2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Identifying Stable Fixed Order Systems from Time and Frequency Response Data

C. Feng*, C. M. Lagoa* and M. Sznaier**

Abstract— **In this paper, we address the problem of identifying fixed order stable SISO systems from time and frequency domain data. Given measurements of time and frequency response corrupted by process and measurement noise, we aim at finding a fixed order stable plant whose response matches the data collected (within the noise bounds) and whose $\mathcal{H}_\infty$ norm is below a prescribed level. It is shown the problem can be solved by finding a point in a set defined by polynomial inequalities and the sparse structure of the polynomials is exploited to develop an efficient system identification algorithm. Further computational improvements are obtained by reformulating the problem as a rank constrained one and using efficient convex relaxations of rank minimization. Numerical examples are provided to illustrate the efficiency of the proposed algorithms.**

I. INTRODUCTION

In this paper, we address the problem of fixed order system identification of single input single output (SISO) systems from both time and frequency data. More precisely, given noisy measurements in the time and frequency domain, we aim at determining a *stable fixed order* system which is consistent to the data within the noise bounds and whose $\mathcal{H}_\infty$ norm is below a prescribed bound. This is accomplished by first reformulating the problem as finding a point in a properly defined semi-algebraic set and exploiting its inherent structure to develop efficient convex relaxations. Finally, to further improve computational efficiency, an equivalent rank constrained problem is developed and known convex relaxations of rank minimization are used to provide a faster, more efficient identification algorithm.

The problem of system identification has been thoroughly studied in the literature. Many of the classical results in these area can be found in many books/notes such as [15], [21]. Moreover, results on how to determine models from time domain data, frequency domain data or both can be found in, e.g., [9], [15], [17] and references therein. Since identification procedure can be only preformed on stable systems in practice, it is desirable to take stability into account in system identification. In [18], it is proven that an autoregressive constraint on the input guarantees stability of an identified transfer function via least-squares (LS) method. When subspace identification is concerned, several conservative methods have been proposed in [3], [16], [20], either by adjusting extended observability matrix or by adding regularization terms to the least squares cost function. In [11], sufficient conditions posed on Lyapunov parameters have been used in subspace identification to ensure stability. In [19], an iterative approach is proposed by incrementally adding constraints to improve stability until a stable solution is found. More recently, in [2], an approach based on Jury's test and polynomial optimization is proposed to compute the upper/lower bound of compatible system parameters, when time-domain data is available; see also [1] for a similar polynomial approach to set-membership error-In-variables identification.

A. Contribution and Organization

In this paper, we also take a polynomial optimization based approach to the problem of stable fixed order system identification. Moreover, we go beyond results available in the literature and consider a general case where 1) both time and frequency domain data are available, 2) both process noise and measurement noise are considered and 3) constraint on the $\mathcal{H}_\infty$ norm of the identified system is imposed. Efficient numerical algorithms are provided to solve the resulting polynomial optimization problem by exploiting its intrinsic sparse structure and using results on rank minimization.

The rest of the paper is organized as follows. In Section II, notation and basic results in polynomial optimization are introduced. The identification problem is formally defined in Section III. In Section IV, the problem is reformulated as an equivalent feasibility problem involving semi-algebraic sets. Semi-definite program (SDP) relaxations are provided with the consideration of sparsity. To further enhance computational efficiency, an equivalent rank minimization problem is proposed in Section V. Numerical examples are provided in Section VI to illustrate the efficacy of the proposed method.

II. PRELIMINARIES

In this section we define the notation used and briefly summarize some results on polynomial optimization used in this paper. For a more detailed exposition, the reader is referred to [4], [12], [13], [14].

A. Notation

x^i	abbreviation for $x_i^{i_1} \cdots x_d^{i_d}$ where d is the dimension of the vector x
$\mathbf{E}_\mu[p(x)]$	the mean value of $p(x)$ w.r.t the probability measure μ on the random variable x
$M \succeq 0$	the matrix M is positive semi-definite

This work was supported by the National Science Foundation under grants CMMI-0838906, ECCS-0731224, ECCS-0501166, IIS0713003, ECCS-0901433; and AFOSR grant FA9550-09-1-0253.

*Department of Electrical Engineering, The Pennsylvania State University, University Park, PA 16802, USA (e-mail: feng@psu.edu, lagoa@engr.psu.edu).

**ECE Department, Northeastern University, Boston, MA 02115, USA (m.sznaier@neu.edu)

B. General Polynomial Optimization

Consider the following constrained polynomial optimization problem:

$$p_K^* := \min_{x \in K} p_0(x) \tag{P1}$$

where $K \subset \mathbb{R}^\ell$ is a compact semi-algebraic set with nonempty interior defined as

$$K \doteq \{x \colon p_i(x) \geq 0, i = 1, \dots d\}$$

where $p_i(x)$ are polynomials with total degree d_i. Consider a related problem in the probability measure space:

$$\tilde{p}_K^* := \min_{\mu \in \mathcal{P}(K)} \int p_0(x)\mu(dx) := \min_{\mu \in \mathcal{P}(K)} \mathbf{E}_\mu [p_0(x)] \tag{P2}$$

where $\mathcal{P}(K)$ is the space of finite Borel probability measures on K. According to [12], these two problems are equivalent. One direct consequence of this equivalency is that, it is possible to develop a convergent sequence of LMI based convex relaxations to problem (P1), where the optimization variables are $m_i \doteq \mathbf{E}_\mu x^i$, the moments of the unknown distribution μ. To this effect, let

$$\begin{aligned} p_N^* = \quad & \min_m \quad \sum_\alpha p_{0,\alpha} m_\alpha \\ & \text{s.t.} \quad M_N(m) \succeq 0, \\ & \qquad\;\; M_{N_i}(p_i m) \succeq 0, i = 1, \dots, d, \end{aligned} \tag{1}$$

where N is the relaxation order; $p_{0,\alpha}$ is the coefficient of x^α in $p_0(x)$; N_i is the smallest integer that no less than $N - d_i/2$; $M_N(m)$ is the so-called *moment matrix* and $M_{N_i}(p_i m)$ is the so-called *localizing matrix*. For illustration and clarity of exposition, consider the case where $x \in R^2$, the moment matrix $M_N(m)$ consists of the block matrix $\{M_{j,k}(m)\}_{0 \leq j,k \leq N}$ is defined as

$$M_N(m) = \begin{bmatrix} M_{0,0}(m) & M_{0,1}(m) & \cdots & M_{0,N}(m) \\ M_{1,0}(m) & M_{1,1}(m) & \cdots & M_{1,N}(m) \\ \vdots & \vdots & \ddots & \vdots \\ M_{N,0}(m) & M_{N,1}(m) & \cdots & M_{N,N}(m) \end{bmatrix}$$

where

$$M_{j,k}(m) = \begin{bmatrix} m_{j+k,0} & m_{j+k-1,1} & \cdots & m_{j,k} \\ m_{j+k-1,1} & m_{j+k-2,2} & \cdots & m_{j-1,k+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_{k,j} & m_{k-1,j+1} & \cdots & m_{0,j+k} \end{bmatrix}.$$

The localizing matrix $M_{N_i}(p_i, m)$ is defined as

$$M_{N_i}(p_i, m)(i,j) = \sum_\alpha p_{i,\alpha} m(\beta(i,j) + \alpha)$$

where $p_{i,\alpha}$ is the coefficient of x^α in $p_i(x)$, $m(i,j)$ is the entry (i,j) of $M_N(m)$ and $\beta(i,j)$ is the subscript of m_β. By the end, according to Theorem 4.2 in [12], we have,

Theorem 1 (General Polynomial Optimization): Under "mild" conditions,

$$p_N^* \uparrow p_K^*. \tag{2}$$

as N increases to infinity,

C. Sparse Polynomial Optimization

If the set of variables in the polynomials satisfies the so-called *running intersection property*, the size of the moment/localizing matrices can be significantly reduced in the resulting SDP relaxations.

Definition 1 (Running Intersection Property [13]): Let I_k, $k = 1, \dots, r$, be the subsets of variables $X \doteq \{x_1, \dots, x_\ell\}$ satisfying $\bigcup_{k=1}^r I_k = X$. If

i) *each constraint polynomial $p_i(x)$ uses only variables in I_k for some $k = \beta(i)$;*
ii) *the objective polynomial can be written as $p_0 = p_{0,1} + \cdots + p_{0,t}$ where each $p_{0,i}$ uses only variables in I_k for some k,*

then running intersection property *is satisfied in* (P1) *if the collection $\{I_1, \dots, I_r\}$ obeys:*

$$I_{k+1} \cap \bigcup_{j=1}^k I_j \subseteq I_s \text{ for some } s \leq k, \tag{3}$$

for every $k = 1, \dots, r-1$.

Similar to the result stated in Theorem 1, for a sparse polynomial optimization problem that satisfies the *running intersection property*, the convergence property holds as well. For simplicity, we denote $M_N(m, I_k)$ the moment matrix for the reduced variables in the set I_k. Similarly, we denote $M_{N_i}(p_j m, I_k)$ the localizing matrix with the reduced variables in I_k. We now restate Theorem 3.6 in [13] that formalize this approach.

Theorem 2 (Sparse Polynomial Optimization): Assume that (P1) *satisfies running intersection property. Let*

$$\begin{aligned} p_N^* = \quad & \min_m \quad \sum_\alpha p_\alpha m_\alpha \\ & \textit{s.t.} \quad M_N(m, I_k) \succeq 0, \quad k = 1, \dots, r \\ & \qquad\;\; M_{N_i}(p_i m, I_{\beta(i)}) \succeq 0, \quad i = 1, \dots, d. \end{aligned} \tag{4}$$

Then, as N increases to infinity,

$$p_N^* \uparrow p_K^*. \tag{5}$$

D. Positive Trigonometric Polynomials on Unit Circle

Consider a univariate trigonometric polynomial

$$R(z) = \sum_{k=-d}^{d} r_k z^k, r_{-k} = r_k,$$

it is non-negative on the unit circle if and only if it can be represented as

$$R(z) = [1, z^{-1}, \dots, z^{-d}] Q [1, z, \dots, z^d]^T,$$

where Q is a positive semi-definite matrix; e.g., see [4]. This is summarized as follows.

Theorem 3 (Theorem 2.5 in [4]): The univariate polynomial $R(z)$ is non-negative on the unit circle if and only if there exist a positive semi-definite matrix

$$Q \doteq \begin{pmatrix} q_{0,0} & \cdots & q_{0,d_r} \\ \vdots & \ddots & \vdots \\ q_{d_r,0} & \cdots & q_{d_r,d_r} \end{pmatrix}$$

such that

$$r_k = \sum_{i=k}^{d_r+k} q_{i,i-k}, k = 0, \dots, d_r. \quad (6)$$

III. Problem Formulation

In this paper, we consider SISO systems of the form

$$G(z) = \frac{b(z)}{a(z)}, \quad (7)$$

where

$$\begin{aligned} a(z) &= 1 + a_1 z^{-1} + \ldots + a_n z^{-n}, & (8)\\ b(z) &= b_1 z^{-1} + \ldots + b_n z^{-m}. & (9) \end{aligned}$$

It is assumed that one only has accesses to noisy measurements and that the input is perturbed by noise, i.e., given measurements y_k and input u_k, the following holds,

$$y_k = -\sum_{i=1}^{n} a_i (y_{k-i} - v_{k-i}) + \sum_{i=1}^{m} b_i (u_{k-i} + w_{k-i}) + v_k + e_k,$$

where w_k, v_k, e_k are input noise, output noise and process noise, respectively; a and b are system parameters; $u_k = 0, y_k = 0$ for $k \le 0$. The noises are assumed to be bounded in ℓ_∞ norm, i.e., there exist known $\bar{w}$, $\bar{v}$ and $\bar{e}$ so that

$$|w_k| \le \bar{w}, |v_k| \le \bar{v}, |e_k| \le \bar{e}, \quad k = 1, \ldots, L_t. \quad (10)$$

It is also assumed that noisy measurements of the frequency response are available. More precisely,

$$\hat{G}_k = G(e^{-j\omega_k}) + \epsilon_k, \quad k = 1, \ldots, L_f, \quad (11)$$

are the (noisy) measurements of frequency response. Moreover, ϵ_k is the frequency measurement noise assumed to be bounded, i.e., $|\epsilon_k| \le \bar{\epsilon}$ for some known constant $\bar{\epsilon}$. We assume that the system $G(z)$ is stable (exponentially stable), i.e., the roots of $a(z)$ lie in the unit disk (a disk with radius $\rho < 1$) and that the $\mathcal{H}_\infty$ norm of $G(z)$ is bounded by a known positive real number $C_\mathcal{H}$, i.e.,

$$\|G(z)\|_\infty \le C_\mathcal{H}. \quad (12)$$

The objective is to identify a compatible model, if any, that is consistent with all a priori information and all measurement data, i.e., we aim at solving the following problem.

Problem 1: Given fixed order (n, m)*, time-domain data* (u_k, y_k)*,* $1 \le k \le L_t$*, frequency-domain data* $(\omega_k, \hat{G}(e^{-j\omega_k}))$*,* $1 \le k \le L_f$*, bounds* ($\bar{w}$, $\bar{v}$, $\bar{e}$) *on the* ℓ_∞ *norm of measurement noise* (w, v, e) *(i.e.* $|w_k| \le \bar{w}$, $|v_k| \le \bar{v}$, $|e_k| \le \bar{e}$ *for all* k*), a bound* $C_\mathcal{H}$ *on the* $\mathcal{H}_\infty$ *norm of* $G(z)$ *and the root radius* ρ *(i.e., all the roots of* $a(z)$ *locate in the disk of radius* $\rho \le 1$*), find a linear model of the form* (7) *that is consistent with all a priori information and the measurement data, or conclude that none exists.*

IV. Algebraic Reformulation and its SDP Relaxations

In this section, it is shown that Problem 1 can be reformulated as a polynomial optimization problem. Furthermore, SDP relaxations are proposed based on the results introduced in Section II with the consideration of sparsity.

A. Conditions on Consistency with Time-Domain Data

We first define the *regression vector* ϕ_k, the *disturbance vector* $\Delta\eta_k$ and the system *parameter vector* Θ as follows,

$$\begin{aligned} \phi_k &= [-y_{k-1}, \ldots, -y_{k-n}, u_k, \ldots, u_{k-m}]^T, \\ \Delta\eta_k &= [v_{k-1}, \ldots, v_{k-n}, w_{k-1}, \ldots, w_{k-m}]^T, \\ \Theta &= [a_1, \ldots, a_n, b_1, \ldots, b_m]^T. \end{aligned}$$

Hence, the output y_k can be written as

$$y_k = \Theta^T \cdot \phi_k + \Theta^T \cdot \Delta\eta_k + e_k + v_k. \quad (13)$$

Therefore, $G(z)$ is consistent with the measurements and a priori noise information if and only if

$$|y_k - \Theta^T \cdot \phi_k - \Theta^T \cdot \Delta\eta_k - v_k| \le \bar{e},$$

for all integers $k \in [1, L_t]$.

Now let's define polynomials $p_t(\Theta, v, w)$ as

$$\begin{aligned} p_{t,6k-5} &= \bar{e} + y_k - \Theta^T \cdot \phi_k - \Theta^T \cdot \Delta\eta_k - v_k, \\ p_{t,6k-4} &= \bar{e} - y_k + \Theta^T \cdot \phi_k + \Theta^T \cdot \Delta\eta_k + v_k, \\ p_{t,6k-3} &= \bar{v} - v_k, \\ p_{t,6k-2} &= \bar{v} + v_k, \\ p_{t,6k-1} &= \bar{w} - w_k, \\ p_{t,6k} &= \bar{w} + w_k, \end{aligned}$$

for $k = 1, \ldots, L_t$ and a semi-algebraic set $\mathcal{K}_t$ as

$$\mathcal{K}_t \doteq \{(\Theta, v, w) : p_{t,k} \ge 0, k = 1, \ldots, 6L_t\}. \quad (14)$$

Then, we have the following result.

Lemma 1: There exists at least one model $G(z)$ *in the form of* (7) *that is compatible with time-domain data and a priori information if and only if the set* $\mathcal{K}_t$ *is non-empty.*

B. Conditions on Consistency with Frequency-Domain Data

If noisy frequency responses are available and the complex valued noise is assumed to be norm bounded by a known positive constant, a compatible model (7) must satisfy

$$|G(e^{-j\omega_k}) - \hat{G}_k| \le \bar{\epsilon},$$

for all integers $k \in [1, L_f]$, where $\hat{G}_k$ is the measured response at frequency $\omega = \omega_k$.

Define polynomials

$$\begin{aligned} p_{f,k}(\Theta) = \; & \bar{\epsilon}^2 (R_d^2 + I_d^2)^2 - (R_n R_d + I_n I_d - \\ & \alpha_k R_d^2 - \alpha_k I_d^2)^2 + (I_n R_d - R_n I_d - \\ & \beta_k R_d^2 - \beta_k I_d^2)^2 \end{aligned}$$

for $k = 1, \ldots, L_f$, where $\alpha_k = \text{Re}\{\hat{G}_k\}$, $\beta_k = \text{Im}\{\hat{G}_k\}$ and

$$\begin{aligned} R_n &= \text{Re}\{b(e^{-j\omega_k})\}, \\ I_n &= \text{Im}\{b(e^{-j\omega_k})\}, \\ R_d &= \text{Re}\{a(e^{-j\omega_k})\}, \\ I_d &= \text{Im}\{a(e^{-j\omega_k})\}, \end{aligned}$$

are polynomial in Θ. Moreover, a semi-algebraic set $\mathcal{K}_f$ is defined as

$$\mathcal{K}_f \doteq \{\Theta : p_{f,k}(\Theta) \ge 0, k = 1, \ldots, L_f\}. \quad (15)$$

Then, we have the following result.

Lemma 2: There exists at least one model $G(z)$ in the form of (7) that is compatible with frequency-domain data and a priori information if and only if $\mathcal{K}_f$ is non-empty.

C. Bound on $\mathcal{H}_\infty$ Norm

Now we provide an algebraic formulation on bounding the $\mathcal{H}_\infty$ norm of $G(z)$. A stable LTI system satisfies

$$\|G(z)\|_\infty = \left\| \frac{b(z)}{a(z)} \right\|_\infty \leq C_{\mathcal{H}},$$

if and only if

$$R(z) \doteq C_{\mathcal{H}}^2 a(z)a(z^{-1}) - b(z)b(z^{-1}) \geq 0, \tag{16}$$

for all $z \in \mathcal{C}$ with $|z| = 1$. Note that $R(z) = R(z^{-1})$. Hence,

$$R(z) = \sum_{k=-d_r}^{d_r} r_i z^i, r_i = r_{-i} \tag{17}$$

where $d_r = \max\{m-1, n\}$ and r_k are (quadratic) polynomial of Θ. To apply Theorem 3 to build polynomial inequalities, it is necessary to introduce additional variables $\lambda_i, i = 1, \ldots, \frac{d_r(d_r+1)}{2}$. To illustrate, consider the case where $n = 2$ and $m = 1$. Then,

$$Q = \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \lambda_3 & r_2 \\ * & \lambda_2 & r_1 - \lambda_1 \\ * & * & r_0 - \lambda_1 - \lambda_2 \end{pmatrix}.$$

It is positive semi-definite if and only if its principal minors are non-negative. Denote the principal minors of Q by $p_{H,k}$, $k = 1, \ldots, d_r + 1$ and define a semi-algebraic set

$$\mathcal{K}_H \doteq \{(\Theta, \lambda) : p_{H,k} \geq 0, k = 1, \ldots, d_r + 1\}. \tag{18}$$

Then, we have the following result.

Lemma 3: There exists at least one model $G(z)$ in the form of (7) whose $\mathcal{H}_\infty$ norm is bounded by $C_{\mathcal{H}}$ if and only if the set $\mathcal{K}_{\mathcal{H}}$ is non-empty.

D. Conditions on Stability

To obtain the conditions on the system parameters for stability, we use Jury's criterion, which is a necessary and sufficient condition for BIBO stability of the discrete-time system $G(z)$; e.g., see [10]. The criterion has been used in [2] to address stability constraints in identification. For completeness, we briefly restate this well-known result.

TABLE I
JURY'S ARRAY

a_n	a_{n-1}	a_{n-2}	$\cdots$	a_2	a_1	1
1	a_1	a_2	$\cdots$	a_{n-2}	a_{n-1}	a_n
$J_{1,n-1}$	$J_{1,n-2}$	$J_{1,n-3}$	$\cdots$	$J_{1,1}$	$J_{1,0}$	
$J_{1,0}$	$J_{1,1}$	$J_{1,2}$	$\cdots$	$J_{1,n-2}$	$J_{1,n-1}$	
$J_{2,n-2}$	$J_{2,n-3}$	$J_{2,n-4}$	$\cdots$	$J_{2,0}$		
$J_{2,0}$	$J_{2,1}$	$J_{2,2}$	$\cdots$	$J_{2,n-2}$		
$\vdots$	$\vdots$	$\vdots$				
$J_{n-2,2}$	$J_{n-2,1}$	$J_{n-2,0}$				

Lemma 4 (Jury's test [10]): The system $G(z)$ in (7) is stable, i.e., all the roots of $a(z)$ locate inside the unit circle, if and only if the following inequalities hold,

$$\begin{aligned} 1 + \sum_{i=1}^n a_i &\geq 0 \\ 1 + \sum_{i=1}^n (-1)^i a_i &\geq 0 \\ |a_n| &\leq 1 \\ |J_{k,n-k}| &\leq |J_{k,0}|, 1 \leq k \leq n-2 \end{aligned}$$

where $J_{k,i}, k = 1, \ldots, n-2, i = 0, \ldots, n-k$ are the elements in the Jury's array as shown in Table 1 that

$$J_{k,i} \doteq \begin{vmatrix} J_{k-1,n-k} & J_{k-1,i} \\ 1 & J_{k-1,n-k-i} \end{vmatrix},$$

for $k = 1, \ldots, n-2$ and $i = 0, \ldots, n-k$.

Note that every element in the Jury's array is a polynomial of the system parameters $a_1, \ldots, a_n$. Hence, a semi-algebraic set is defined as

$$\mathcal{K}_s \doteq \{(a_1, \ldots, a_n) : p_{s,k}(a) \geq 0, k = 1, \ldots, 2n\} \tag{19}$$

where

$$\begin{aligned} p_{s,1} &= 1 + \sum_{i=1}^n a_i, \\ p_{s,2} &= 1 + \sum_{i=1}^n (-1)^i a_i, \\ p_{s,3} &= 1 + a_n, \\ p_{s,4} &= 1 - a_n, \\ p_{s,2k+3} &= J_{k,0} - J_{k,n-k}, k = 1, \ldots, n-2, \\ p_{s,2k+4} &= J_{k,0} + J_{k,n-k}, k = 1, \ldots, n-2. \end{aligned}$$

Then the following lemma is a direct consequence of Jury's criterion.

Lemma 5: There exist at least one model $G(z)$ in the form of (7) that is stable if and only if the set $\mathcal{K}_s$ is non-empty.

Remark 1: If exponential stability is presumed, i.e., the radius of the roots, $\rho < 1$, is known at a priori, Lemma 5 can be modified by setting $\hat{a}_k = a_k/\rho^k, k = 1, \ldots, n$ and substitute them correspondingly in the polynomials that define $\mathcal{K}_s$.

E. An equivalent reformulation and its SDP relaxations

In this section, we state a polynomial optimization problem that is equivalent to Problem 1 and provide SDP relaxations that take into consideration sparsity. Combing the results from Lemma 1, 2, 3 and 5, we have the following result.

Proposition 1: There exist at least one model $G(z)$ in the form of (7) that satisfies the following conditions,

1) *it is consistent with input/output data (u, y) and a priori information on the noise (v, w);*
2) *it is consistent with the measurements in frequency-domain $((\omega, \hat{G})$ and a priori information on noise e;*

3) $\|G(z)\|_\infty \leq C_{\mathcal{H}}$;
4) $G(z)$ *is stable,*

if and only if the intersection of the sets $\mathcal{K}_t$, $\mathcal{K}_f$, $\mathcal{K}_H$ *and* $\mathcal{K}_s$ *are non-empty.*

Proof: This follows directly from Lemma 1, 2, 3 and 5. □

Thus, to solve Problem 1, it suffices to find a point, if there exists any, in the semi-algebraic set. One may choose any polynomial objective function on the system parameters to form this feasibility problem to a general polynomial optimization. For example, one may set $p_0 = a_1$ or $p_0 = -a_1$ to find the lower/upper bound on a_1.

According to Section II, the polynomial optimization problem can be asymptotically solved by solving a hierarchy of SDP relaxations. However, the size of the SDPs can be very large if the number of variables in the polynomial optimization problem is large. To address this issue, it is desirable to exploit the inherent sparse structure of the problem. It should be noted that the variables appearing in any of the polynomials that define $\mathcal{K}_t$ are included in one of the following sets.

$$I_k \doteq \{\Theta, v_{k-1}, \ldots, v_{k-n}, w_{k-1}, \ldots, w_{k-m}, e_k\}, \tag{20}$$

for $k = 1, \ldots, L_t$. Similarly, the variables in any of the polynomials that define $\mathcal{K}_f$, $\mathcal{K}_H$ and $\mathcal{K}_s$ are included in

$$I_{L_t+1} \doteq \{\Theta, \lambda\}. \tag{21}$$

In general, the objective polynomials p_0 of interest is a function of the parameters of the system, like the examples previously mentioned aim at finding upper/lower bounds on the parameters. If this is the case, then the polynomial optimization problem with objective $p_0(\Theta)$ satisfies *running intersection property*, according to Definition 1.

Now we are ready to state the first main results.

Theorem 4: Given an polynomial objective function $p_0(\Theta)$*, measurements* (u, y) *in time-domain and* $(\omega, \hat{G})$ *in frequency domain, noise bounds* $\bar{v}, \bar{w}, \bar{e}$ *and* $\mathcal{H}_\infty$ *norm bound* $C_{\mathcal{H}}$*, consider the following optimization problem*

$$\begin{aligned} p_s^* = \quad & \min_{m} \quad \sum_\alpha p_{0,\alpha} m_\alpha \\ \textit{s.t.} \quad & M_N(m, I_k) \succeq 0, \quad k = 1, \ldots, L_t + 1 \\ & M_{N_i}(p_{t,i} m, I_k) \succeq 0, \quad i = 1, \ldots, L_t, \\ & M_{N_i}(p_{f,i} m, I_{L_t+1}) \succeq 0, \quad i = 1, \ldots, L_f, \\ & M_{N_i}(p_{H,i} m, I_{L_t+1}) \succeq 0, \quad i = 1, \ldots, L_H, \\ & M_{N_i}(p_{s,i} m, I_{L_t+1}) \succeq 0, \quad i = 1, \ldots, 2n. \end{aligned} \tag{22}$$

Then, if there exists at least one compatible model of the form (7) *for Problem 1,* (22) *is feasible for any relaxation order N. Conversely, if* (22) *is feasible, rank* $M_N(m, I_k) =$ *rank* $M_{N_i}(m, I_k)$ *for all k, and rank* $M_N(m, I_k \cap I_j) = 1$ *for all pairs* (j, k) *with* $I_k \cap I_j \neq \emptyset$*, then, there exists at least one compatible model.*

Proof: This is a direct consequence of Theorem 2, given the fact that running intersection property holds for the collection of the variable sets I_k defined in (20) and in (21). □

Remark 2: The rank condition rank $M_N(m, I_k) =$ rank $M_{N_i}(m, I_k)$ and rank $M_N(m, I_k \cap I_j) = 1$ is a sufficient condition to guarantee that the optimum of the SDP relaxation is the same to the one of the corresponding polynomial optimization problem, see e.g. [8], [13]. With this rank condition being satisfied, an algorithm is given in [8], which can always extract an optimal moment sequence corresponding to a probability measure with point support.

As discussed above, by taking into account sparsity, the computational complexity can be substantially reduced. Hence, the optimization problem can be solved if its size is relatively small. However, the relaxation order N in (22) is not known at a priori, in general one needs to gradually increasing N until the rank conditions are satisfied. Secondly, additional variables λ have been introduced in Section IV.C to formulate the polynomial inequalities associate with bounding the $\mathcal{H}_\infty$ norm. This is also undesirable as the number of variables are then increased to $O(n^2)$.

To overcome these numerical difficulties, we formulate an equivalent rank minimization problem based on the fact that efficient convex relaxations on rank minimization are available in the literature, especially when the matrices that are symmetric and positive semi-definite. This is described in the next section.

V. An Efficient Reformulation via Rank Minimization

Motivated by the fact that the system parameters can be extracted from the optimizer of (22) only if the rank conditions are satisfied, we impose a stronger condition on the rank of the moment matrices, and, hence, simplify (22).

Fix relaxation order $N = d_{\max}/2$ where $d_{\max}$ is the highest total degree of all polynomials $p_{t,k}$, $p_{f,k}$ and $p_{s,k}$. Define a block matrix whose diagonal elements are the moment matrices in (22), i.e.,

$$M = \begin{pmatrix} M_N(m, I_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_N(m, I_{L_t+1}) \end{pmatrix}. \tag{23}$$

Consider the following rank constrained problem,

$$\begin{aligned} p_s^* = \quad & \min_{m,Q} \quad \sum_\alpha p_{0,\alpha} m_\alpha \\ \text{s.t.} \quad & \text{rank}(M) = L_t + 1, \\ & M \succeq 0, \\ & \sum_\alpha p_{t,i,\alpha} m^\alpha \geq 0, i = 1, \ldots, L_t, \\ & \sum_\alpha p_{f,i,\alpha} m^\alpha \geq 0, i = 1, \ldots, L_f, \\ & \sum_\alpha p_{s,i,\alpha} m^\alpha \geq 0, i = 1, \ldots, L_H, \\ & Q \doteq [q_{ij}] \succeq 0, \\ & r_k = \sum_{i=k}^{n+k} q_{i,i-k}, k = 0, \ldots, n. \end{aligned} \tag{24}$$

where $p_{t,k,\alpha}$, $p_{f,k,\alpha}$ and $p_{s,k,\alpha}$ are the corresponding coefficients of polynomials $p_{t,k}$, $p_{f,k}$ and $p_{s,k}$, respectively; r_k

are the coefficients of $R(z)$, as defined in (17) and m^α are the corresponding elements in M. Then we have our second main result.

Theorem 5: There exists at least one compatible model in Problem 1 if and only if problem (24) *is feasible.*

Proof: Assume that there exists such a compatible model with parameters Θ admissible noise (v, w, e). Then set $m^\alpha = (\Theta, v, w)^\alpha$, i.e., the moment sequence m is assigned with a Dirac distribution at the point (Θ, v, w). Hence, the moment matrices $M_N(m, I_k)$, $k = 1, \ldots, L_f$ are of rank one and all the polynomials $p_{t,k}$, $p_{f,k}$ and $p_{s,k}$ are non-negative. Moreover, according to Theorem 3, there exists a positive semi-definite matrix Q such that (6) holds.

On the other hand, let m and Q be an optimizer of (24). Since $\text{rank}(M_N(m, I_k)) \geq 1$ and $\text{rank}(M) = L_t$, we must have

$$\text{rank}(M_N(m, I_k)) = 1, k = 1, \ldots, L_t.$$

Hence, the sequence m is associated with a Dirac distribution. Assume that the Dirac probability density function is only non-zero at the point (Θ, v, w), then, a system associated with parameters Θ is a compatible model for Problem 1. □

Since the rank condition $\text{rank}(M) = L_t$ is in general difficult to address directly, we rewrite problem (24) as

$$\begin{aligned}
\sigma = \min_{m,Q} \quad & \text{rank}(M) && (25)\\
\text{s.t.} \quad & \sum_\alpha p_{0,\alpha} m_\alpha \geq \gamma, && (26)\\
& \sum_\alpha p_{t,i,\alpha} m^\alpha \geq 0, i = 1, \ldots, L_t, && (27)\\
& \sum_\alpha p_{f,i,\alpha} m^\alpha \geq 0, i = 1, \ldots, L_f, && (28)\\
& \sum_\alpha p_{s,i,\alpha} m^\alpha \geq 0, i = 1, \ldots, L_H, && (29)\\
& M \succeq 0, && (30)\\
& Q \doteq [q_{ij}] \succeq 0, && (31)\\
& r_k = \sum_{i=k}^{n+k} q_{i,i-k}, k = 0, \ldots, n, && (32)
\end{aligned}$$

where γ is a preset constant. It is easy to see this problem is equivalent to Problem 1 and (24) in the following sense.

Corollary 1: There exists at least one compatible model in Problem 1 if and only if problem (25) *has an optimal value of* L_t *for some* γ. *Moreover,* (25) *has optimal value* L_t *for some* γ *if and only if the minimum of* (24) *satisfies*

$$p_s^* \geq \gamma. \quad (33)$$

Remark 3: Note that in problem (24) and (25), the relaxation order N is fixed and is only determined by the highest total degree of all the polynomials p_0, p_t, p_f and p_s, and, hence, is known at a priori. There is no need to increase the value of N since we aim at finding rank one moment matrices.

Although rank minimization is NP-hard, efficient convex relaxations are available. In particular, good approximate solutions can be obtained by using a log-det heuristic that relaxes rank minimization to a sequence of convex problems. Thus, we use the log-det heuristic algorithm proposed in [5], which has been proven to be efficient in practice, to solve problem (25), as summarized below.

Algorithm 1 Rank Minimization

Set $X \leftarrow M(m)$, $X_0 \leftarrow I$, $\delta \leftarrow 0$, $k \leftarrow 0$.
repeat
Solve

$$X_{k+1} \leftarrow \quad \arg\min \quad \text{Tr}(X_k + \delta I)^{-1} X \qquad \text{s.t.} \quad (26) \text{ - } (32).$$

Decompose the symmetric matrix $X_k = T^{-1} D T$.
Set $\delta \leftarrow \min \text{diag}(D) + \delta_0$.
Set $k \leftarrow k + 1$.
until a convergence criterion is reached.
return X_k

VI. Numerical Results

In this section, we present two simple numerical examples to illustrate the efficiency of the proposed algorithms.

A. Example 1

We first consider a second order system with transfer function

$$G(z) = \frac{1}{1 - 1.8z^{-1} + 0.9997z^{-2}}.$$

This system is stable with roots at $0.9 \pm j0.4355$ and $\mathcal{H}_\infty$ norm being 6757. To perform the identification process, the system was excited with $L_t = 30$ random input between -1 and 1 and process noise bounded by 0.3, i.e., $\bar{w} = 0, \bar{v} = 0, \bar{e} = 0.3$ The frequency responses were also measured at 2 randomly select frequency points, with disturbance being complex numbers that are norm bounded by 0.2, i.e., $\bar{\epsilon} = 0.2$.

In the first experiment, the system was identified without imposing conditions on stability. The corresponding identified system is

$$\hat{G}(z) = \frac{0.7833}{1 - 1.807z^{-1} + 1.014z^{-2}}.$$

It can be verified that the identified system is NOT stable.

After imposing conditions on exponential stability and on the $\mathcal{H}_\infty$ norm, i.e., $\rho \leq 0.9997$ and $\|\hat{G}(z^{-1})\|_\infty \leq 7000$ the identified system is

$$\hat{G}(z) = \frac{0.8685}{1 - 1.787z^{-1} + 0.9994z^{-2}},$$

which is stable with $\mathcal{H}_\infty$ norm being 3228.

B. Example 2

A third order system is considered with transfer function

$$G(z) = \frac{2z^{-1} - 1.2z^{-2} + 0.3z^{-3}}{1 - 0.815z^{-1} - 0.7738z^{-2} + 0.9842z^{-3}}.$$

This system is stable with H_∞ norm $\|G(z)\|_\infty = 1750.97$ and root radius $\rho = 0.9996$. To perform the identification process, the system was excited with $L_t = 300$ random input between -1 and 1 and process noise bounded by 0.2, i.e., $\bar{w} = 0, \bar{v} = 0, \bar{e} = 0.2$. The frequency responses were also measured at 10 randomly select frequency points, with disturbance being complex numbers that are norm bounded by 0.2, i.e., $\bar{\epsilon} = 0.2$.

The system was first identified without posing conditions on stability and $\mathcal{H}_\infty$ norm. The corresponding identified system is

$$\hat{G}(z) = \frac{1.991z^{-1} - 1.2z^{-2} + 0.3056z^{-3}}{1 - 0.8148z^{-1} - 0.7743z^{-2} + 0.9846z^{-3}}.$$

It can be verified that the identified system is stable but with $\mathcal{H}_\infty$ norm being 2020.2, which is significantly larger than that of the true system.

Next, the system was identified with conditions that the system is stable and the $\mathcal{H}_\infty$ norm is less than 1900. Then, the corresponding identified system is

$$\hat{G}(z) = \frac{1.985z^{-1} - 1.186z^{-2} + 0.3023z^{-3}}{1 - 0.8156z^{-1} - 0.7729z^{-2} + 0.9834z^{-3}}.$$

It can be verified that this identified system is stable and its $\mathcal{H}_\infty$ norm is 1359.7, which is less than that of the true system.

Moreover, with the same experimental data, Algorithm 1 was performed multiple times to find the upper/lower bounds on a_1. Without imposing stability constraint and norm condition, it is found $a_1 \in [-0.83, -0.87]$. On the other hand, with stability constraint and norm condition, it is found that $a_1 \in [-0.82, -0.85]$. Thus, it can be seen that imposing conditions on stability and $\mathcal{H}_\infty$ norm can lead to a tighter bound on system parameters.

VII. Concluding Remarks

In this paper, we have addressed the problem of identifying fixed order stable systems from time and frequency domain measurements. We started by showing that this problem is equivalent to finding a point in a suitably defined semi-algebraic set. To efficiently solve this problem, the sparse structure of the polynomials involved in the description of the set was exploited in building a sequence of computationally efficient convergent SDP relaxations. Moreover, to further improve efficiency, an equivalent formulation is derived involving rank constraints. This is motivated by the fact that efficient convex relaxations are available for rank minimization problems. Finally, two simple numerical examples are provided to illustrate the efficiency of the proposed algorithm. Further effort is now being put on improving the computational efficiency of the proposed algorithms.

References

[1] Vito Cerone, Dario Piga, and Diego Regruto. Bounding the parameters of linear systems with stability constraints. In *Proc. American Control Conference*, 2010.

[2] Vito Cerone, Dario Piga, and Diego Regruto. Set-membership eiv identification through lmi relaxation techniques. In *Proc. American Control Conference*, 2010.

[3] N. L. C. Chui and J. M. Maciejowski. Realization of stable models with subspace methods. *Automatica*, 32(11):1587–1595, 1996.

[4] B. Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer, 2007.

[5] M. Fazel, H. Hindi, and S. Boyd. Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices. In *Proc. American Control Conference*, 2003.

[6] C. Feng, C. Lagoa, N. Ozay, and M. Sznaier. Hybrid system identification: an SDP approach. In *Proc. IEEE Conf. Dec. Contr.*, 2010.

[7] C. Feng, C. M. Lagoa, and M. Sznaier. Hybrid system identification via sparse polynomial optimization. In *Proc. American Control Conference*, 2010.

[8] D. Henrion and J. B. Lasserre. *Positive polynomials in control*, chapter Detecting global optimality and extracting solutions in GloptiPoly. Lecture Notes on Control and Information Sciences. Springer Verlag, 2005.

[9] T. Inanc, M. Sznaier, P. A. Parrilo, and R. S. Pena. Robust identification with mixed parametric/nonparametric models and time/frequency-domain experiments: theory and an application. *IEEE Trans. Contr. Syst. Technol.*, 9(4):608–617, 2001.

[10] E. Jury. *The theory and application of Z-transform method*. Wiley, New York, 1964.

[11] S.L. Lacy and D.S. Bernstein. Subspace identification with guaranteed stability using constrained optimization. *IEEE Trans. Automat. Contr.*, 48(7):1259 – 1263, jul. 2003.

[12] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.

[13] Jean B. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM J. on Optimization*, 17(3):822–843, 2006.

[14] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. *Emerging applications of algebraic geometry*, 149:157–270, 2009.

[15] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, 1999.

[16] J. M. Maciejowski. Guaranteed stability with subspace methods. *Syst. Control Lett.*, 26(2):153–156, 1995.

[17] R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach*. IEEE Press, Piscataway, NJ, 2001.

[18] P.A. Regalia and P. Stoica. Stability of multivariable least-squares models. *Signal Processing Letters, IEEE*, 2(10):195 –196, oct. 1995.

[19] S. Siddiqi, B. Boots, and G. Gordon. A constraint generation approach to learning stable linear dynamical systems. In *Advances in Neural Information Processing Systems*, December 2007.

[20] T. Van Gestel, J.A.K. Suykens, P. Van Dooren, and B. De Moor. Identification of stable models in subspace identification by using regularization. *IEEE Trans. Automat. Contr.*, 46(9):1416 –1420, sep. 2001.

[21] M. Verhaegen and V. Verdult. *Filtering and System Identification: An Introduction*. TU-Delft/ITS, lecture notes for the course SC4040(ET4094) edition, 2003.