

# A Moments-Based Approach To Estimation and Data Interpolation for a Class of Wiener Systems

Mustafa Ayazoglu

Mario Sznaier

Constantino Lagoa

Octavia Camps

**Abstract**—This paper addresses the problems of estimating the values of both the outputs and the internal signals for a class of Wiener systems consisting of the cascade of an unknown linear time invariant systems and a known, rational, generically non-invertible nonlinearity, based solely on past input/output data corrupted by noise. This situation arises in many scenarios of practical interest where an explicit a-priori model of the linear system is not available. Examples include extracting geometric 3D structure from a sequence of 2D images (structure from motion), and nonlinear dimensionality reduction via manifold embedding. The main result of the paper is a simple, computationally efficient algorithm that is capable of handling intermittent measurements and does not entail identifying first the unknown linear dynamics. Rather, the problem of estimating the internal signals and interpolating missing data is recast into a rank-constrained feasibility problem. Although this problem depends polynomially in the data, we show that, by appealing to classical results on moments optimization, it can be reduced to a rank-constrained Linear Matrix Inequality optimization and efficiently solved using existing techniques. The potential of the proposed approach is illustrated by solving structure from motion problems using real data.

## I. INTRODUCTION AND MOTIVATION

Many processes of practical interests can be modeled as the output of a Wiener system, consisting of the cascade of a memoryless, static nonlinearity and a Linear Time Invariant (LTI) plant. Examples include domains as varied as communications [4], biology [2], nonlinear dimensionality reduction [21], and computer vision [14]. Consider for instance the problem of estimating the 3D geometry of a scene, using 2 dimensional data generated by a moving perspective camera that has a focal length  $f$  and a principal point with coordinates  $(c_u, c_v)$ . This situation can be modeled by the following system (see for instance [14]):

$$\begin{aligned} \zeta_{k+1} &\doteq \mathbf{A}\zeta_k + \mathbf{B}e_k \\ \mathbf{P}_{kj} &\doteq \begin{bmatrix} X_{kj} \\ Y_{kj} \\ Z_{kj} \end{bmatrix} = \mathbf{C}_j \zeta_k \\ u_{kj} &= f \frac{X_{kj}}{Z_{kj}} - c_u, \\ v_{kj} &= f \frac{Y_{kj}}{Z_{kj}} - c_v \end{aligned} \quad (1)$$

This work was supported in part by NSF grants ECCS-0731224, IIS-0713003, CMMI-0838906 and ECCS-0901433; AFOSR grant FA9550-09-1-0253; and DHS grant 2008-ST-061-ED0001. M. Ayazoglu, O. Camps and M. Sznaier are with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, 02115. Constantino Lagoa is with the Department of Electrical Engineering, Penn State University, University Park, PA 16802

Here the pair  $(\mathbf{A}, \mathbf{B})$  and the associated state vector  $\zeta_k$  model the 3D motion of the camera in response to the input  $e^a$ ,  $\mathbf{P}_{kj}$ ,  $j = 1, \dots, N$ , denotes the 3D coordinates of point  $P_j$  at time  $k$ ,  $i = 1, \dots, F$ , and  $(u_{kj}, v_{kj})$  denotes the 2D coordinates of the image of  $\mathbf{P}_{kj}$ . Clearly, this model above is a special case of a Wiener system, where the nonlinearity is rational. In this context, the problem of reconstructing the 3D geometry of the scene can be formalized as a nonlinear filtering one: estimating  $\mathbf{P}_{kj}$  from  $(u_{kj}, v_{kj})$ , potentially corrupted by measurement noise. In principle, this nonlinear filtering problem can be solved using the techniques proposed in [14], [8], [1], [6], [9]. However, proceeding in this fashion requires first obtaining a model of the Wiener system. Note that in the application above, if the camera is calibrated, the nonlinearity is perfectly known, and only the linear portion of the system (e.g. a pair  $(\mathbf{A}, \mathbf{B})$  and the projection matrices  $\mathbf{C}_j$ ,  $j = 1, \dots, N_p$ ) needs to be identified. However, as shown in [19], even this simpler case leads to NP-hard problems in the presence of persistent measurement noise. A further complication that prevents the use of classical filtering techniques arises from the fact that data is often missing: not all points appear in all frames, due for instance to occlusion or misidentified correspondences. The goal of this paper is to develop a computationally tractable framework for data interpolation/estimation for a class of Wiener systems that embodies the main features of problem described above: an unknown linear component of the system and noisy, fragmented measurements. Our main result shows that in principle this problem can be recast into minimizing the rank of a matrix that depends polynomially on the measured data. As we show in the paper, exploiting some classical results on moment optimization [12] allows for recasting the problem into a rank-constrained LMI optimization that can be efficiently solved using existing techniques [7], [17]. These results are illustrated in a problem that has been the object of considerably attention in the computer vision community, 3D reconstruction from video, where the proposed approach outperforms existing techniques.

## II. PRELIMINARIES

### A. Notation

$$\begin{aligned} x_k & \text{ } k^{\text{th}} \text{ element of a vector } \mathbf{x}. \\ \|\mathbf{x}\|_\infty & \text{ } \sup \text{ norm of } \mathbf{x} \in R^n: \|\mathbf{x}\|_\infty \doteq \\ & \max_{1 \leq k \leq n} |x_k| \end{aligned}$$

<sup>a</sup>Typically  $\eta_k$  contains the past values of the camera position, velocity, etc.

$\ell^\infty$	Space of vector sequences equipped with the norm: $\ x\ _{\ell^\infty} \doteq \sup_i \ x_i\ _\infty$
$\mathbf{A}^T$	conjugate transpose of matrix $\mathbf{A}$ .
$\mathbf{A} \succ 0$ ( $\succeq$ )	$\mathbf{A} = \mathbf{A}^T$ is positive (semi)definite.
$\mathbf{A} \succ \mathbf{B}$	$(\mathbf{A} - \mathbf{B}) \succ 0$
$\text{rank}(\mathbf{A})$	rank of matrix $\mathbf{A}$ .

Given a vector sequence  $\{y_k\}_{k=1}^{n+l-1}$ , in the sequel we denote its corresponding Hankel matrix as:

$$\mathbf{H}_{n,l}(\mathbf{y}) \doteq \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_l \\ \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_{l+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_n & \mathbf{y}_{n+1} & \cdots & \mathbf{y}_{l+n-1} \end{bmatrix}$$

Moreover,  $\mathbf{H}_{n,n}(\cdot)$  will be simply denoted by  $\mathbf{H}_n(\cdot)$ .

### III. PROBLEM STATEMENT

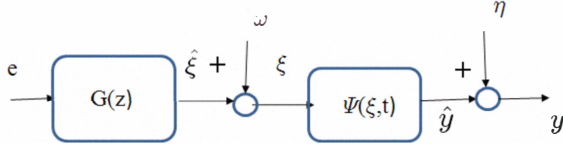


Fig. 1. Wiener System Structure

Consider the (discrete time) Wiener system shown in Figure 1 consisting of the interconnection of a LTI system  $G(z)$  and a memoryless nonlinearity  $\psi(\cdot)$ . The corresponding equations are given by:

$$\begin{aligned} \xi_t &= (g * e)_t + \omega_t \\ y_t &= \psi(\xi_t, t) + \eta_t \end{aligned} \quad (2)$$

where  $g$  is the convolution kernel of  $G(\cdot)$ ,  $*$  denotes convolution and the signals  $e \in R^{n_e}$  and  $y \in R^{n_y}$  represent a known input and its corresponding output, corrupted by measurement noise  $\eta$ . The signal  $\omega$  represents an (unmeasurable) disturbance, that can be used to model for instance noise at the input of the nonlinearity, or temporal correlations in the outputs of the nonlinearity. Finally,  $\Psi(\cdot, t)$  represents a memoryless, potentially time-varying nonlinearity. In this paper we will further assume that  $\Psi$  is rational, e.g.

$$\Psi(x, t) = \frac{P(x, t)}{Q(x, t)} \quad (3)$$

where  $P$  and  $Q$  are multivariate time-varying polynomials of the form:

$$P(x, t) = \sum_{\alpha} p_{\alpha,t} \mathbf{x}^{\alpha}, \quad Q(x, t) = \sum_{\alpha} q_{\alpha,t} \mathbf{x}^{\alpha} \quad (4)$$

where  $\mathbf{x}^{\alpha} \doteq x_1^{m_1} x_2^{m_2} \dots x_n^{m_n}$ . Then, in its simplest form, the problem addressed in this paper can be stated as:

**Problem 1:** [Consistency and Estimation] Given:

- 1) A system of the form shown in Figure 1 consisting of the interconnection of a known nonlinearity of the form (3) and an unknown linear plant  $G$ , with McMillan degree bounded by some known  $n_g$ .

- 2) *A priori* information consisting of a set membership descriptions of the measurement noise  $\eta \in \mathcal{N}$  and the disturbance  $\omega \in \mathcal{W}$ .
- 3) *A posteriori* experimental information consisting of  $n_m$  input/output measurements  $\{e_i, y_i\}_{i=k-n_m+1}^k$ .

Then:

- 1) Establish whether the experimental data is consistent with the a-priori assumptions.
- 2) If so, estimate the values of  $\xi_i$  and  $\hat{y}_i$ ,  $i = 1, \dots, t$ .

That is, if the model is not invalidated by the data, we want to estimate the values of both, the external and internal signals, based on the measurements  $y_i$ . A variation of this problem that will be discussed later includes the case where the measurements  $y_i$  are not available for  $i$  in some set  $\mathcal{I}$ .

#### A. The Problem of Moments

The key step in establishing our main result is to transform the problem from an optimization over noise sequences to an optimization over probability measures. To this effect, we will use several classical results establishing the necessary and sufficient conditions for a given sequence to be the moments of a probability distribution. For completeness, these results are briefly quoted below. Details are given for instance in [5], [12].

Let  $\mathbf{K}$  be a closed subset of  $\mathfrak{R}^D$  and let  $\alpha$  be a multi-index (i.e.  $\alpha \in \mathbb{N}^D$ ) representing the powers of a monomial in  $D$  variables. Given a sequence of scalars  $\{m_{\alpha}\}$ , the **K-moment problem** is to determine whether there exists a probability measure  $\mu$  supported on  $\mathbf{K}$  such that it has each  $m_{\alpha}$  as its  $\alpha^{\text{th}}$  moment. That is:

$$m_{\alpha} = \mathbf{E}_{\mu}(\mathbf{x}^{\alpha}) = \int_{\mathbf{K}} \mathbf{x}^{\alpha} \mu(dx) \quad (5)$$

where  $\mathbf{x}^{\alpha} = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_D^{\alpha_D}$ . In particular, in the sequel we are interested in probability measures that are supported on balls of radius  $\epsilon$  centered at the origin. Next, we recall a necessary and sufficient condition for the existence of such a measure [15].

**Theorem 1:** Let  $p = \sum_{\alpha} c_{\alpha} \mathbf{x}^{\alpha} \in \mathcal{P}$  denote a generic (multivariate) polynomial. Given a sequence  $\{m_{\alpha}\}$ , there exists a linear functional  $\mathbf{E}: \mathcal{P} \rightarrow \mathcal{R}$  such that

$$\mathbf{E}(p) = \sum_{\alpha} c_{\alpha} m_{\alpha} \quad (6)$$

and  $\{m_{\alpha}\}$  are the moments of a distribution supported on  $\|\mathbf{x}\|_2 \leq \epsilon$ , if and only if the following two conditions hold for all  $p \in \mathcal{P}$ :

$$\mathbf{E}(p^2) \geq 0 \quad (7)$$

$$\mathbf{E}((\epsilon^2 - (x_1^2 + \dots + x_D^2))p^2) \geq 0 \quad (8)$$

**Remark 1:** The conditions given in the above theorem consist of infinite semidefinite quadratic forms which can be converted into (infinite) linear matrix inequalities (LMIs) in the moment variables  $\{m_{\alpha}\}$ .

Next, we briefly discuss how to build a matrix representation of a given sequence  $\mathbf{m}$  that contains all the moments up to order  $2\delta$ . Assume for simplicity that the moments are

ordered according to a graded reverse lexicographic order (grevlex) of the corresponding monomials so that we have  $\mathbf{0} = \boldsymbol{\alpha}^{(1)} < \dots < \boldsymbol{\alpha}^{(M_\delta)}$ , where  $M_\delta \doteq \binom{\delta + D}{D}$  is the number of moments in  $\mathfrak{R}^D$  up to order  $\delta$ . Then, the moment conditions take the form:

$$\begin{aligned} \mathbf{L}^{(\delta)}(\mathbf{m}) &\succeq 0 \\ \mathbf{K}^{(\delta)}(\epsilon, \mathbf{m}) &\succeq 0 \end{aligned} \quad (9)$$

where

$$\begin{aligned} \mathbf{L}^{(\delta)}(i, j) &= m_{\boldsymbol{\alpha}^{(i)} + \boldsymbol{\alpha}^{(j)}} \text{ for all } i, j \leq M_\delta \\ \mathbf{K}^{(\delta)}(i, j) &= (\epsilon^2 m_{\boldsymbol{\alpha}^{(i)} + \boldsymbol{\alpha}^{(j)}} - m_{\boldsymbol{\alpha}^{(i)} + \boldsymbol{\alpha}^{(j)} + (2, 0, \dots, 0)} - \dots - m_{\boldsymbol{\alpha}^{(i)} + \boldsymbol{\alpha}^{(j)} + (0, \dots, 0, 2)}) \text{ for all } i, j \leq M_\delta - 1 \end{aligned}$$

It can be shown [12] that the linear matrix inequalities (9) are necessary conditions for the existence of a measure  $\mu$  supported in the  $\epsilon$ -ball that has the sequence  $\mathbf{m}$  as its moments. Moreover, as  $\delta \uparrow \infty$ , (9) becomes equivalent to conditions (7)–(8) in Theorem 1; hence it is (asymptotically) sufficient as well. Thus, increasing the size of the moment matrices yields progressively better finite-dimensional approximations to the infinite dimensional conditions (7)–(8) [12]. It is also worth noting that if as  $\delta$  increases, the rank of the moment matrices stops increasing, the so-called *flat extension* property ([5]) is satisfied. In this case, the finite dimensional conditions (9) corresponding to this value of  $\delta$  are necessary and sufficient for the existence of a measure supported in the  $\epsilon$  ball.

#### IV. CONSISTENCY AND ESTIMATION AS A RANK CONSTRAINED FEASIBILITY PROBLEM

Next, we show how Problem 1 can be reduced to a rank constrained feasibility form. Given an input/output sequence  $\{\mathbf{e}_k, \mathbf{y}_k\}_{k=1}^m$ , define the consistency set:

$$\begin{aligned} \mathcal{T}(\mathbf{y}) &\doteq \left\{ \begin{array}{l} \{\boldsymbol{\xi}_k\}_{k=1}^m: \quad \boldsymbol{\xi}_k = (G * \mathbf{e})_k + \boldsymbol{\omega}_k \\ \quad \quad \quad \mathbf{y}_k = N(\boldsymbol{\xi}_k, k) + \boldsymbol{\eta}_k \\ \text{for some sequences } \boldsymbol{\eta}_k \in \mathcal{N}, \boldsymbol{\omega}_k \in \mathcal{W} \text{ and some} \\ \text{LTI system } G \text{ with McMillan degree } n_g \end{array} \right\} \end{aligned}$$

Intuitively, this is the set of all (unknown) inputs to the nonlinearity that are compatible with both, existing a-priori information and a-posteriori measurements, and thus indistinguishable based on the existing data. It is well known that, under mild conditions [16], existence of an LTI system  $G$  with McMillan degree  $n_g$  that interpolates a given input output sequence  $(\mathbf{e}_k, \hat{\boldsymbol{\xi}}_k)_{k=1}^m$  with  $m \geq 2n_g$  is equivalent to the following rank condition:

$$\text{rank} \begin{bmatrix} \mathbf{H}(\hat{\boldsymbol{\xi}}, \mathbf{e}) \\ \mathbf{H}(\hat{\boldsymbol{\xi}}, \mathbf{e}) \end{bmatrix} = n_g + n_e, \quad \mathbf{H}(\hat{\boldsymbol{\xi}}, \mathbf{e}) \doteq \begin{bmatrix} \mathbf{H}(\hat{\boldsymbol{\xi}}) \\ \mathbf{H}(\mathbf{e}) \end{bmatrix}$$

where  $\mathbf{H}(\mathbf{e})$  and  $\mathbf{H}(\hat{\boldsymbol{\xi}})$  denote the Hankel matrices corresponding to the input and output sequences, respectively, and  $n_e \doteq \text{rank}[\mathbf{H}(\mathbf{e})]$ . As we show next, this fact allows for recasting Problem 1 into a feasibility one, albeit one where the constraints depends polynomially on the measurements.

From equations (2) and (4), it follows that the input to the nonlinearity satisfies

$$(y_k - \eta_k)Q(\xi_k, k) = P(\xi_k, k) \quad (10)$$

Thus, establishing consistency and finding an admissible sequence  $\boldsymbol{\xi}_k$  reduces to the following problem:

*Problem 2:* [Feasibility] Find admissible sequences  $\omega \in \mathcal{W}$  and  $\eta \in \mathcal{N}$  such that:

$$\begin{aligned} \text{rank} [\mathbf{H}_{\frac{m}{2}}(\xi - \omega, \mathbf{e})] &\leq n_g + n_e \\ \text{subject to (10)} \end{aligned} \quad (11)$$

A difficulty with the above formulation is that while  $\mathbf{H}$  is affine in  $\omega_k$ , it depends polynomially on  $\eta_k$ , through (10). However, as we show next, (11) can be recast into a rank constrained feasibility problem, where all the constraints are affine in the optimization variables.

*Theorem 2:* Assume that the measurement noise is characterized by an  $\ell^\infty$  bound, e.g.  $\mathcal{N} = \{\eta: \|\eta\|_{\ell^\infty} \leq \epsilon\}$ . Then, the feasibility problem (11) is equivalent to finding a point in the following set

$$\text{rank} [\mathbf{M}_t(m)] \leq n_g + n_e \quad (12)$$

$$\mathbf{N}(\mathbf{m}_i) = 0 \quad i = 1, \dots, t \quad (13)$$

$$\mathbf{L}_i(\mathbf{m}_i) \succeq 0, \quad i = 1, \dots, t \quad (14)$$

$$\mathbf{K}_i(\epsilon, \mathbf{m}_i) \succeq 0, \quad i = 1, \dots, t \quad (15)$$

$$\text{rank}[\mathbf{L}_i(\mathbf{m}_i)] = 1, \quad i = 1, \dots, t \quad (16)$$

where  $\mathbf{N} \doteq y_k \mathbf{E}(Q(\xi_k, k) - P(\xi_k, k))$ ,  $\mathbf{M} \doteq \mathbf{E}(\mathbf{H}_t)^b$ , and  $\mathbf{L}_i$  and  $\mathbf{K}_i$  are the moment matrices, defined in (9), of the joint probability distribution of  $\xi_k$  and  $\eta_k$  which is also used to compute the expected values.

*Proof.* Let  $\zeta_k = [\xi_k^T \eta_k^T]^T$ . Note that, in the feasibility problem above, the matrices do not contain any cross moments between  $\zeta_i$  and  $\zeta_j$ ,  $i \neq j$ . Hence, one can consider the distribution of  $\zeta_i$  as being independent from the one of  $\zeta_j$  for  $i \neq j$ . Furthermore, the constrains

$$\text{rank} [\mathbf{L}_i(\mathbf{m}_i)] = 1, \quad i = 1, \dots, t$$

imply that we are only considering distributions whose support contains only one element. Therefore, feasible solutions of the inequalities in the theorem above are moment sequences  $\mathbf{m}_i$  corresponding to atomic distributions whose mass is concentrated at some point  $\zeta_i^*$ ,  $i = 1, \dots, t$ . This implies that, for any feasible distribution,  $\mathbf{E}[f(\zeta_j)] = f(\zeta_j^*)$  for any function  $f(\cdot)$ . Hence,  $\zeta_i^*$  solve Problem 2. Conversely, if  $\zeta_i^*$ ,  $i = 1, \dots, t$ , is a feasible solution for Problem 2, then the moment sequence of an atomic measure concentrated at  $\zeta_i^*$ ,  $i = 1, \dots, t$ , is a feasible point for the inequalities in the theorem.  $\square$

The rank constrained LMI problem (12)–(16) can be solved directly using LMIRank [17], or, alternatively using the convex relaxation of rank proposed in [7].

<sup>b</sup>Here the expectation operator  $\mathbf{E}$  acts elementwise on  $\mathbf{H}(\boldsymbol{\xi}) = \mathbf{E}(\hat{\boldsymbol{x}} + \boldsymbol{\omega})$ . That is,  $\mathbf{M}$  is constructed by replacing all the monomials in  $\boldsymbol{\xi}$  in  $\mathbf{H}(\boldsymbol{\xi})$  with the corresponding moments.

## V. ILLUSTRATIVE EXAMPLES

Next we illustrate our approach using several examples, including the cases of missing measurements and randomly switching nonlinearities. For simplicity, in all cases the input was chosen to be an impulse.

### A. Estimation in the presence of noise

In this example we considered a Wiener system composed of a randomly generated 4th order system cascaded with the nonlinearity  $\Psi_s(\xi) = \frac{5+5\xi}{1+\xi^2}$ . The goal was to estimate  $\xi$  and  $\hat{y}$  from the measurements  $y = \hat{y} + \eta$ , with  $\|\eta\|_{\ell^\infty} \leq 0.05$ , roughly corresponding to 6% noise level. Figures 2 and 3 show the noiseless ( $\hat{y}$ ) and noisy ( $y$ ) outputs and the true (blue) and estimated (red)  $\xi$ . As illustrated there, the algorithm is able to recover  $\xi$ , with a maximum error of about 2%.

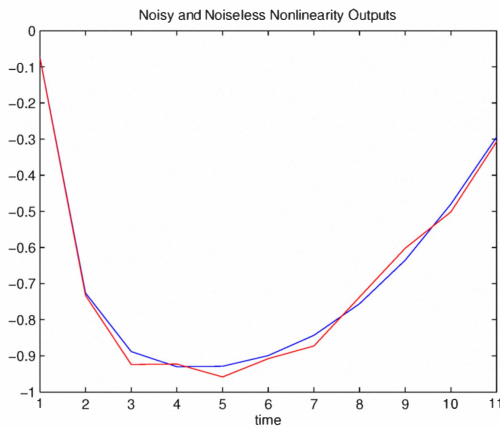


Fig. 2. Noiseless (blue) and noisy (red) measurements for Example V-A

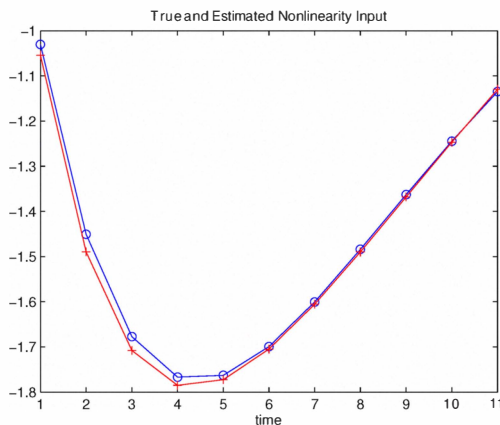


Fig. 3. Actual (blue) versus estimated (red)  $\xi$  (input to the nonlinearity) for Example V-A

### B. Interpolating Missing Data

Next we show the ability of our approach to handle missing data. To this effect we considered the same system

as before, but with 4% measurement noise, 7% process noise and 27% of the measurements missing. Figures 4 and 5 show the measured and reconstructed data, respectively. As seen there, our approach is able to successfully reconstruct  $\hat{\xi}$ , with about 5% error.

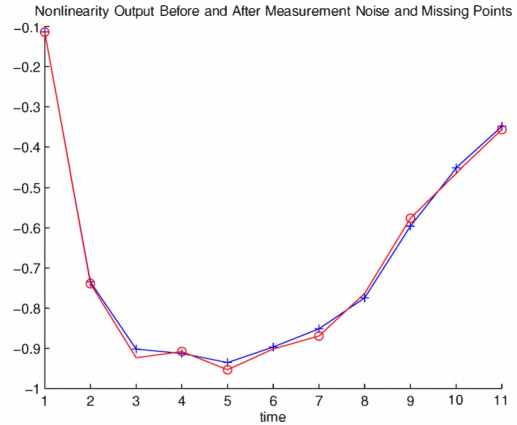


Fig. 4.  $\hat{y}$  (blue) and  $y$  (red) for Example V-B. Red circles indicate measurements used in the reconstruction

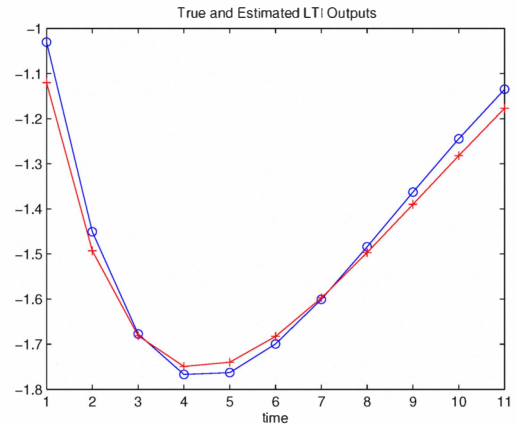


Fig. 5. True (blue) versus estimated (red) LTI System Outputs for Example V-B

### C. Randomly Switching Nonlinearities

Next, we consider the case of a switching Wiener system, where the nonlinearity switches randomly, at unknown times, amongst those on a given, known set. As we show next, this case can be also handled by our framework, with minimal modifications. For simplicity, we will assume that the system switches between two nonlinearities  $\Psi_1(\xi) = \frac{P_1(\xi)}{Q_1(\xi)}$  and  $\Psi_2(\xi) = \frac{P_2(\xi)}{Q_2(\xi)}$ , and that perfect output measurements are available. However, both assumptions can be easily removed. Since at any given time instant either  $\Psi_1$  or  $\Psi_2$  is active, then

the following constraint must hold at all times<sup>c</sup>:

$$\begin{aligned} 0 &= (y_k - \Psi_1(\xi_k))(y_k - \Psi_2(\xi_k)) \iff \\ 0 &= y_k^2 Q_1(\xi_k) Q_2(\xi_k) - y_k (P_1(\xi_k) Q_2(\xi_k) \\ &+ Q_1(\xi_k) P_2(\xi_k)) + P_1(\xi_k) P_2(\xi_k) \end{aligned} \quad (17)$$

Clearly, the constraint above can be incorporated to the optimization problem (11) as an affine constraint by considering the appropriate moments. This idea is illustrated next using a randomly generated second order system cascaded with the switching nonlinearities  $\Psi_1(\xi) = \frac{1+\xi}{\xi}$  and  $\Psi_2(\xi) = \frac{1-\xi}{\xi}$  leading to the following equivalent constraint:  $(y^2 - 1)\xi^2 - 2y\xi + 1 = 0$ . Figures 6 (a) and (b) show the output of the randomly switching nonlinearity and the switching times (here red and green indicate that  $\Psi_1(\cdot)$  or  $\Psi_2(\cdot)$  was active, respectively). Finally, Figure 6 (c) shows the result of applying our algorithm. Note that once  $\xi$  is estimated, detecting the switching times merely entails comparing  $y$  against  $\Psi_i(\xi)$ ,  $i = 1, 2$ . In this particular example, this leads to exact recovery of the switching times.

## VI. APPLICATION: 3D STRUCTURE FROM PERSPECTIVE VIDEO

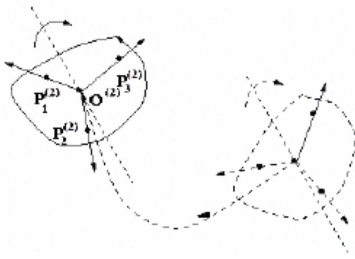


Fig. 7. 3D Motion of a rigid object is determined by the trajectories of 4 points

In this section we apply our approach to the structure from motion (SfM) problem discussed in the introduction: extracting the 3D geometry of moving objects from a sequence of 2D images taken with a perspective camera. Existing approaches can be roughly divided into geometric based and filtering based ones. Geometric methods (see for instance [18], [3], [13], [10]) use an iterative approach to determine a set of projective depths that forces a *measurement* matrix containing the coordinates of the tracked points to have at most rank 4. However, this approach yields only the affine, rather than Euclidian geometry, that is, the 3D coordinates are recovered up to an axis dependent scaling. Obtaining the 3D geometry from these coordinates entails solving a challenging non-linear, non-convex optimization problem. On the other hand, filtering-type approaches require either the availability of a motion model for the target [14], [8], [6] or the solution to a non-linear optimization (bundle adjustment) problem [11]. As we show below, these difficulties can be circumvented by reformulating the problem as the estimation

<sup>c</sup>This constraint is similar to the hybrid decoupling constraint used in GPCA [22]

of the internal signals in a Wiener system with an unknown linear component, and solved using the framework developed in this paper. Preliminary results along these lines were presented in [20] for the case of noiseless measurements. The starting point is the fact that the 3D trajectories of points  $\mathbf{P}_i$  on a rigid body satisfy a model of the form:

$$\mathbf{P}_{ti} - \mathbf{O}_t = \{\mathcal{L}[\mathbf{P}_{oi} - \mathbf{O}_o]\}_t, \quad t = 1, 2, \dots \quad (18)$$

for some point  $\mathbf{O}$  (the center of the motion), where  $\mathbf{P}_{ti}$  denote the coordinates of point  $\mathbf{P}_i$  at time  $t$ , and where  $\mathcal{L}$  is some LTI operator. For example, for a constant rotation  $R$  about a moving axis we have  $\{\mathcal{L}[\mathbf{P}_{oi} - \mathbf{O}_o]\}_k = R^k[\mathbf{P}_{oi} - \mathbf{O}_o]$ . Further, note that the 3D trajectories of all points in the rigid are fully determined once the trajectories of 4 points are given (see Figure 7). Thus, in the sequel we concentrate on recovering the trajectories of 4 points. Once this is accomplished, the trajectories of the other points can be recovered (frame by frame) from the rigidity constraints  $\|\mathbf{P}_{ti} - \mathbf{P}_{tj}\|_2 = \text{constant}$ . Next, given two points  $\mathbf{P}_i, \mathbf{P}_j$ , consider the trajectory of the difference  $\xi_{ij}(t) \doteq \mathbf{P}_{ti} - \mathbf{P}_{tj}$ <sup>d</sup>. Since  $\mathbf{P}_i, \mathbf{P}_j$  belong to the rigid, it follows from (18) that  $\xi_{ij}(t)$  can be modeled as the impulse response of a system of the form:

$$\begin{aligned} \mathbf{x}_{ij}(t) &= \mathbf{A}\mathbf{x}_{ij}(t) + \mathbf{B}_{ij}\delta(t) \\ \xi_{ij}(t) &= \mathbf{C}\mathbf{x}_{ij}(t) \end{aligned} \quad (19)$$

where the triple  $(\mathbf{A}, \mathbf{B}_{ij}, \mathbf{C})$  provide a state space realization of the operator  $\mathcal{L}$  in (18). Note that the model can be chosen so that all matrices except  $\mathbf{B}_{ij}$  are independent of which pair of points in the rigid is selected<sup>e</sup>. It follows that, given 4 points, the motion of the 3 difference vectors  $\mathbf{P}_1 - \mathbf{P}_j$ ,  $j = 2, 3, 4$  can be modeled as the impulse response of a system with the following (not minimal) state space realization:

$$\begin{aligned} \mathbf{A}_S &\doteq \text{diag}(\mathbf{A}, \mathbf{A}, \mathbf{A}),; \mathbf{B}_S \doteq [\mathbf{B}_{12}, \mathbf{B}_{13}, \mathbf{B}_{14}] \\ \mathbf{C}_S &\doteq [\mathbf{C} \quad \mathbf{C} \quad \mathbf{C}] \end{aligned} \quad (20)$$

in the sense that  $\mathbf{P}_{t1} - \mathbf{P}_{tj}$  is given precisely by its output at time  $t$  in response to an impulse applied at the  $j^{\text{th}}$  input. Further, if the pairs  $(\mathbf{A}, \mathbf{B}_{1j})$  are controllable, then the realization above is also generically controllable, but unobservable. In fact, it can be easily shown that the rank of the corresponding Hankel matrix satisfies  $\text{rank}(\mathbf{H}) = n$ , the rank of the observability matrix of the pair  $(\mathbf{A}, \mathbf{C})$ . Thus, it follows that the SfM problem can be recast into the feasibility form (11), by considering a Wiener system with 3 inputs and 2 outputs (the difference of the 2D image coordinates  $(u_1, v_1) - (u_j, v_j)$ ) and where the nonlinearity is given by:

$$\begin{aligned} u_{ti} &= f \frac{X_{ti}}{Z_{ti}} - c_u, \\ v_{ti} &= \alpha f \frac{Y_{ti}}{Z_{ti}} - c_v \end{aligned} \quad (21)$$

<sup>d</sup>It can be shown that considering the Hankel matrix of the points instead of differences leads to direction depending scaling. The proof is omitted for space reasons, but intuitively hinges on the fact that, given  $x$ ,  $\min_y \text{rank}([H_x \quad \alpha H_y])$  is the same for all  $\alpha \neq 0$ , but  $\min_y \text{rank}([H_x - \alpha H_y]) = 0$  only for  $\alpha = 1, y = x$ .

<sup>e</sup>This is accomplished by including all the information about the initial condition of  $\mathbf{P}_i - \mathbf{P}_j$  in  $\mathbf{B}_{ij}$ .

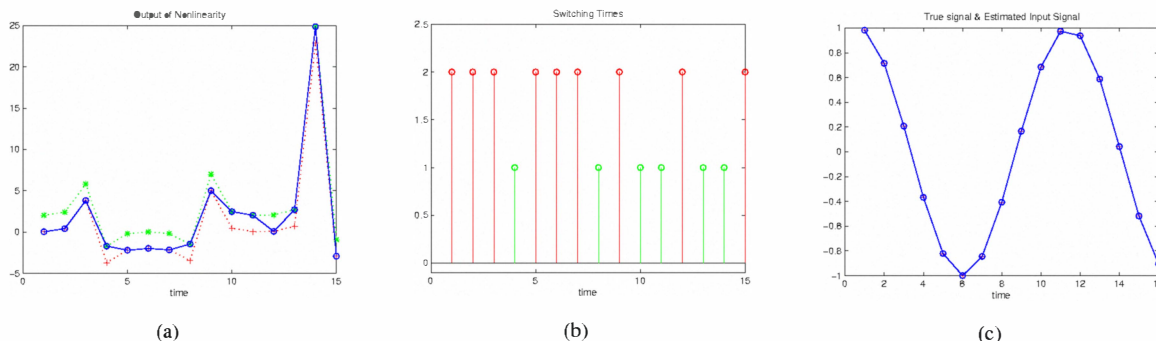


Fig. 6. Output  $y$  and switching times for Example V-C: (a) Blue: measured output  $y$ , red and green, values of  $\Psi_{r1}(\xi)$  and  $\Psi_{r2}(\xi)$ , respectively. (b) Nonlinearity switching times: a marker at  $(k, i)$  indicates that at time  $k$ ,  $\Psi_{ri}(\xi)$  is the active nonlinearity. (c) Actual and estimated values of  $\xi$ .

where  $(X_{t,i}, Y(t, i), Z(t, i))$  denote the 3D coordinates of point  $P_i$  at time  $t$ . Note that in this application in general *a priori* information about the order of the linear subsystem is not available. However, from the controllability analysis outlined above it can be shown (see [20] for details) that the correct 3D geometry corresponds to the minimum order interpolant. Thus, this geometry can be recovered from the 2D measurements by solving a slightly modified version of Problem (11) that minimizes the rank of a Hankel matrix formed by placing side by side (rather than stacking on top of each other) the Hankel matrices corresponding to each of the outputs of the linear system. Finally, note that in this particular case the nonlinearity is homogeneous of degree zero, e.g.  $(X, Y, Z)$  and  $\beta(t)(X, Y, Z)$  lead to the same pair  $(u, v)$ . Thus, this approach can recover that 3D geometry only up to a time varying scaling factor. Nevertheless, since rigid motion preserves distances, the actual 3D geometry can be recovered from this data, up to an overall, time-invariant, scaling factor, by simply scaling the solution by  $\beta(t) = \|\mathbf{P}_{t1} - \mathbf{P}_{t2}\|$ .

Next, we use real data to compare the performance of the proposed approach against two standard algorithms HTSFM [10] and MHSFM [13] for solving perspective SfM problems. The 2D data was generated by projecting the *noisy* 3D coordinates of special markers attached to an umbrella and to a human sitting on a swivel chair that were measured using a VICON motion capture system<sup>f</sup> as shown in Figure 8. Comparisons between the 3D reconstructions and ground truth are displayed in Figures 9 and 10 and quantitative results are summarized in Table I.

TABLE I  
3D MEDIAN ERROR (IN MM).

Data Set	Rank Minimization	MHSFM	HTSFM
Umbrella	3.50e+1	8.22e+1	8.32e+1
Human	4.10e+1	1.37e+2	1.51e+2

<sup>f</sup>It should be noted that the objects used in these experiments are flexible. Furthermore, the markers are about 1cm. in diameter and hence have a significant depth which affects the measurement of their location as the object moves in front of the motion capturing system.



Fig. 8. (a) Umbrella (b) Human on a chair.

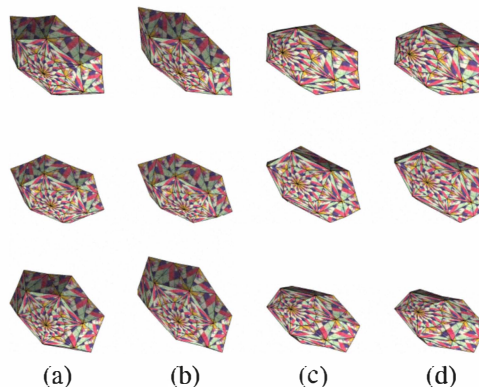


Fig. 9. Frames 1, 6 and 12 of the umbrella sequence (a) and 3D structure recovered using (b) Rank Minimization, (c) MHSFM and (d) HTSFM.

## VII. CONCLUSIONS

In this paper we considered the problem of estimating the values of both the output and the internal signals for a class of Wiener systems consisting of the cascade of an unknown linear time invariant system and a known, rational, generically non-invertible nonlinearity, based solely on past input/output data corrupted by noise. This situation arises in many scenarios of practical interest where an explicit model of the linear system is not available *a priori*. In principle this problem could be solved using a two-tiered approach, identifying first the plant and then designing a filter proceeding for instance as in [14]. However, as shown in [19], the problem of identifying the linear portion of

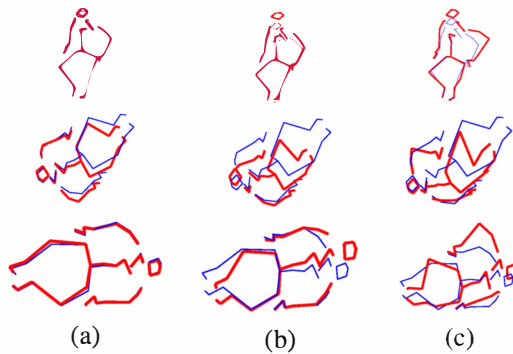


Fig. 10. Frames 1, 7, 14 frames of the human on a chair sequence with ground truth data (blue) superimposed with 3D structure (red) recovered using (a) Rank Minimization, (b) MHSFM and (c) HTSFM.

a Wiener system, even when the nonlinearity is perfectly known, is NP-hard. The present paper avoids this difficulty by exploiting results from the theory of the moments to recast the problem into a rank-constrained LMI feasibility form that can be efficiently solved using existing relaxations. As shown in the paper, the resulting algorithm can also handle the case where some of the data is missing. These results were illustrated using both academic examples and a challenging problem in computer vision: reconstructing the 3D structure of a moving object from a sequence of two dimensional images, taken with a perspective camera. As shown here, a rank-minimization approach based upon the idea of recasting the problem into a Wiener's system estimation form is able to recover this geometry, without the need to have an explicit model of the object motion, outperforming existing techniques.

## REFERENCES

- [1] Rixat Abdursul, Hiroshi Inaba, and Bijoy K. Ghosh. Nonlinear observers for perspective time-invariant linear systems. *Automatica*, 40(3):481–490, 2004.
- [2] P. Celka and P. Colditz. Nonlinear nonstationary wiener model of infant eeg seizures. *IEEE Transactions on Biomedical Engineering*, 49:556–564, 2002.
- [3] G. Chen and G. Medioni. Efficient iterative solutions to m-view projective reconstruction problem. In *IEEE Computer Vision and Pattern Recognition*, volume 2, pages 55–61, 1999.
- [4] S. C. Cripps. *RF Power Amplifiers for Wireless Communications*. Artech House, Norwood, MA, 1999.
- [5] Raul E. Curto and Lawrence A. Fialkow. Truncated k-moment problems in several variables. *Journal of Operator Theory*, 54(1):189–226, 2005.
- [6] A. J. Davison, I. D. Reid, N. D. Molton, and O. Stasse. Monoslam: Real time single camera slam. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(6):1052–1067, 2007.
- [7] M. Fazel, H. Hindi, and S. Boyd. Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices. In *American Control Conference*, June 2003.
- [8] João Pedro Hespanha. State estimation and control for systems with perspective outputs. In *Proc. of the 41st Conf. on Decision and Contr.*, Dec. 2002.
- [9] Guoqiang Hu, D. Aiken, S. Gupta, and W.E. Dixon. Lyapunov-based range identification for paracatadioptric systems. *Automatic Control, IEEE Transactions on*, 53(7):1775–1781, aug. 2008.
- [10] Y.S. Hung and W.K. Tang. Projective reconstruction from multiple views with minimization of 2d reprojection error. *International Journal of Computer Vision*, 66(3):305–317, 2006.
- [11] Georg Klein and David Murray. Parallel tracking and mapping for small AR workspaces. In *Proc. Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)*, Nara, Japan, November 2007.
- [12] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optimization*, 11:796–817, 2001.
- [13] S. Mahamud and M. Hebert. Iterative projective reconstruction from multiple views. In *IEEE Computer Vision and Pattern Recognition*, volume 2, pages 430–437, 2000.
- [14] A. Matveev, X. Hu, R. Frezza, and H. Rehbinder. Observers for systems with implicit output. *Automatic Control, IEEE Transactions on*, 45(1):168–173, jan 2000.
- [15] J. L. McGregor. Solvability criteria for certain n-dimensional moment problems. *Journal of Approximation Theory*, 30:315–333, 1980.
- [16] M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle. On- and off-line identification of linear state-space models. *International Journal of Control*, 49(1):219–232, 1989.
- [17] R. Orsi. LMIRank: software for rank constrained lmi problems. (web page and software), 2005. <http://rsise.anu.edu.au/robert/lmirank/>.
- [18] Peter Sturm and Bill Triggs. A factorization based algorithm for multi-image projective structure and motion. In B. Buxton and Roberto Cipolla, editors, *Proceedings of the 4th European Conference on Computer Vision, Cambridge, England*, volume 1065 of *Lecture Notes in Computer Science*, pages 709–720. Springer-Verlag, April 1996.
- [19] M. Sznaier. Computational complexity of set membership Hammerstein and Wiener systems identification. *Automatica*, 45(3):701–705, 2009.
- [20] M. Sznaier, M. Ayazoglu, and O. Camps. Using dynamics to recover 3dimensional euclidian structure from 2dimensional perspective projections. In *Proc 2009 IEEE Conf. Dec. Control*, pages 2414–2419, Dec. 2009.
- [21] M. Sznaier et al. The role of dynamics in extracting information sparsely encoded in high dimensional data streams. In P. Pardalos M. Hirsch, R. Murphey and D. Grundle, editors, *Dynamics of Information Systems: Theory and Applications*. Springer-Verlag, 2010. to appear.
- [22] R. Vidal, S. Soatto, Y. Ma, and S. Sastry. An algebraic geometric approach to the identification of linear hybrid systems. In *IEEE Conference on Decision and Control*, pages 167–172, December 2003.