be solved using the ADMM algorithm demonstrated in Hansson, A et al. (2012). We concisely review ADMM in order to understand the theory developed in Section 3. For a detailed explanation of ADMM, readers are referred to Boyd, S et al. (2011). The optimization variables from (3) are split into two (primal) variables: $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$ (with $n_\mathbf{x} = p_y(N + m + (s-1)p_y)$) and $\mathbf{X} \in \mathbb{R}^{p \times q}$. They contain the optimization variables $T_{u,s}, T_{y,s}$ and $\hat{Y}_s$ of problem defined in (3). Accordingly, the problem is reformulated as:

$$\begin{aligned} \min_{\mathbf{x},\mathrm{X}} \quad & f(\mathbf{x}) + g(\mathrm{X}) \\ \text{subject to} \quad & \mathcal{A}(\mathbf{x}) - \mathrm{X} = \mathcal{B} \end{aligned} \tag{5}$$

Since ADMM is a primal-dual algorithm (Komodakis, N and Pesquet, J (2015)), it employs a dual variable Z along with the primal variables $\mathbf{x}$ and $\mathbf{X}$ to solve the problem defined in equation (5). ADMM maximizes a dual function with respect to the dual variable; the dual function is defined as:

$$g_{t_p}(Z) = \inf_{\mathbf{x},X} \; L_{t_p}(\mathbf{x}, \mathrm{X}, Z) \tag{6}$$

The function $L_{t_p}(\mathbf{x}, \mathrm{X}, Z)$ here is the Augmented Lagrangian defined as:

$$L_{t_p} = f(\mathbf{x}) + g(\mathrm{X}) + \mathrm{trace}(Z^T(\mathcal{A}(\mathbf{x}) - \mathrm{X} - \mathcal{B})) + \frac{t_p}{2}||\mathcal{A}(\mathbf{x}) - \mathrm{X} - \mathcal{B}||_F^2$$

where t_p is the penalty parameter. Each iteration proceeds by maximizing the dual function, which is concave, and finds the minimum of the primal variables for each dual update Z. The stopping criteria, which is defined in terms of primal and dual residual norms (r_p and r_d respectively) and primal and dual tolerances (ϵ_p and ϵ_d respectively) is defined as:

$$r_p \leq \epsilon_p, \quad r_d \leq \epsilon_d \tag{7}$$

Once the stopping criteria has been reached, the ADMM iterates satisfy residual, objective and dual variable convergence. Note that primal variables need not converge to optimal values (Boyd, S et al. (2011)). Using the variables resulting from the convergence of ADMM, the matrices that describe the system in (1) can be extracted from $T_{u,s}$ and $T_{y,s}$ (Verhaegen, M and Hansson, A (2015)), hence solving the system identification problem.

The aim of this paper is to develop a recursive version of the identification method reviewed in this section. Accordingly we will consider, in the next section, some conditions that facilitate such identification.

3. RECURSIVE IDENTIFICATION

Recursive identification can be considered to be a repetitive application of the batchwise identification process described in the previous section. Such an implementation would be of specific interest for adaptive control.

At a particular time instant, a system is identified using its past N measurement values as described in Section 2. Let us call it the i^{th} identification cycle and accordingly denote the i^{th} identification problem, which is equation (3), as $\mathcal{Q}^i$:

$$\mathcal{Q}^i : \min_{\hat{Y}_s^i, T_{u,s}^i, T_{y,s}^i} \quad \frac{\lambda}{N}\sum_{k=i-N+1}^{i}||y^i(k) - \hat{y}^i(k)||_2^2 + ||\hat{Y}_s^i - T_{u,s}^i U_s^i - T_{y,s}^i Y_s^i||_\star \tag{8}$$

The corresponding optimization problem that is solved using ADMM is $\mathcal{P}^i$ which is defined as:

$$\mathcal{P}^i : \quad \min_{\mathbf{x}, X} \quad f^i(\mathbf{x}) + g^i(\mathrm{X}) \quad \text{subject to} \quad \mathcal{A}^i(\mathbf{x}) - \mathrm{X} = \mathcal{B} \tag{9}$$

The variables in (9) are $\mathbf{x}$, X and Z, and the resulting values after the algorithm converges are denoted with superscript i. Accordingly, the solution of $\mathcal{P}^i$ yields the optimal dual variable, denoted Z^i, and the primal variables $\mathbf{x}^i$ and X^i such that the objective has converged to its optimal.

We now consider the problem of solving the $(i+1)^{th}$ identification with a lesser computational burden and an improved speed, for an online implementation. As shown in the next section, such a speed-up can be provided with a warm start by using as the initial condition the results of the i^{th} cycle instead of the default zero initial matrices for primal and dual variables of the ADMM. To further improve the speed, we next revisit and modify the stopping criteria.

3.1 Choice of initial condition

Knowledge of the previous optimum can be exploited in speeding up the optimization process at each time instant that a system identification estimate is required. Apart from λ and the optimization variables in (8), the definition of the problem $\mathcal{Q}^i$ reveals that it is entirely made up of the past N measurement values – from time instant $(i-N+1)$ to i. Since the $(i+1)^{th}$ identification is performed immediately after the i^{th} identification, the problem $\mathcal{Q}^{i+1}$ comprises of the measurements from $(i-N+2)$ to $(i+1)$. Thus, the problems $\mathcal{Q}^i$ and $\mathcal{Q}^{i+1}$ differ by just one measurement value. This is reflected in two ways: (1) there is a shift structure in U_s and Y_s and (2) by the proximity of the dual functions of $\mathcal{P}^i$ and $\mathcal{P}^{i+1}$. In this paper, we focus on the latter (exploitation of the shift structure in the Hankel matrices is a part of future work). The (concave) dual functions of $\mathcal{P}^i$ and $\mathcal{P}^{i+1}$ are such that they are less than $\epsilon > 0$ apart:

$$|g_{t_p}^{i+1}(Z) - g_{t_p}^{i}(Z)| < \epsilon \tag{10}$$

From the solution of $\mathcal{Q}^i$, the optimal dual variable (Z^i) is available. Since the aim is to solve the $(i+1)^{th}$ identification, we now try to gain insight about Z^{i+1} using the knowledge of Z^i and using (10). First, however, a useful result is stated in the following lemma.

Lemma 1. Consider two convex functions $h_1 : \mathbb{R}^{p\times q} \to \mathbb{R}$ and $h_2 : \mathbb{R}^{p\times q} \to \mathbb{R}$ such that

$$|h_1(x) - h_2(x)| < \epsilon \qquad \forall x \in \mathbb{R}^{p\times q} \tag{11}$$

Let

$$x_1^* = \underset{x\in\mathbb{R}^{p\times q}}{\operatorname{argmin}} \; h_1(x)$$
$$x_2^* = \underset{x\in\mathbb{R}^{p\times q}}{\operatorname{argmin}} \; h_2(x)$$

$$\mathcal{M} := \{x \in \mathbb{R}^{p\times q} : h_1(x) - \epsilon < h_1(x_1^*) + \epsilon\}, \qquad \mathcal{M} \subset \mathbb{R}^{p\times q}$$

Then we have the following:

(1) $x_1^* \in \mathcal{M}$
(2) $x_2^* \in \mathcal{M}$

Proof. The first part of the lemma follows trivially from

$$(h_1(x_1^*) - \epsilon) < (h_1(x_1^*) + \epsilon)$$

We now proceed to prove $x_2^* \in \mathcal{M}$. From equation (11),

$$h_2(x) < (h_1(x) + \epsilon)$$

So,

$$h_2(x_1^*) < (h_1(x_1^*) + \epsilon) \tag{12}$$

Consider $\tilde{x} \in \mathbb{R}^{p\times q}$ such that

$$h_2(\tilde{x}) \;\geq\; (h_1(x_1^*) + \epsilon)$$

From equation (12),

$$h_2(\tilde{x}) \;>\; h_2(x_1^*)$$

$\implies \tilde{x}$ is not x_2^*

$\implies x_2^*$ must satisfy the condition

$$h_2(x_2^*) \;<\; (h_1(x_1^*) + \epsilon) \tag{13}$$

From equation (11) we have,

$$(h_1(x) - \epsilon) \;<\; h_2(x), \quad \forall x \in \mathbb{R}^{p\times q} \tag{14}$$

Therefore, from equation (13) and equation (14), we have:

$$x_2^* \in \mathcal{M}$$

Application of Lemma 1 to the concave functions $g_{t_p}^i : \mathbb{R}^{p\times q} \to \mathbb{R}$ and $g_{t_p}^{i+1} : \mathbb{R}^{p\times q} \to \mathbb{R}$, reveals that Z^i – the optimal dual variable solution of the problem $\mathcal{P}^i$ – and Z^{i+1} belong to the same set. Thus, it immediately follows that Z^i is a suitable choice of initial condition for the maximization of the function $g_{t_p}^{i+1}$.

Furthermore, we know that ADMM proceeds by maximizing the dual function, and along with it, finding the minimizers of the primal variables in each iteration. Therefore, from application of Lemma 1, since $\mathcal{P}^{i+1}$ is convex, $\mathbf{x}^i$ and X^i can be used as the initial condition for solving $\mathcal{P}^{i+1}$. Thus, with $\mathbf{x}^i$, X^i and Z^i as the initial condition for solving the problem $\mathcal{P}^{i+1}$, we start in the same set which contains Z^{i+1}.

With the choice of initializing the optimization variables in ADMM using a so-called 'warm-start', it is thus possible to improve the rate of convergence of a recursive online identification strategy. However, since there is no theory to comment on the convergence of ADMM using zero initial condition, it should be noted that the convergence with the developed choice of initial condition cannot be guaranteed to be better than the default zero initial condition. In the next section, we consider the modification of the stopping criteria in ADMM to further improve the convergence characteristics.

3.2 Modification of the stopping criteria

It was reviewed in Section 2 that convergence of ADMM is considered to be achieved when the stopping criteria, which depends on primal and dual residual norms and tolerances, given in (7) are reached. It is known that the ADMM algorithm is characterized by a very slow convergence rate for very high accuracy (Boyd, S et al. (2011)). This can be seen in Fig. 1 which shows the behavior of primal and dual residual norms and tolerances as the iterations of ADMM progress for an identification. The values in this demonstrative example correspond to identification (using N2SID as in Verhaegen, M and Hansson, A (2015)) of a CD player arm system, whose identification dataset is available at Moor, B (2012).

If the application does not require extremely high accuracy then the ADMM algorithm can be terminated early i.e., before the stopping criteria as in equation (7) has been reached. Exploiting the fact that most of the convergence in ADMM occurs in the first few iterations, the stopping criteria is restated in terms of the derivatives of residual norms. Accordingly, ADMM iterations are terminated when the derivatives of the primal and dual residual norms are lesser than a specified tolerance. The modified stopping criteria is thus given by:

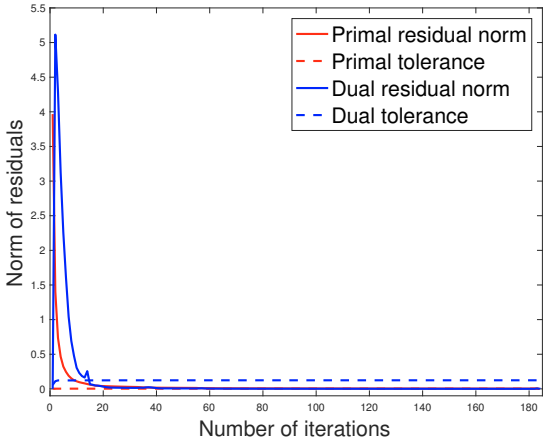

Figure 1. Convergence analysis of ADMM through behavior of residual norms and tolerances with progressing iterations

$$\frac{dr_p}{dk_a} \leq tol_p, \quad \frac{dr_d}{dk_a} \leq tol_q \tag{15}$$

where k_a is the iteration number in ADMM. The idea behind this modification is to improve the speed while maintaining the accuracy of identification.

In the next section, we will study the identification behavior using two examples to evaluate the effectiveness of the formulated stopping criteria and present the results of providing a warm start for identification.

4. RESULTS

In this section, the recursive identification performed using the theory that is outlined in Sections 2 and 3 is tested in simulations on Linear Time-Varying (LTV) systems. In the first example, the emphasis lies on analyzing the practical implications of the theory presented in Section 3. This is carried out by comparing the results of identification for warm-start and for the modified stopping criteria with those obtained from N2SID as reviewed in Section 2. The two modifications are then combined and the resulting method (which we call Recursive N2SID) is compared with the other cases. In the second example, RN2SID is compared with a Recursive PBSID method developed in Houtzager, I et al. (2009).

Outline of RN2SID: Identification is recursively performed at every discrete timestep, i.e., with every new measurement sample, over T timesteps. Each identification is initialized with the results of the previous identification (as presented in Section 3.1) and terminated using the revised stopping criteria given in Section 3.2. Further, each identification is performed using past N measurement

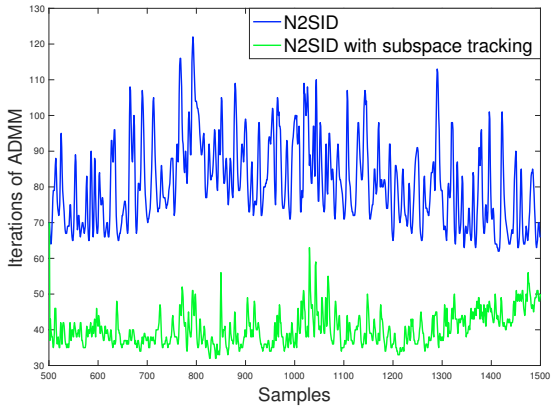

Figure 2. Number of iterations required for convergence

values. In order to maintain constant size of the problem, for every latest measurement value that is appended the oldest value is discarded. Hence the identification dataset with fixed size N is sliding over T samples.

4.1 Example 1

To evaluate the theory presented so far, we perform system identification using benchmark dataset provided by Lataire, J et al. (2015). The system described therein is varying, whose dynamics are influenced by an external input referred to as the scheduling signal "p". The system has another input u and an output y. The scheduling signal is considered to be unknown and is not included in the identification dataset.

The LTV system is recursively identified over $T = 1000$ discrete timesteps, with $N = 500$, for the dataset `MS_Ramp_N15640_RMS140_P2P700` which is obtained by a ramp variation of the scheduling signal p. Since p is constantly changing over time, the system dynamics vary with time. The number of iterations required for convergence of ADMM in each identification cycle is shown for both N2SID and for N2SID with warm-start in Fig. 2. Since with the presented choice of initial condition we start in the same set as the optima, early convergence is expected. Accordingly, the number of iterations required for convergence is less than that required with the traditional N2SID algorithm with cold-start.

The results of each modification are summarized in Table 1 along with the results for RN2SID that is outlined early in this Section. The values tabulated correspond to one identification and they are obtained by averaging the results over T samples. It is worth noting that the time taken for one cycle has reduced by almost 70% from N2SID to RN2SID, while maintaining roughly the same accuracy.

Table 1. Summary of recursive identification

	VAF(%)	Computation time (s)	Iterations
N2SID	90.45	1.59	81
With warm-start	90.39	0.89	45
Modified convergence	90.24	0.82	40
RN2SID	90.23	0.48	24

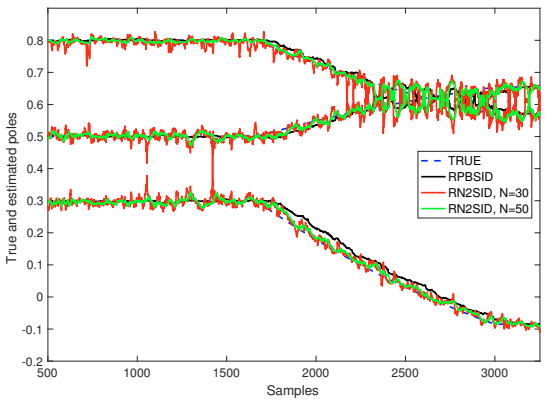

Figure 3. Trajectories of true and estimated poles in slowly changing environment for $\lambda = 100$

4.2 Example 2

The tracking speed of RN2SID is analyzed by performing recursive identification on the state-space system from Mercére, G et al. (2008). The loop is closed with a state-feedback control law and the resulting closed-loop measurement values is used for identification. The system is the same as considered in Houtzager, I et al. (2009), the method of which RN2SID will be compared with. Two cases are considered to evaluate the tracking performance: a slowly changing environment, and an abrupt-change case. The analysis is carried out for different values of the tuning parameters N and λ.

In the case of slowly changing environment RN2SID is analyzed by fixing λ to 100 and varying N, as shown in Fig. 3. The forgetting factor employed in RPBSID is 0.98 which roughly corresponds to the length of identification dataset being 50. For increase in the length of identification dataset (N), the variance of the estimates decreases but the bias increases. Comparison with RPBSID shows that the bias of the estimates obtained using RN2SID is lesser than that of RPBSID. Moreover, the convergence of RPBSID takes more samples than that required by RN2SID. However, the computational time taken by RPBSID is considerably lesser than RN2SID. On the other hand, we see that in RN2SID there is a faster detection of system dynamics, even with the reduced $N = 30$, but at the price of higher variance. However, the behavior of RPBSID corresponding to $N = 30$ (not plotted) is almost the same as with $N = 50$.

Analysis of RN2SID is extended to the system in an abruptly-varying environment by varying the regularization parameter λ for $N = 50$ shown in Fig. 4. Since nuclear norm minimization is a form of regularization, it induces a bias in the estimates. With decrease in weighting on the nuclear norm term (corresponding to increase in λ), there is a distinct reduction in the bias of the estimates. The observed behavior intuitively reflects the term λ/N in (3).

5. CONCLUSION

A novel, recursive form of N2SID was developed in order to facilitate the extension of this algorithm to online adaptive control. The convergence speed of the algorithm was enhanced by enforcing a warm-start in ADMM and

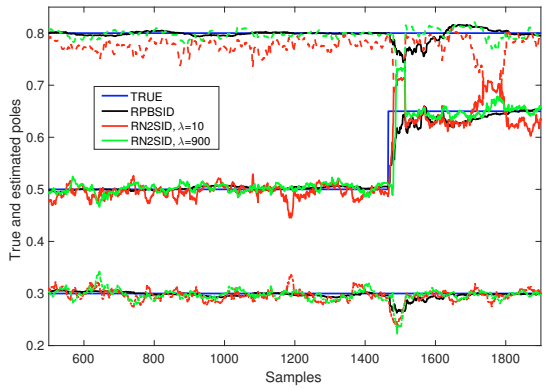

Figure 4. Trajectories of true and estimated poles in an abrupt time-varying environment for $N = 50$

modifying the convergence conditions. From the examples presented, the computational time decreased by up to 70% with negligible loss in accuracy. The tracking speed of the developed methodology was comparable to that of the state-of-the-art RPBSID in simulation. Future work will focus on exploiting the shift structure in the Hankel matrices U_s and Y_s to further reduce the computational complexity.

REFERENCES

Ayazoglu, M and Sznaier, M (2012). An algorithm for fast constrained nuclear norm minimization and applications to systems identification. *IEEE Decision and Control*, 3469–3475.

Boyd, S, Parikh, N, Chu, E, Peleato, B, and Eckstein, J (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 1, 1–122.

Cai, J. F, Candés, E, and Shen, Z (2008). A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, 20.

Candés, E. J and Tao, T (2009). The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 2053–2080.

Chuiso, A (2007). The role of vector auto regressive modeling in predictor-based subspace identification. *Automatica*, 43, 1034–1048.

Favoreel, W, Moor, B, and Van Overschee, P (1999). Model-free subspace-based lqg-design. *Proceedings of ACC*, 5, 3372–3376.

Fazel, M, Hindi, H, and Boyd, S (2001). A rank minimization heuristic with application to minimum order system approximation. *Proceedings of American Control Conference*, 4734–473.

Gevers, M (2006). A personal view on the development of system identification. *IEEE Control Systems*, 26, 93–105.

Hansson, A, Liu, Z, and Vandenberghe, L (2012). Subspace system identification via weighted nuclear norm optimization. *IEEE Decision and Control*, 3439–3444.

Houtzager, I, van Wingerden, J.W, and M. Verhaegen (2009). Fast-array recursive closed-loop subspace model identification. *IFAC Symposium on System Identification*.

Ji, S and Ye, J (2009). An accelerated gradient method for trace norm minimization. *Proceedings of the 26th International Conference on Machine Learning*.

Komodakis, N and Pesquet, J (2015). Playing with Duality: An Overview of Recent Primal-Dual Approaches for Solving Large-Scale Optimization Problems. *IEEE Signal Processing Magazine*, 32.

Larimore, W. E. (1990). Canonical variate analysis in identification, filtering and adaptive control. *Proceedings of 29th Conference on Decision and Control*, 596–604.

Lataire, J, Louarroudi, E, Pintelon, R, and Rolain, Y (2015). Benchmark data on a linear time- and parameter-varying system. *17th IFAC Symposium on System Identification*.

Liu, Z and Vandenberghe, L (2009). Semidefinite programming methods for system realization and identification. *Proceedings of the Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, 4676–4681.

Lovera, M, Gustafsson, T, and Verhaegen, M (2000). Recursive subspace identification of linear and nonlinear wiener state-space models. *Automatica*, 36.

Mercére, G, Bako, L, and Lecoeuche, S (2008). Propagator-based methods for recursive subspace model identification. *Signal Processing*, 88, 468–491.

Mercére, G, Lecoeuche, S, and Lovera, M (2004). Recursive subspace identification based on instrumental variable unconstrained quadratic optimization. *International Journal of Adaptive Control and Signal Processing*, 18, 771–797.

Mohan, K and Fazel, M (2012). Iterative reweighted algorithms for matrix rank minimization. *Journal of Machine Learning Research*, 3441–3473.

Moor, B (2012). DaISy: Database for the Identification of Systems. *http://homes.esat.kuleuven.be/ smc/daisy/*.

Navalkar, S.T and van Wingerden, J.W (2016). Nuclear norm-enhanced recursive subspace identification: Closed-loop estimation of rapid variations in system dynamics. *Proceedings of ACC*.

Qin, J (2006). An overview of subspace identification. *Computers and chemical engineering*, 1502–1513.

Van Overschee, P and Moor, B (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30, 75–93.

Verhaegen, M and Dewilde, P (1992). Subspace model identification part I: The output-error state space model identification class of algorithms. *International Journal of Control*, 56, 1187–1210.

Verhaegen, M and Hansson, A (2014). Nuclear Norm Subspace Identification (N2SID) for short data batches. *IFAC World Congress*.

Verhaegen, M and Hansson, A (2015). N2SID: Nuclear norm subspace identification. *CoRR, http://arxiv.org/abs/1401.4273*.

Verhaegen, M and Verdult, V (2007). *Filtering and System Identification*. Cambridge University Press, Cambridge CB2 8RU, UK.

Yang, B (1995). Projection approximation subspace tracking. *IEEE Transactions on Signal Processing*, 43, 95–107.