2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Realization from moments: The linear case

Jin Gyu Lee and Alessandro Astolfi

***Abstract*— By exploiting the time-domain notion of moments we recover a time-domain counterpart of the fact that a certain number of steady-state responses is sufficient to realize a linear system. This may pave the way to a realization theory for nonlinear systems based on their steady-state responses.**

I. Introduction

This paper is motivated by the simple observation that a certain number of steady-state responses is sufficient to realize a SISO linear system, that is (almost) any $2n$ 0-moments [1], i.e., any $2n$ pairs $(s_i, G(s_i)) \in \mathbb{C}^2$, $i = 1, \dots, 2n$, with

$$G(s) = \frac{b_1 s^{n-1} + \cdots + b_n}{s^n + a_1 s^{n-1} + \cdots + a_n},$$

can be used to "recover" the transfer function itself. In fact, the parameters a_i and b_i, $i = 1, \dots, n$, can be obtained by solving the $2n$ linear equations $(i = 1, \dots, 2n)$

$$G(s_i)[s_i^n + a_1 s_i^{n-1} + \cdots + a_n] = b_1 s_i^{n-1} + \cdots + b_n,$$

which are always solvable provided that $s_i \neq s_j$, for $i \neq j$, and that each s_i is not a pole of $G(\cdot)$.

The goal of this paper is to recover a *time-domain* counterpart of this fact. This yields several advantages, among which we are interested in the possibility of developing a nonlinear enhancement, thus possibly paving the way to a realization theory for nonlinear systems based on their steady-state responses. For this purpose, we utilize the time-domain notion of moment, developed in the theory of model reduction by moment matching [1], [2]. In addition to the aforementioned advantage, this also provides a good starting point for our investigation: a simple and direct parameterization of all the reduced order models (approximations) achieving moment matching, i.e., the moments (steady-state responses) of the approximation and the system coincide at the interpolation points.

To begin with, we identify a necessary and sufficient condition for sufficiently many other moments of the approximation and of the original system to match, as this is necessary for "recovering" the original system. It turns out that this condition implies that the approximation is identical to the original system, possibly after pole-zero cancellations and a coordinates transformation. Then, to complete the paper, we provide guidelines to actually find such an approximation having a stable pole-zero cancellation, using only moments. By following these guidelines, we utilize exactly $2n$ moments to realize the original system: this recovers the fact that with a certain number of steady-state responses one can realize a linear system.

This means that among the infinitely many moments, only a finite (but sufficiently many) of them determine the entire response, that is any finite (but sufficiently many) number of steady-state responses can reproduce not only all the steady-state responses but also all the transient responses.

To make the paper self-contained, in Section II we briefly review the problem of model reduction by moment matching for linear systems in a time-domain perspective. Then, the realization of a linear system using moments is discussed in Sections III and IV. Conclusions are given in Section V and some technical results are given in the Appendix.

Notation: Throughout the paper, we use standard notation. $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$ denote the set of real numbers, of n-dimensional vectors with real components, and of $n \times m$-dimensional matrices with real entries, respectively. $\mathbb{C}$ and $\mathbb{C}^n$ denote the set of complex numbers and of n-dimensional vectors with complex components, respectively, $\mathbb{C}^0$ denotes the set of complex numbers with zero real part, and $\mathbb{C}^-$ denotes the set of complex numbers with negative real part. $\sigma(A)$ denotes the spectrum of the matrix $A \in \mathbb{R}^{n\times n}$. $\emptyset$ denotes the empty set. For vectors or matrices a and b, $\mathrm{col}(a, b) := \begin{bmatrix} a^T & b^T \end{bmatrix}^T$. For matrices $A_1, \dots, A_k$, we denote by $\mathrm{diag}(A_1, \dots, A_k)$ the block diagonal matrix with diagonal blocks $A_1, \dots, A_k$.

II. A brief review of model reduction by moment matching for linear systems

A. Notion of moment

Consider a linear, single-input, single-output, system described by the equation

$$\begin{aligned} \dot{x} &= Ax + Bu \in \mathbb{R}^n, \\ y &= Cx \qquad\;\; \in \mathbb{R}, \end{aligned} \tag{1}$$

the single-output signal generator

$$\begin{aligned} \dot{w}_g &= S_g w_g \in \mathbb{R}^{N_g}, \\ u &= L_g w_g \in \mathbb{R}, \end{aligned} \tag{2}$$

and the single-input filter

$$\dot{w}_f = S_f w_f + L_f y \in \mathbb{R}^{N_f}. \tag{3}$$

This work has been supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 739551 (KIOS CoE); by the Italian Ministry for Research in the framework of the 2017 and of the 2020 Program for Research Projects of National Interest (PRIN), Grant No. 2017YKXYXJ and 2020RTWES4; and by the EPSRC grant "Model Reduction from Data", Grant No. EP/W005557. This work was done while Jin Gyu Lee was with Imperial College London.

J. G. Lee is with Inria, University of Lille, CNRS, UMR 9189 - CRIStAL, F-59000 Lille, France. A. Astolfi is with Control and Power Research Group, Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom. jin-gyu.lee@inria.fr, a.astolfi@imperial.ac.uk

Let $G(s) = C(sI - A)^{-1}B$. The signal generator and the filter capture the requirement that one is interested in studying the input-output behavior of the system (1) only in *specific circumstances*. In what follows we assume that A, S_g, and S_f do not have common eigenvalues, that is the following assumption holds.

Assumption 1: $\sigma(A) \cap \sigma(S_g) = \sigma(A) \cap \sigma(S_f) = \sigma(S_g) \cap \sigma(S_f) = \emptyset$. □

Definition 1: The forward moment of the system (1) is the matrix

$$M_g := C\Pi_g,$$

where Π_g is the unique solution of the Sylvester equation

$$A\Pi_g + BL_g = \Pi_g S_g. \tag{4}$$

Correspondingly, the backward moment of the system (1) is the matrix

$$M_f := \Pi_f B,$$

where Π_f is the unique solution of the Sylvester equation

$$\Pi_f A + L_f C = S_f \Pi_f. \tag{5}$$

□

Note that moments are associated with the input-output behavior of the system (1), the signal generator (2), and the filter (3), hence in what follows we make the following standing assumption.

Assumption 2: The system (1) is controllable and observable, the signal generator (2) is observable, and the filter (3) is controllable. □

Definition 1 provides a generalized time-domain notion of moments. As regarding to the *frequency-domain* notion of moments we have the following result.

Definition 2 ([1]): The 0-moment of the system (1) at $s^* \in \mathbb{C}$ is the complex number

$$\eta_0(s^*) = C(s^*I - A)^{-1}B.$$

The k-moment of the system (1) at $s^* \in \mathbb{C}$ is the complex number

$$\begin{aligned}\eta_k(s^*) &= \frac{(-1)^k}{k!}\left[\frac{d^k}{ds^k}(C(sI-A)^{-1}B)\right]_{s=s^*}\\ &= C(s^*I - A)^{-(k+1)}B.\end{aligned}$$

□

Lemma 1 ([2]): Consider the system (1) and $s^* \in \mathbb{R}$. Suppose $s^* \notin \sigma(A)$. Then the moments $\eta_0(s^*), \ldots, \eta_k(s^*)$ are in one-to-one relationship with the forward moment $M_g = C\Pi_g$ of the signal generator (2) with S_g any non-derogatory[1] real matrix such that

$$\det(sI - S_g) = (s - s^*)^{k+1}, \tag{6}$$

and L_g such that the pair (L_g, S_g) is observable. □

Lemma 2 ([2]): Consider the system (1) and $s^* \in \mathbb{C} \setminus \mathbb{R}$. Suppose $s^* \notin \sigma(A)$. Then the moments $\eta_0(s^*), \eta_0(\overline{s}^*), \ldots, \eta_k(s^*)$, and $\eta_k(\overline{s}^*)$ are in one-to-one relationship with the forward moment $M_g = C\Pi_g$ of the signal generator (2) with S_g any non-derogatory real matrix such that

$$\det(sI - S_g) = ((s - s^*)(s - \overline{s}^*))^{k+1}, \tag{7}$$

and L_g such that the pair (L_g, S_g) is observable. □

Remark 1: One can easily derive similar results for the backward moments. □

The concept of moments has a tight relation with the concept of steady-state responses as illustrated in what follows.

Proposition 1 ([2]): Consider the system (1), $s^* \in \mathbb{C}$, and $k \geq 0$. Assume $\sigma(A) \subset \mathbb{C}^-$ and $s^* \in \mathbb{C}^0$. Let

$$N_g = \begin{cases} k+1 & \text{if } s^* \in \mathbb{R}, \\ 2(k+1) & \text{if } s^* \in \mathbb{C} \setminus \mathbb{R}, \end{cases}$$

and $S_g \in \mathbb{R}^{N_g \times N_g}$ be any non-derogatory matrix with characteristic polynomial as in (6), if $s^* \in \mathbb{R}$, or as in (7), if $s^* \in \mathbb{C} \setminus \mathbb{R}$.

Consider the interconnection of systems (1) and (2), and L_g such that the pair (L_g, S_g) is observable.

Then the moments $\eta_0(s^*), \ldots, \eta_k(s^*)$ are in one-to-one relationship with the (well-defined) steady-state response of the output of the interconnected system. □

B. Moment matching

We are now in a position to define, precisely, the notion of approximating system.

Definition 3: The system

$$\begin{aligned}\dot{\hat{x}} &= \hat{A}\hat{x} + \hat{B}u \in \mathbb{R}^{\hat{n}}\\ y &= \hat{C}\hat{x} \qquad \in \mathbb{R}\end{aligned} \tag{8}$$

is an approximation of the system (1) with regards to the signal generator (2) (to the filter (3)) if for a minimal realization $(\bar{C}, \bar{A}, \bar{B})$ of (8) with dimension $\bar{n}$, we have

$$\sigma(\bar{A}) \cap \sigma(S_g) = \emptyset \quad (\sigma(\bar{A}) \cap \sigma(S_f) = \emptyset) \tag{9}$$

and

$$C\Pi_g = \bar{C}\bar{\Pi}_g \quad (\Pi_f B = \bar{\Pi}_f \bar{B}), \tag{10}$$

where Π_g (Π_f) is the unique solution of the Sylvester equation (4) ((5)), with L_g (L_f) such that the pair (L_g, S_g) is observable ((S_f, L_f) is controllable), and $\bar{\Pi}_g$ ($\bar{\Pi}_f$) is the unique solution of the Sylvester equation

$$\bar{A}\bar{\Pi}_g + \bar{B}L_g = \bar{\Pi}_g S_g \quad (\bar{\Pi}_f \bar{A} + L_f \bar{C} = S_f \bar{\Pi}_f). \tag{11}$$

□

Remark 2: In Definition 3 we have implicitly extended our definition of moments, given in Definition 1, to systems that are not minimal. We emphasize that when $\sigma(\hat{A}) \cap \sigma(S_g) = \emptyset$ ($\sigma(\hat{A}) \cap \sigma(S_f) = \emptyset$), the Sylvester equation

$$\hat{A}\hat{\Pi}_g + \hat{B}L_g = \hat{\Pi}_g S_g \quad (\hat{\Pi}_f \hat{A} + L_f \hat{C} = S_f \hat{\Pi}_f)$$

has a unique solution and it satisfies the condition

$$C\Pi_g = \bar{C}\bar{\Pi}_g = \hat{C}\hat{\Pi}_g \quad (\Pi_f B = \bar{\Pi}_f \bar{B} = \hat{\Pi}_f \hat{B}),$$

[1] A matrix is non-derogatory if its characteristic and minimal polynomials coincide.

regardless of the minimality (see Proposition 3 and Lemma 5). This justifies the extension. □

According to [2], a parameterized family of approximations of the system (1) with regards to the signal generator (2) (to the filter (3)) described by equations of the form (8) with

$$\hat{A} = S_g - K_g L_g, \quad \hat{B} = K_g, \quad \hat{C} = C\Pi_g$$
$$(\hat{A} = S_f - L_f H_f, \quad \hat{B} = \Pi_f B, \quad \hat{C} = H_f)$$

namely

$$\begin{aligned} \dot{\chi}_g &= (S_g - K_g L_g)\chi_g + K_g u \in \mathbb{R}^{N_g}, \\ y &= M_g \chi_g \qquad\qquad\qquad\quad \in \mathbb{R}, \end{aligned} \tag{12}$$

$$\left(\begin{aligned} \dot{\chi}_f &= (S_f - L_f H_f)\chi_f + M_f u \in \mathbb{R}^{N_f}, \\ y &= H_f \chi_f \qquad\qquad\qquad\quad \in \mathbb{R}, \end{aligned} \right) \tag{13}$$

with the matrix parameter K_g (H_f) is *complete*, i.e., the family (12) ((13)) contains all approximations of dimension N_g (N_f) achieving moment matching.

Proposition 2 ([2]): Consider an approximation of the system (1) with regards to the signal generator (2) of dimension $N_g \leq n$, and let $\hat{G}(s)$ be its transfer function. Then, there exists a unique K_g such that $\hat{G}(s) = M_g(sI - (S_g - K_g L_g))^{-1} K_g$, i.e., the family of systems (12) contains all approximations of the system (1) with regards to the signal generator (2) of dimension N_g. □

Remark 3: One can easily derive a similar result for the backward moments. □

III. A NECESSARY AND SUFFICIENT CONDITION FOR AN APPROXIMATION TO MATCH OTHER MOMENTS

In [2], special emphasis has been devoted to the parameterization of approximations (which is briefly reviewed in Section II-B), and to the selection of parameters under additional constraints, such as to preserve the properties of the original system (e.g., stability or passivity) [3]. On the contrary to other solutions in the linear framework, e.g., projection methods and interpolation theory, the parameterization as in Section II-B provides a unifying perspective, i.e., the selection of parameters does have a simple and direct interpretation [2]. Motivated by this, to recover a time-domain counterpart of the fact introduced at the beginning of this paper, we consider the selection of parameters with the additional property of matching (sufficiently many) other moments, as this is necessary for "recovering" the original system.

To begin with, we find a necessary and sufficient condition for the forward moments $\bar{M}_g^k := C\bar{\Pi}_g^k$, $k = 1, \ldots, o_g$ of the system (1) with regards to (sufficiently many) other signal generators

$$\begin{aligned} \dot{\bar{w}}_g^k &= \bar{S}_g^k \bar{w}_g^k \in \mathbb{R}^{\bar{N}_g^k}, \\ u &= \bar{L}_g^k \bar{w}_g^k \in \mathbb{R}, \qquad k = 1, \ldots, o_g, \end{aligned} \tag{14}$$

to match each forward moment $\widehat{M}_g^k := \widehat{C}\widehat{\Pi}_g^k$, $k = 1, \ldots, o_g$ of the approximation (12) with regards to the signal generator (14), where $(\widehat{C}, \widehat{A}, \widehat{B})$ is a minimal realization (with dimension $\widehat{n}$) of the approximation (12) and $\bar{\Pi}_g^k$ and $\widehat{\Pi}_g^k$ are the unique solution of the Sylvester equations

$$A\bar{\Pi}_g^k + B\bar{L}_g^k = \bar{\Pi}_g^k \bar{S}_g^k \tag{15}$$

and

$$\widehat{A}\widehat{\Pi}_g^k + \widehat{B}\bar{L}_g^k = \widehat{\Pi}_g^k \bar{S}_g^k, \tag{16}$$

respectively.

Theorem 1: Consider the system (1), the signal generator (2), and the approximation (12). Let Assumptions 1 and 2 hold. Then, the condition

$$C\bar{\Pi}_g^k = \bar{M}_g^k = \widehat{M}_g^k = \widehat{C}\widehat{\Pi}_g^k, \quad k = 1, \ldots, o_g$$

hold for a set of signal generators (14) such that $(\bar{L}_g^k, \bar{S}_g^k)$ is observable, $\sigma(A) \cap \sigma(\bar{S}_g^k) = \emptyset$ and $\sigma(\widehat{A}) \cap \sigma(\bar{S}_g^k) = \emptyset$, and

$$\left| \bigcup_{k=1,\ldots,o_g} \sigma(\bar{S}_g^k) \right| \geq n + \widehat{n} \quad \text{(i.e., sufficiently many)}, \tag{17}$$

where Π_g, $\bar{\Pi}_g^k$, and $\widehat{\Pi}_g^k$ are the unique solution of the Sylvester equations (4), (15), and (16), respectively, if and only if

$$N_g \geq n \quad \text{and} \quad \Pi_g K_g = B. \tag{18}$$

In particular, if (18) is satisfied, then (C, A, B) is a minimal realization of the approximation (12), hence the forward moments of the system (1) and of the approximation (12) of any other signal generators $\dot{\bar{w}}_g = \bar{S}_g \bar{w}_g$, $u = \bar{L}_g \bar{w}_g$, such that $(\bar{L}_g, \bar{S}_g)$ is observable and $\sigma(A) \cap \sigma(\bar{S}_g) = \emptyset$, exist and match. □

Proof:

(⇐) As $M_g = C\Pi_g$, $\Pi_g K_g = B$, and

$$\Pi_g(S_g - K_g L_g) = \Pi_g S_g - BL_g = A\Pi_g,$$

(C, A, B) is a minimal realization of the approximation (12), by (18).

(⇒) By collecting the Sylvester equations (15) and (16), for $k = 1, \ldots, o_g$, we have that

$$\begin{aligned} &\mathcal{A}\mathcal{P}_g + \mathcal{B}\bar{L}_g \\ &\quad := \begin{bmatrix} A & 0 \\ 0 & \widehat{A} \end{bmatrix} \begin{bmatrix} \bar{\Pi}_g \\ \widehat{\Pi}_g \end{bmatrix} + \begin{bmatrix} B \\ \widehat{B} \end{bmatrix} \bar{L}_g = \begin{bmatrix} \bar{\Pi}_g \\ \widehat{\Pi}_g \end{bmatrix} \bar{S}_g = \mathcal{P}_g \bar{S}_g, \end{aligned} \tag{19}$$

where $\bar{L}_g = \begin{bmatrix} \bar{L}_g^1 & \cdots & \bar{L}_g^{o_g} \end{bmatrix}$, $\bar{S}_g = \text{diag}(\bar{S}_g^1, \ldots, \bar{S}_g^{o_g})$,

$$\bar{\Pi}_g = \begin{bmatrix} \bar{\Pi}_g^1 & \cdots & \bar{\Pi}_g^{o_g} \end{bmatrix}, \quad \text{and} \quad \widehat{\Pi}_g = \begin{bmatrix} \widehat{\Pi}_g^1 & \cdots & \widehat{\Pi}_g^{o_g} \end{bmatrix}.$$

The moment matching condition is now represented as $\bar{M}_g := C\bar{\Pi}_g = \widehat{C}\widehat{\Pi}_g =: \widehat{M}_g$.

Now, perform a "controllable" decomposition and an "observable" decomposition of the pair $(\mathcal{A}, \mathcal{B})$ and $(\bar{L}_g, \bar{S}_g)$, respectively. Then, we have coordinates transformation matrices $T = \text{col}(\bar{T}, \tilde{T})$ and $P = \text{col}(\bar{P}, \tilde{P})$ such that

$$\begin{bmatrix} \bar{T} \\ \tilde{T} \end{bmatrix} \mathcal{A} \begin{bmatrix} \bar{T}^\dagger & \tilde{T}^\dagger \end{bmatrix} = \begin{bmatrix} \bar{\mathcal{A}} & \check{\mathcal{A}} \\ 0 & \tilde{\mathcal{A}} \end{bmatrix}, \quad \begin{bmatrix} \bar{T} \\ \tilde{T} \end{bmatrix} \mathcal{B} = \begin{bmatrix} \bar{\mathcal{B}} \\ 0 \end{bmatrix},$$

and the pair $(\bar{\mathcal{A}}, \bar{\mathcal{B}}) \in \mathbb{R}^{\bar{N}\times\bar{N}} \times \mathbb{R}^{\bar{N}\times 1}$ is controllable and

$$\begin{bmatrix}\bar{P}\\ \tilde{P}\end{bmatrix}\bar{S}_g\begin{bmatrix}\bar{P}^\dagger & \tilde{P}^\dagger\end{bmatrix}=\begin{bmatrix}\widehat{S}_g & 0\\ \check{S}_g & \tilde{S}_g\end{bmatrix}, \quad \bar{L}_g\begin{bmatrix}\bar{P}^\dagger & \tilde{P}^\dagger\end{bmatrix}=\begin{bmatrix}\widehat{L}_g & 0\end{bmatrix},$$

and the pair $(\widehat{L}_g, \widehat{S}_g) \in \mathbb{R}^{1\times\widehat{N}_g} \times \mathbb{R}^{\widehat{N}_g\times\widehat{N}_g}$ is observable, respectively, where

$$T^{-1} =: \begin{bmatrix}\bar{T}^\dagger & \tilde{T}^\dagger\end{bmatrix} \quad \text{and} \quad P^{-1} =: \begin{bmatrix}\bar{P}^\dagger & \tilde{P}^\dagger\end{bmatrix}.$$

Note that, by (17) and by the observability of the pairs $(\bar{L}_g^k, \bar{S}_g^k)$, $k = 1, \dots, o_g$, we have that $\widehat{N}_g \geq n + \widehat{n}$.

By our assumptions, we have $\sigma(\mathcal{A}) \cap \sigma(\bar{S}_g) = \emptyset$, and thus, from Proposition 3 and Lemma 5, the unique solution $\mathcal{P}_g$ of the Sylvester equation (19) satisfies the condition $\mathrm{rank}(\mathcal{P}_g) = \min\{\bar{N}, \widehat{N}_g\} = \bar{N} \leq n + \widehat{n} \leq \widehat{N}_g$ and

$$\begin{bmatrix}\bar{\Pi}_g\\ \widehat{\Pi}_g\end{bmatrix} = \mathcal{P}_g = \bar{T}^\dagger \bar{\mathcal{P}}_g \bar{P},$$

where $\bar{\mathcal{P}}_g$ is the unique solution of the Sylvester equation

$$\bar{\mathcal{A}}\bar{\mathcal{P}}_g + \bar{\mathcal{B}}\widehat{L}_g = \bar{\mathcal{P}}_g\widehat{S}_g. \tag{20}$$

Now, $\bar{\mathcal{P}}_g$ has full row rank as $\widehat{N}_g \geq n + \widehat{n} \geq \bar{N}$, and this implies that there exists $\mathcal{K}_g$ such that

$$\bar{\mathcal{B}} = \bar{\mathcal{P}}_g\mathcal{K}_g.$$

This further implies that

$$\begin{bmatrix}B\\ \widehat{B}\end{bmatrix} = \mathcal{B} = \bar{T}^\dagger\bar{\mathcal{B}} = \bar{T}^\dagger\bar{\mathcal{P}}_g\mathcal{K}_g = \mathcal{P}_g\bar{P}^\dagger\mathcal{K}_g =: \begin{bmatrix}\bar{\Pi}_g\\ \widehat{\Pi}_g\end{bmatrix}\bar{K}_g.$$

Therefore, we have that

$$A\bar{\Pi}_g = \bar{\Pi}_g(\bar{S}_g - \bar{K}_g\bar{L}_g),$$
$$\widehat{A}\widehat{\Pi}_g = \widehat{\Pi}_g(\bar{S}_g - \bar{K}_g\bar{L}_g),$$

and this further provides the conditions:

$$\begin{aligned}CA^iB = CA^i\bar{\Pi}_g\bar{K}_g &= C\bar{\Pi}_g(\bar{S}_g - \bar{K}_g\bar{L}_g)^i\bar{K}_g\\ &= \widehat{C}\widehat{\Pi}_g(\bar{S}_g - \bar{K}_g\bar{L}_g)^i\bar{K}_g\\ &= \widehat{C}\widehat{A}^i\widehat{\Pi}_g\bar{K}_g = \widehat{C}\widehat{A}^i\widehat{B}, \quad \forall i \geq 0.\end{aligned}$$

This means that the transfer functions $C(sI - A)^{-1}B$ and $\widehat{C}(sI - \widehat{A})^{-1}\widehat{B}$ are identical, hence $n = \widehat{n} \leq N_g$.

Finally, the fact that $(\widehat{C}, \widehat{A}, \widehat{B})$ is a minimal realization of the approximation (12) ensures that there exists a matrix $\Pi \in \mathbb{R}^{n\times N_g}$ such that $\mathrm{rank}(\Pi) = n$ and

$$M_g = C\Pi, \quad \Pi K_g = B, \quad \Pi(S_g - K_gL_g) = A\Pi.$$

This implies that

$$\Pi S_g = A\Pi + BL_g,$$

hence by the uniqueness of the solution of the Sylvester equation (4), we have that $\Pi = \Pi_g$, and thus, $\Pi_g K_g = B$. ■

Theorem 1 can be interpreted as follows: the necessary and sufficient condition (18) is such that the approximation (12) is identical to the system (1) after pole-zero cancellation and a coordinates transformation. In particular, selecting the variables[2] $\hat{x}_g = \Pi_g\chi_g$ and $\tilde{x}_g = \tilde{\Pi}_g\chi_g$ we have that

$$\begin{aligned}\dot{\hat{x}}_g &= \Pi_g(S_g - K_gL_g)\Pi_g^\dagger\hat{x}_g + \Pi_g(S_g - K_gL_g)\tilde{\Pi}_g^\dagger\tilde{x}_g + Bu\\ &= A\hat{x}_g + Bu,\\ \dot{\tilde{x}}_g &= \tilde{\Pi}_g(S_g - K_gL_g)\tilde{\Pi}_g^\dagger\tilde{x}_g + \tilde{\Pi}_g(S_g - K_gL_g)\Pi_g^\dagger\hat{x}_g\\ &\quad + \tilde{\Pi}_gK_gu\\ &=: \tilde{A}_g\tilde{x}_g + \tilde{D}_g\hat{x}_g + \tilde{B}_gu,\\ y &= C\hat{x}_g,\end{aligned} \tag{21}$$

where $\tilde{\Pi}_g$ is such that the matrix

$$O_g := \begin{bmatrix}(\Pi_g\Pi_g^T)^{-1/2}\Pi_g\\ \tilde{\Pi}_g\end{bmatrix}$$

is orthogonal (that is $O_g^TO_g = I$), so that

$$T_g^{-1} := \begin{bmatrix}\Pi_g\\ \tilde{\Pi}_g\end{bmatrix}^{-1} = \begin{bmatrix}\Pi_g^T(\Pi_g\Pi_g^T)^{-1} & \tilde{\Pi}_g^T\end{bmatrix} =: \begin{bmatrix}\Pi_g^\dagger & \tilde{\Pi}_g^\dagger\end{bmatrix}, \tag{22}$$

and we have utilized the fact that

$$\Pi_g(S_g - K_gL_g)\Pi_g^\dagger = (\Pi_gS_g - BL_g)\Pi_g^\dagger = A\Pi_g\Pi_g^\dagger = A,$$
$$\Pi_g(S_g - K_gL_g)\tilde{\Pi}_g^\dagger = (\Pi_gS_g - BL_g)\tilde{\Pi}_g^\dagger = A\Pi_g\tilde{\Pi}_g^\dagger = 0.$$

This implies that a necessary and sufficient condition to match sufficiently many other moments of the system (1) is that the approximation (12) contains the system (1) as an internal model. It also implies that n forward moments (the forward moment with regards to a signal generator with dimension n) are sufficient to build the basic structure (12) for the realization.

In the next section, we provide guidelines on the selection of the design parameter K_g that satisfies (18) (hence to realize the system (1) by constructing an approximation (12) that contains the system (1) as an internal model), using only the knowledge of the signal generator, the filter, and the corresponding moments. We note that in the selection process of K_g, we should also guarantee that

$$\tilde{A}_g = \tilde{\Pi}_g(S_g - K_gL_g)\tilde{\Pi}_g^\dagger \tag{23}$$

is Hurwitz for a stable pole-zero cancellation.

Remark 4: One can easily derive similar results for the backward moments. □

IV. Guidelines on the Selection of the Parameter

Note first that for any parameter $\widehat{K}_g$ that satisfies (18), i.e., $\Pi_g\widehat{K}_g = B$, all K_g can be parameterized as

$$K_g = \widehat{K}_g + \tilde{K}_g,$$

with the additional parameter $\tilde{K}_g$ which is such that $\Pi_g\tilde{K}_g = 0$. Since $N_g \geq n$ and $\mathrm{rank}(\Pi_g) = n$ (see Lemma 4), at least one such $\widehat{K}_g$ exists, e.g., $\Pi_g^\dagger B$.

[2] By Assumptions 1 and 2, we have that $\mathrm{rank}(\Pi_g) = n$ (see Lemma 4).

Therefore, in the first part of this section we provide guidelines to find $\widehat{K}_g$ using only the knowledge of the signal generator, the filter, and the corresponding moments.

For this purpose, recall that the backward moment of the system (1) has the form: $M_f = \Pi_f B$, where Π_f is the unique solution of the Sylvester equation (5). This implies that the backward moment contains indirect information of B ($\text{rank}(\Pi_f) = n$), and thus, can be used to find $\widehat{K}_g$ or even can be a suitable candidate for $\widehat{K}_g$.

However, for the backward moment to be a suitable candidate, it must satisfy the condition $\Pi_g M_f = \Pi_g \Pi_f B = B$, and this significantly restricts the choice of the filter (3).

In this regard, we suggest an additional step, which is to find an appropriate matrix D such that $\widehat{K}_g = DM_f$ satisfies the condition $\Pi_g \widehat{K}_g = B$. Among the infinitely many possible choices, we select

$$D = \Pi_g^T(\Pi_g\Pi_g^T)^{-1}(\Pi_f^T\Pi_f)^{-1}\Pi_f^T,$$

which is the Moore-Penrose inverse of $\Pi_m := \Pi_f\Pi_g$, denoted as $\Pi_m^\dagger$. It satisfies the condition $\Pi_g\widehat{K}_g = B$ since

$$\begin{aligned}\Pi_g\widehat{K}_g &= \Pi_g\Pi_m^\dagger M_f = \Pi_g\Pi_m^\dagger\Pi_f B\\ &= \Pi_g\Pi_g^T(\Pi_g\Pi_g^T)^{-1}(\Pi_f^T\Pi_f)^{-1}\Pi_f^T\Pi_f B = B.\end{aligned}$$

This additional step allows selecting any filter (3) that satisfies Assumptions 1 and 2.

Note that $\Pi_m = \Pi_f\Pi_g$ satisfies the Sylvester equation

$$\Pi_m S_g + (L_f M_g - M_f L_g) = S_f \Pi_m, \quad (24)$$

which contains only the knowledge of (S_g, L_g, M_g) and (S_f, L_f, M_f). This means that if the signal generator (2) and the filter (3) satisfy the condition $\sigma(S_g) \cap \sigma(S_f) = \emptyset$, then $\Pi_m = \Pi_f\Pi_g$ becomes the unique solution of the Sylvester equation (24), hence can be found using only the knowledge of the signal generator (S_g, L_g), of the filter (S_f, L_f), and of the corresponding moments M_g and M_f.

Lemma 3: Under Assumptions 1 and 2, the matrix $\Pi_m = \Pi_f\Pi_g$ satisfies the equation (24), where Π_g and Π_f are the unique solution of the Sylvester equations (4) and (5), respectively. □

Proof: By multiplying Π_f to the left of the Sylvester equation (4) we have that

$$\Pi_f A\Pi_g + \Pi_f B L_g = \Pi_f\Pi_g S_g,$$

and by multiplying Π_g to the right of the Sylvester equation (5) we have that

$$\Pi_f A\Pi_g + L_f C\Pi_g = S_f\Pi_f\Pi_g,$$

from which we obtain

$$\Pi_f\Pi_g S_f + (L_f C\Pi_g - \Pi_f B L_g) = S_f\Pi_f\Pi_g.$$

■

This implies that one can find an approximation of the system (1) having stable pole-zero cancellation as

$$\begin{aligned}\dot{\chi}_g &= (S_g - (\Pi_m^\dagger M_f + \tilde{K}_g)L_g)\chi_g + (\Pi_m^\dagger M_f + \tilde{K}_g)u \in \mathbb{R}^{N_g},\\ y &= M_g\chi_g \qquad\qquad \in \mathbb{R},\end{aligned} \quad (25)$$

where $\tilde{K}_g$ is a free design parameter to be selected so that $\tilde{A}_g$ in (23) is Hurwitz and such that $\Pi_g\tilde{K}_g = 0$.

Here, the constraint $\Pi_g\tilde{K}_g = 0$ can be rephrased as $\Pi_m\tilde{K}_g = \Pi_f\Pi_g\tilde{K}_g = 0$ ($\text{rank}(\Pi_f) = n$), to be described by only the given knowledge of the signal generator, the filter, and the corresponding moments. Accordingly, the constraint to make $\tilde{A}_g$ Hurwitz can be rephrased as to make the matrix

$$\tilde{S}_g(q) := S_g - \Pi_m^\dagger\Pi_m(S_g + qI) - \tilde{K}_g L_g, \quad q \geq 0 \quad (26)$$

Hurwitz. This is because

$$T_g\tilde{S}_g(q)T_g^{-1} = \begin{bmatrix} -qI & 0 \\ * & \tilde{\Pi}_g(S_g - \tilde{K}_g L_g)\tilde{\Pi}_g^\dagger \end{bmatrix}$$

and

$$\begin{aligned}\tilde{\Pi}_g(S_g - \tilde{K}_g L_g)\tilde{\Pi}_g^\dagger &= \tilde{\Pi}_g(S_g - (\Pi_m^\dagger M_f + \tilde{K}_g)L_g)\tilde{\Pi}_g^\dagger\\ &= \tilde{\Pi}_g(S_g - K_g L_g)\tilde{\Pi}_g^\dagger = \tilde{A}_g,\end{aligned}$$

where T_g is defined as in (22) and we have utilized the fact that $\Pi_g\Pi_m^\dagger\Pi_m = \Pi_g$, $\Pi_g\tilde{K}_g = 0$, and $\tilde{\Pi}_g\Pi_m^\dagger = 0$.

Since $(L_g\tilde{\Pi}_g^\dagger, \tilde{\Pi}_g S_g\tilde{\Pi}_g^\dagger)$ is observable[3] there exists at least one $\check{K}_g \in \mathbb{R}^{N_g - n}$ that makes $\tilde{\Pi}_g S_g\tilde{\Pi}_g^\dagger - \check{K}_g L_g\tilde{\Pi}_g^\dagger$ Hurwitz. As a result, there exists at least one $\tilde{K}_g = \tilde{\Pi}_g^\dagger\check{K}_g$ that satisfies the constraint $\Pi_g\tilde{K}_g = 0$ and that makes $\tilde{A}_g$ Hurwitz ($\tilde{\Pi}_g\tilde{K}_g = \check{K}_g$).

In this regard, we provide guidelines on the realization of a linear system using $N_g + N_f$ moments (the forward moment with regards to the signal generator (2) with dimension N_g and the backward moment with regards to the filter (3) with dimension N_f) as illustrated in what follows.

1) Find $\widehat{K}_g$ from Π_m, which is the solution of the Sylvester equation (24) that requires only the knowledge of the signal generator, the filter, and the corresponding moments.
2) Find $\tilde{K}_g$ so that $\tilde{S}_g(q)$ in (26) is Hurwitz and $\Pi_m\tilde{K}_g = 0$.

Remark 5: Given the knowledge of the system dimension n, the guideline given in this section simplifies to

1) let $N_g = N_f = n$ and find the solution Π_m of (24).

Then, the unique K_g can be found as $\Pi_m^{-1}M_f$: this gives the realization of the original system utilizing exactly $2n$ moments. □

Remark 6: One can easily derive similar results for the backward moments. □

V. CONCLUSION

The existence condition of the forward moment is equivalent to the observability condition of the auxiliary system having the forward moment as a state variable, which can be derived from the interconnection of the signal generator (2) and the system (1). This gives the intuition that the provided guidelines to realize a linear system using moments can be translated into guidelines utilizing only the input-output data given by the corresponding interconnection. In addition to this, an extension to the nonlinear case is of future interest.

[3]This is because, from the observability of (L_g, S_g), we have the observability of $(L_g, S_g - \widehat{K}_g L_g)$, which implies the observability of $(L_g\tilde{\Pi}_g^\dagger, \tilde{\Pi}_g(S_g - \widehat{K}_g L_g)\tilde{\Pi}_g^\dagger)$ since we have $\Pi_g(S_g - \widehat{K}_g L_g)\tilde{\Pi}_g^\dagger = 0$ according to (21).

References

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. SIAM Advances in Design and Control, 2005.

[2] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2321–2336, 2010.

[3] A. C. Antoulas, "A new result on passivity preserving model reduction," *Systems & Control Letters*, vol. 54, pp. 361–374, 2005.

[4] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX + XB = C$," *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.

Appendix

In the Appendix, we study the existence, uniqueness, and rank of the solutions $\Pi \in \mathbb{R}^{n\times N}$ of the Sylvester equation:

$$\Pi S = A\Pi + BL, \tag{27}$$

where $S \in \mathbb{R}^{N\times N}$, $L \in \mathbb{R}^{1\times N}$, $A \in \mathbb{R}^{n\times n}$, and $B \in \mathbb{R}^{n\times 1}$. For this purpose, we perform a "controllable" decomposition and an "observable" decomposition of the pair (A, B) and (L, S), respectively, from which we have coordinates transformation matrices $T = \mathrm{col}(\bar{T}, \tilde{T}) \in \mathbb{R}^{[\bar{n}+(n-\bar{n})]\times n}$ and $P = \mathrm{col}(\bar{P}, \tilde{P}) \in \mathbb{R}^{[\bar{N}+(N-\bar{N})]\times N}$ such that

$$\begin{bmatrix}\bar{T}\\ \tilde{T}\end{bmatrix} A \begin{bmatrix}\bar{T}^\dagger & \tilde{T}^\dagger\end{bmatrix} = \begin{bmatrix}\bar{A} & \tilde{A}_{12}\\ 0 & \tilde{A}_{22}\end{bmatrix}, \quad \begin{bmatrix}\bar{T}\\ \tilde{T}\end{bmatrix} B = \begin{bmatrix}\bar{B}\\ 0\end{bmatrix},$$

and the pair $(\bar{A}, \bar{B})$ is controllable and

$$\begin{bmatrix}\bar{P}\\ \tilde{P}\end{bmatrix} S \begin{bmatrix}\bar{P}^\dagger & \tilde{P}^\dagger\end{bmatrix} = \begin{bmatrix}\bar{S} & 0\\ \tilde{S}_{21} & \tilde{S}_{22}\end{bmatrix}, \quad L\begin{bmatrix}\bar{P}^\dagger & \tilde{P}^\dagger\end{bmatrix} = \begin{bmatrix}\bar{L} & 0\end{bmatrix},$$

and the pair $(\bar{L}, \bar{S})$ is observable, respectively, where

$$T^{-1} =: \begin{bmatrix}\bar{T}^\dagger & \tilde{T}^\dagger\end{bmatrix} \quad \text{and} \quad P^{-1} =: \begin{bmatrix}\bar{P}^\dagger & \tilde{P}^\dagger\end{bmatrix}.$$

Proposition 3 ([4]): If $\sigma(A) \cap \sigma(S) = \emptyset$, then for any $C \in \mathbb{R}^{n\times N}$, the Sylvester equation $\Pi S = A\Pi + C$ has a unique solution. On the other hand, if there exists $C \in \mathbb{R}^{n\times N}$ such that the Sylvester equation $\Pi S = A\Pi + C$ has a unique solution Π, then $\sigma(A) \cap \sigma(S) = \emptyset$. □

Lemma 4: There exists a matrix $\bar{\Pi} \in \mathbb{R}^{\bar{n}\times\bar{N}}$ such that

$$\bar{\Pi}\bar{S} = \bar{A}\bar{\Pi} + \bar{B}\bar{L} \tag{28}$$

only if $\sigma(\bar{A}) \cap \sigma(\bar{S}) = \emptyset$. If it exists, then it is unique and $\mathrm{rank}(\bar{\Pi}) = \min\{\bar{n}, \bar{N}\}$. □

Proof: Assume that there exists $\bar{\Pi}$ such that (28) is satisfied and $\sigma(\bar{A}) \cap \sigma(\bar{S}) \neq \emptyset$, i.e., there exists $\lambda \in \sigma(\bar{A}) \cap \sigma(\bar{S}) \subset \mathbb{C}$. Then, there exist non-zero vectors $v \in \mathbb{C}^{\bar{n}}$ and $w \in \mathbb{C}^{\bar{N}}$ such that $\bar{S}w = \lambda w$ and $v^\dagger \bar{A} = \lambda v^\dagger$. This implies that $\lambda v^\dagger \bar{\Pi} w = v^\dagger \bar{\Pi}\bar{S}w = v^\dagger \bar{A}\bar{\Pi}w + v^\dagger \bar{B}\bar{L}w = \lambda v^\dagger \bar{\Pi} w + v^\dagger \bar{B}\bar{L}w$, from which we obtain that $v^\dagger \bar{B}\bar{L}w = 0$. This implies that either $v^\dagger \bar{B} = 0$ or $\bar{L}w = 0$. Without loss of generality assume that $v^\dagger \bar{B} = 0$ then, we obtain that $v^\dagger \bar{A}^i \bar{B} = \lambda^i v^\dagger \bar{B} = 0$, which implies that $v^\dagger \begin{bmatrix}\bar{B} & \bar{A}\bar{B} & \cdots & \bar{A}^{\bar{n}-1}\bar{B}\end{bmatrix} = 0$, hence $v = 0$, which is a contradiction.

Now, if $\mathrm{rank}(\bar{\Pi}) = \bar{r} < \min\{\bar{n}, \bar{N}\}$, then there exist orthonormal matrices $\bar{U} \in \mathbb{R}^{\bar{n}\times\bar{r}}$, $\bar{V} \in \mathbb{R}^{\bar{N}\times\bar{r}}$, and a nonsingular diagonal matrix $\bar{\Lambda} \in \mathbb{R}^{\bar{r}\times\bar{r}}$, such that $\bar{\Pi} = \bar{U}\bar{\Lambda}\bar{V}^T$. Let $\tilde{U} \in \mathbb{R}^{\bar{n}\times(\bar{n}-\bar{r})}$ and $\tilde{V} \in \mathbb{R}^{\bar{N}\times(\bar{N}-\bar{r})}$ be such that $\begin{bmatrix}\bar{U} & \tilde{U}\end{bmatrix}$ and $\begin{bmatrix}\bar{V} & \tilde{V}\end{bmatrix}$ are orthogonal. Then, we have $\tilde{U}^T\bar{B}\bar{L}\tilde{V} = 0$. This implies that either $\tilde{U}^T\bar{B} = 0$ or $\bar{L}\tilde{V} = 0$. Without loss of generality assume that $\tilde{U}^T\bar{B} = 0$ then, since $(\bar{A}, \bar{B})$ is controllable, we also have that the pair

$$\begin{aligned}&\left(\begin{bmatrix}\bar{U}^T\\ \tilde{U}^T\end{bmatrix}(\bar{A} + \bar{B}\bar{L}\bar{V}\bar{\Lambda}^{-1}\bar{U}^T)\begin{bmatrix}\bar{U} & \tilde{U}\end{bmatrix}, \begin{bmatrix}\bar{U}^T\bar{B}\\ \tilde{U}^T\bar{B}\end{bmatrix}\right)\\ &= \left(\begin{bmatrix}\bar{U}^T(\bar{A} + \bar{B}\bar{L}\bar{V}\bar{\Lambda}^{-1}\bar{U}^T)\bar{U} & \bar{U}^T\bar{A}\tilde{U}\\ 0 & \tilde{U}^T\bar{A}\tilde{U}\end{bmatrix}, \begin{bmatrix}\bar{U}^T\bar{B}\\ 0\end{bmatrix}\right)\end{aligned}$$

is controllable, which is a contradiction. Here, we have utilized the fact that $\tilde{U}^T\bar{A}\bar{U}\bar{\Lambda}\bar{V}^T + \tilde{U}^T\bar{B}\bar{L} = \tilde{U}^T(\bar{A}\bar{\Pi} + \bar{B}\bar{L}) = \tilde{U}^T\bar{\Pi}\bar{S} = 0$, hence $\tilde{U}^T(\bar{A}\bar{U} + \bar{B}\bar{L}\bar{V}\bar{\Lambda}^{-1})\bar{\Lambda}\bar{V}^T = \tilde{U}^T\bar{A}\bar{U}\bar{\Lambda}\bar{V}^T + \tilde{U}^T\bar{B}\bar{L} - \tilde{U}^T\bar{B}\bar{L}\tilde{V}\tilde{V}^T = 0$. Therefore, $\mathrm{rank}(\bar{\Pi}) = \min\{\bar{n}, \bar{N}\}$. ■

Lemma 5: Any solution $\Pi \in \mathbb{R}^{n\times N}$ of the Sylvester equation (27) satisfies

$$\mathrm{rank}(\Pi) \geq \min\{\bar{n}, \bar{N}\}.$$

Moreover, there exists a solution $\Pi \in \mathbb{R}^{n\times N}$ of the Sylvester equation (27) such that $\mathrm{rank}(\Pi) = \min\{\bar{n}, \bar{N}\}$ if and only if $\sigma(\bar{A}) \cap \sigma(\bar{S}) = \emptyset$. If it exists, then it satisfies

$$\Pi = \bar{T}^\dagger\bar{\Pi}\bar{P},$$

where $\bar{\Pi} \in \mathbb{R}^{\bar{n}\times\bar{N}}$ is the unique solution of the Sylvester equation (28), hence Π is also unique. □

Proof: Let Π be a solution of the Sylvester equation (27). Then, we have $r := \mathrm{rank}(\Pi) \geq \min\{\bar{n}, \bar{N}\}$. This is because there exist orthonormal matrices $U \in \mathbb{R}^{n\times r}$, $V \in \mathbb{R}^{N\times r}$, and a nonsingular diagonal matrix $\Lambda \in \mathbb{R}^{r\times r}$, such that $\Pi = U\Lambda V^T$. In particular, let $\tilde{U} \in \mathbb{R}^{n\times(n-r)}$ and $\tilde{V} \in \mathbb{R}^{N\times(N-r)}$ be such that $\begin{bmatrix}U & \tilde{U}\end{bmatrix}$ and $\begin{bmatrix}V & \tilde{V}\end{bmatrix}$ are orthogonal. Then, we have that $\tilde{U}^TBL\tilde{V} = \tilde{U}^T(\Pi S - A\Pi)\tilde{V} = \tilde{U}^TU\Lambda V^TS - AU\Lambda V^T\tilde{V} = 0$. This implies that either $\tilde{U}^TB = 0$ or $L\tilde{V} = 0$. Without loss of generality assume that $\tilde{U}^TB = 0$ then, we obtain that $\tilde{U}^TAU = \tilde{U}^TA\Pi V\Lambda^{-1} = \tilde{U}^T(\Pi S - BL)V\Lambda^{-1} = 0$. Therefore, $r \geq \bar{n} \geq \min\{\bar{n}, \bar{N}\}$.

($\Leftarrow$) If $\sigma(\bar{A}) \cap \sigma(\bar{S}) = \emptyset$, then, by Proposition 3, there exists a unique solution $\bar{\Pi}$ of the Sylvester equation (28). Then, $\Pi = \bar{T}^\dagger\bar{\Pi}\bar{P}$ ($\mathrm{rank}(\Pi) = \min\{\bar{n}, \bar{N}\}$) satisfies the equation

$$\begin{aligned}\Pi S &= \bar{T}^\dagger\bar{\Pi}\bar{P}S = \bar{T}^\dagger\bar{\Pi}\bar{S}\bar{P} = \bar{T}^\dagger\bar{A}\bar{\Pi}\bar{P} + \bar{T}^\dagger\bar{B}\bar{L}\bar{P}\\ &= \bar{T}^\dagger\bar{T}A\bar{T}^\dagger\bar{\Pi}\bar{P} + \bar{T}^\dagger\bar{T}BL\bar{P}^\dagger\bar{P}\\ &= A\Pi + BL - \tilde{T}^\dagger\tilde{T}A\bar{T}^\dagger\bar{\Pi}\bar{P} - \tilde{T}^\dagger\tilde{T}BL\bar{P}^\dagger\bar{P} - BL\tilde{P}^\dagger\tilde{P}\\ &= A\Pi + BL,\end{aligned}$$

hence is a solution of the Sylvester equation (27).

($\Rightarrow$) If there exists a solution Π of the Sylvester equation (27) such that $\mathrm{rank}(\Pi) = \min\{\bar{n}, \bar{N}\}$, then by the former argument, there exists U and V. Note that we have $\tilde{T}U = 0$ and $V^T\tilde{P}^\dagger = 0$. This implies $\tilde{T}\Pi = 0$ and $\Pi\tilde{P}^\dagger = 0$. Now, from $T\Pi P^{-1}PSP^{-1} = TAT^{-1}T\Pi P^{-1} + TBLP^{-1}$, we obtain that $\left(\bar{T}\Pi\bar{P}^\dagger\right)\bar{S} = \bar{A}\left(\bar{T}\Pi\bar{P}^\dagger\right) + \bar{B}\bar{L}$, and thus, by Lemma 4, we obtain that $\sigma(\bar{A}) \cap \sigma(\bar{S}) = \emptyset$ and $\bar{T}\Pi\bar{P}^\dagger = \bar{\Pi}$. Together with $\tilde{T}\Pi = 0$ and $\Pi\tilde{P}^\dagger = 0$, this implies $\Pi = \bar{T}^\dagger\bar{\Pi}\bar{P}$. ■