Proceedings of the 45th IEEE Conference on Decision & Control
Manchester Grand Hyatt Hotel
San Diego, CA, USA, December 13-15, 2006

Risk Adjusted Identification of Wiener Systems

Wenjing Ma Hwasup Lim
Department of Electrical Engineering,
The Pennsylvania State University,
University Park, PA 16802.

Mario Sznaier Octavia Camps.
Electrical and Comp. Engineering Department,
Northeastern University,
Boston, MA 02115.

Abstract— **This paper addresses the problem of identification of Wiener systems from a set–membership standpoint. Its main result shows that, by pursuing a risk–adjusted approach, the problem can be reduced to a convex LMI optimization form that can be efficiently solved. These results are illustrated with a non–trivial problem arising in computer vision: tracking a human in a sequence of frames, where the challenge here arises from the changes in appearance undergone by the target and the large number of pixels to be tracked.**

I. Introduction

Wiener systems are a special type of nonlinear systems consisting of the cascade of a Linear Time Invariant (LTI) plant and a memoryless, static non–linearity. The problem of identifying such models from experimental data has received considerable attention, since these systems arise in many practical situations in a wide range of applications, including control [27], communications, [10], [7], and biology, [4], [5].

Several techniques have been proposed in the past to solve this problem. [12], [27] propose a recursive identification algorithm in the context of a stochastic approximation framework. [26], [13] use subspace methods to separately identify the linear part of the system. [11], uses a nonparametric approach where the invertible part of the nonlinearity is estimated using kernel regression. A potential difficulty with these approaches is that they are stochastic in nature, relying on the use of white Gaussian inputs. [19] addresses this limitation by explicitly modelling the inverse of the nonlinearity, coupled with the use of subspace based approaches. [2] presents an optimal two-stage identification algorithm combining recursive least squares and singular value decomposition and [1] uses a blind approach to recover all the internal variables solely based on the output measurement. All of these approaches require additional assumptions on the nonlinearity (either invertibility or a special structure), which do not hold in several problems of practical interest. Finally, [3] develops a frequency domain based algorithm by exploring the fundamental frequency and harmonics generated by the nonlinearity. However, at the present time this approach is restricted only to frequency domain data (e.g. steady state outputs generated by sinusoidal inputs).

In this paper, we propose an algorithm for time–domain based identification that avoids these difficulties by pursuing a risk-adjusted approach. Here, in return for an (arbitrarily) small risk of not being able to establish consistency of the data, the problem is reduced to a convex optimization problem. In the second part of the paper we illustrate these results with a non–trivial problem arising in computer vision: tracking a human in a sequence of frames. The challenge here arises from the changes in appearance undergone by the target and the large number of pixels to be tracked. By using the proposed identification method, we show that the problem can be solved by modelling the plant as a Wiener system. This formalizes some recent conjectures [16] where it has been argued that this motion can be explained by considering linear dynamics in a low dimensional manifold, accounting for the physics of the motion, followed by a static non-linearity that accounts for appearance changes in the target.

This work was supported by NSF grants ECS–0221562, ITR–0312558 and ECS–050166, and AFOSR grant FA9550–05–1–0437.

II. Preliminaries

For ease of reference, next we summarize the notation used in the paper.

x	column vector.
$\mathbf{A}^H$	conjugate transpose of matrix $\mathbf{A}$.
$\overline{\sigma}(\mathbf{A})$	maximum singular value of $\mathbf{A}$.
$\mathbf{A} > (\geq) 0$	$\mathbf{A} = \mathbf{A}^H$ is positive(semi) definite.
$\mathbf{I}, \mathbf{0}$	the identity and zero matrices of compatible dimensions (when omitted).
$\mathcal{BX}(\gamma)$	open γ-ball in a normed space $\mathcal{X}$: $\mathcal{BX}(\gamma) = \{x \in \mathcal{X} \colon \|x\|_{\mathcal{X}} \leq \gamma\}$.
$\overline{\mathcal{B}}\mathcal{X}(\gamma)$	closure of $\mathcal{BX}(\gamma)$.
ℓ_p	extended Banach space of vector valued real sequences equipped with the norm: $$\|x\|_p \doteq \left(\sum_{i=0}^{\infty} \|\mathbf{x}_i\|_p^p\right)^{\frac{1}{p}},$$ $p \in [1, \infty]$ and $\|x\|_\infty \doteq \sup_i \|\mathbf{x}_i\|_\infty$.

$\mathcal{H}_\infty$ Space of functions with bounded analytic continuation inside the unit disk, equipped with the norm: $\|G\|_\infty \doteq ess\sup_{|z|<1} \overline{\sigma}(G(z))$.

$\mathcal{H}_{\infty,\rho}$ space of transfer functions analytic in $|z| \leq \rho$,equipped with the norm $\|G\|_{\infty,\rho} \doteq ess\sup_{|z|<\rho} \overline{\sigma}(G(z))$.

$\mathcal{BH}_\infty$ open unit ball in the normed space: $\mathcal{BH}_\infty\colon \{h \in \mathcal{H}_\infty, \|h\|_\infty < 1\}$.

$\mathcal{BH}_\infty^N$ set of $(N-1)^{\text{th}}$ order FIR transfer matrices that can be completed to belong to $\mathcal{BH}_\infty$, i.e. $\mathcal{BH}_\infty^N \doteq \{H(z) = \mathbf{h}_0 + \mathbf{h}_1 z + \ldots + \mathbf{h}_{N-1} z^{N-1}\colon H(z) + z^N G(z) \in \mathcal{BH}_\infty$, for some $G(z) \in \mathcal{H}_\infty\}$.
By a slight abuse of notation, in cases where it is clear from the context, we will use $\mathcal{BH}_\infty^N$ also to denote the finite sequence $\{\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_{N-1}\}$.

$\mathbf{T}_x$ lower triangular block Toeplitz matrix associates with any finite sequence $\{x_k, k = 0, 1, \cdots, n-1\}$, or any column vector $\mathbf{x} = [x_0, x_1, \cdots, x_{n-1}]^T$:

$$\mathbf{T}_x = \begin{bmatrix} x_0 & 0 & \ldots & 0 \\ x_1 & x_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ x_{n-1} & x_{n-2} & \ldots & x_0 \end{bmatrix}$$

.

In the case of an LTI system G, we will denote by $\mathbf{T}_G$ the Toeplitz operator associated with an ℓ_∞ stable system. For an input sequence applied in the interval $[0, \infty]$:

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} g_0 & 0 & 0 & \cdots & \cdots \\ g_1 & g_0 & 0 & 0 & \cdots \\ g_2 & g_1 & g_2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix}}_{\mathbf{T}_G} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \end{bmatrix}$$

where u_k and y_k denote the inputs and outputs sequences to the system.

When dealing with finite sequences of length N, we will represent these operators by $\mathbf{T}_G^N$ the finite $N \times N$ upper left sub-matrices of $\mathbf{T}_G$.

III. Problem Statement

Consider the Wiener system shown in Figure 1 consisting of the interconnection of a LTI system $H(z)$ and a memoryless nonlinearity $f(.)$. The corresponding equations are given by:

$$\begin{aligned} \mathbf{y}_k &= f(\omega_k) + \eta_k \\ \omega_k &= (\mathbf{h} * \mathbf{u})_k \end{aligned} \tag{1}$$

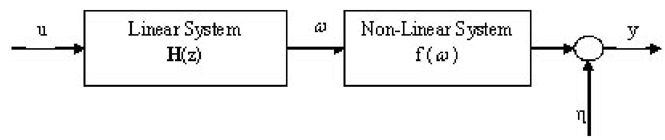

Fig. 1. Wiener System Structure

where $*$ denotes convolution and the signals $\mathbf{u} \in R^{n_u}$ and $\mathbf{y} \in R^{n_y}$ represent the experimental data: a known finite input sequence and its corresponding output sequence, corrupted by unknown but norm–bounded measurement noise η. Note that the intermediate signal $\omega \in R^{n_\omega}$ (the output of the LTI system) is not measurable. Our goal is to, given the experimental data $\{\mathbf{u}, \mathbf{y}\}$ and some *a priori* information about the plant, establish whether they are consistent, and if so, find a model that interpolates the experimental data within the measurement error level.

In the sequel, we will make the following standard assumptions about the *a priori* information:

A1.- The linear portion of the plant belongs to the set $\mathcal{S} \doteq \mathcal{BH}_{\infty,\rho}(K)$, that is, we consider exponentially stable plants with a stability margin of $(\rho - 1)$ and a peak response to complex exponential inputs bounded by some known K.

A2.- The static nonlinearity can be expanded in terms of some known basis functions:

$$f_i(.) = \sum_{j=1}^{n_f} b_{i,j} \psi_{i,j}(.), ; \psi_{i,j}(.) \text{ known}$$

A3.- The measurement noise satisfies:

$$\eta \in \mathcal{N} \doteq \{\eta\colon \|\eta_k\|_\infty \leq \epsilon\}$$

With these assumptions, the problem under consideration can be precisely stated as:

Problem 1: Given the *a priori* information $\mathcal{S}, \mathcal{N}, \psi_{i,j}(.)$ and the *a posteriori* experimental data $\{\mathbf{y}, \mathbf{u}\}$, determine:

1) if the *a priori* and *a posteriori* information are consistent, i.e., the consistency set

$$\begin{aligned} \mathcal{T}(\mathbf{y}) \doteq \{H \in \overline{\mathcal{B}}\mathcal{H}_{\infty,\rho}(K)\colon\ & y_k = \mathbf{B\Psi}\left[(h * u)_k\right] + \eta_k, \\ & \text{for some } \mathbf{B} \in R^{n_y \times n_f} \text{ and some sequence} \\ & \eta_k \in \mathcal{N}, k = 0, 1, \ldots, N-1\} \end{aligned} \tag{2}$$

is nonempty, where $\mathbf{\Psi} = (\psi_{i,j})$

2) If $\mathcal{T} \neq \emptyset$, find a nominal model $\{H, f(.)\}$ that interpolates the experimental data[1]

[1] If $\mathcal{T} = \emptyset$, then the experimental data $\{\mathbf{y}, \mathbf{u}\}$ invalidates the *a priori* assumptions about the class of models and noise, that is, the experimental data cannot be explained by models in these sets.

IV. Risk Adjusted Set–Membership Identification of Wiener Systems

In this section we present a solution to Problem 1 based on the use of risk adjusted ideas. We begin by recasting the problem into an equivalent, albeit non–convex, optimization form.

A. Establishing consistency

Lemma 1: Given $K > 0$, $\rho > 1$, a nonlinear matrix function $\mathbf{\Psi}(\cdot)$, and two vector sequences of experimental data $\mathbf{y} = [y_0, y_1, \cdots, y_{N-1}]^T$ and $\mathbf{u} = [u_0, u_1, \cdots, u_{N-1}]^T$, there exist a linear operator $\mathbf{H}(z) \in \overline{\mathcal{B}}\mathcal{H}_{\infty,\rho}(K)$ and a nonlinear mapping $f(\cdot) = \mathbf{B\Psi}(\cdot)$ such that the consistency set $\mathcal{T}(\mathbf{y})$ is nonempty, if and only if there exist a vector $\mathbf{h}$ and a matrix $\mathbf{B}$ satisfying:

$$\begin{gathered}\mathbf{M}(\mathbf{h}) \doteq \begin{bmatrix} K\mathbf{R}^{-2} & \left(\mathbf{T}_h^N\right)^T \\ \mathbf{T}_h^N & K\mathbf{R}^2 \end{bmatrix} \geq 0 \\ \mathbf{y} - \mathbf{B} \cdot \mathbf{\Psi}(\mathbf{T}_h^N \mathbf{u}) \in \mathcal{N}\end{gathered} \tag{3}$$

where $\mathbf{T}_u^N$ and $\mathbf{T}_h^N$ are the lower Toeplitz matrix associated with the sequences $\mathbf{u}$ and vector $\mathbf{h}$ respectively, and $\mathbf{R} = diag[1, \rho, \rho^2, \cdots, \rho^{N-1}]$.

Proof: From theorem 2.3.6 in [6] it follows that, given a finite sequence $\{\mathbf{h}_i\}_{i=0}^{N-1}$ there exists $H_o \in \overline{\mathcal{B}}\mathcal{H}_\infty$ such that $H_o = \mathbf{h}_o + \ldots + \mathbf{h}_{N-1}z^{N-1} + \ldots$.. if and only if $\left(\mathbf{T}_h^N\right)^T \mathbf{T}_h^N \leq \mathbf{I}$. The first condition in (3) follows now from the fact that

$$H(z) \in \overline{\mathcal{B}}\mathcal{H}_{\infty,\rho}(K) \Longleftrightarrow \frac{1}{K}H(\rho z) \in \overline{\mathcal{B}}\mathcal{H}_\infty$$

followed by a Schur complement argument. The second condition is a restatement of the conditions

$$\begin{aligned} \omega_k &= (h * u)_k, \\ y_k &= f(\omega_k) + \eta_k \text{ for some } \eta_k \in \mathcal{N} \end{aligned}$$

■

Note that the second constraint in (3) leads to a computationally hard to solve, non–convex optimization problem, even in the case where $\mathbf{\Psi}(.)$ is affine[2]. In the next section we propose to circumvent this difficulty by pursuing a risk–adjusted approach.

B. A Risk-Adjusted Relaxation

In this section we exploit some recently introduced results on sampling systems in $\overline{\mathcal{B}}\mathcal{H}_\infty$ in an attempt to find a system such that, when cascaded with the nonlinearity $\mathbf{B\Psi}$, explains the observer input/output data. This renders the second constraint in (3) convex, allowing for recasting the identification problem into a convex LMI optimization form. This idea is formalized in the following algorithm:

Algorithm 1:

1.- Generate N_s samples $\{\hat{\mathbf{h}}\}_{i=1}^{N_s}$ of the set $\overline{\mathcal{B}}\mathcal{H}_\infty^N, \rho(K)$ by using the algorithm proposed in [24] to sample $\overline{\mathcal{B}}\mathcal{H}_\infty^N$, followed by the transformation $\mathbf{h}^i = K \cdot [1, \rho^{-1}, \rho^{-2}, \cdots, \rho^{-(N-1)}] \cdot \hat{\mathbf{h}}^i$. Set $i = 1$.

2.- Solve the following optimization problem:

$$\mu(\mathbf{h}_i) = \min_{\mathbf{B}} \|\mathbf{y} - \mathbf{B\Psi}(\mathbf{T}_{h_i}^N \mathbf{u})\|_\infty \tag{4}$$

3.- If $\mu(\mathbf{h}_i) \leq \epsilon$ or $i = N_s$, stop. Otherwise, set $i = i + 1$ and go back to step 2.

The algorithm finishes either by finding one feasible pair $\{\mathbf{h}, \mathbf{B}\}$ or after N_s steps, in which case the *a posteriori* experimental data is deemed to invalidate the *a priori* information. As we show next, if the number of samples N_s is large enough, the risk of incorrectly concluding that $\mathcal{T}(\mathbf{y}) = \emptyset$ can be made arbitrarily small, possibly at the expense of increased computational time.

Lemma 2: Let (ν, δ) be two positive constants in $(0, 1)$. If N_s is chosen such that

$$N_s \geq \frac{\ln(1/\delta)}{\ln(1/(1-\nu))},$$

then, with probability greater than $1 - \delta$, the probability of not finding a feasible pair $\{\mathbf{h}, \mathbf{B}\}$ when one exists is smaller that ν.

Proof: Note that $\mathcal{T}(\mathbf{y}) \neq \emptyset$ if there exists at least one $\mathbf{h} \in \overline{\mathcal{B}}\mathcal{H}_\infty^N$ such that $\mu(\mathbf{h}) \leq \epsilon$. Direct application of the results in [25] shows that if the number of samples is at least N_s then

$$\begin{aligned}&\textbf{Prob}\big\{\textbf{Prob}\{\exists \mathbf{h} \in \overline{\mathcal{B}}\mathcal{H}_\infty^N, \rho \colon \mu(\mathbf{h}) \leq \epsilon \colon \\ &\quad \{\mu(\mathbf{h}_i)\}_{i=1}^{N_s} > \epsilon\}\big\} \leq \nu\big\} \geq (1-\delta),\end{aligned} \tag{5}$$

which yields the desired result. ■

V. Application: Human Motion Modelling and Tracking

The problems of modelling and tracking human motion using as input images from a sequence of video frames has been the subject of extensive research in the computer vision community, see for instance [18], [14], [9], [15], [20], [21], [23] and references therein. A difficulty with these approaches stems from the need to search very high dimensional spaces in order to find the correct values of the parameters of the underlying model. This is due to the fact that even using small size images and considering only human silhouettes still requires processing hundreds of pixels from each frame. In [16] it has been conjectured that the problem can be partitioned into two decoupled problems: (a) a tracking problem in a low dimensional manifold, accounting for the *dynamics* of the motion, and (b) a nonlinear, static mapping that accounts for the changes in appearance of the target. Moreover, the results there strongly suggest that the dynamics that account for the motion in the low dimensional manifold are linear, although no formal proof of the fact is given.

[2]Since in this case the resulting inequality is bilinear in the variables $\mathbf{B}, \mathbf{h}$.

Motivated by these results, in this section we apply the proposed Wiener systems identification framework to the problem of human motion modelling and tracking. The starting point is to *postulate* that such motion can be modelled as the impulse response of a Wiener system whose output is the observed images. The proposed framework can then be used to substantiate this hypothesis by establishing consistency of the *a priori* assumptions and the *a posteriori* experimental data, and to find a suitable model. In order to accomplish this, we will make the following assumptions concerning the *a priori* information:

1.- The output of the LTI part, ω, evolves in a 3-dimensional space (this hypothesis is motivated by the physics of the problem, where ω is related to the coordinates of the centroid of the target).

2.- The static nonlinearity $f(\omega)$ is given:

$$\begin{aligned} f(\omega) &= \mathbf{B}\mathbf{\Psi}(\omega) \\ &= \mathbf{B}\left[\phi(|\omega - t_1|), \phi(|\omega - t_2|), 1, \omega^T\right]^T \end{aligned}$$

with $\phi(x) = \exp(-0.8x^2)$ and

$$\begin{bmatrix} t_1 & t_2 \end{bmatrix} = \begin{bmatrix} 0.6833 & -0.7552 \\ -0.4521 & 0.4997 \\ -0.0033 & 0.0036 \end{bmatrix}$$

This hypothesis is motivated by the nonlinear dimensionality reduction method, Local Linear Embedding (LLE) proposed in [22], and the bases proposed in [8], [16] to map human silhouettes to lower dimensional spaces.

3.- Measurements of the pixels values are corrupted by measurement noise of up to 10% of their peak value. (This accounts for both actual measurement noise and errors in establishing pixel correspondences across frames).

The experimental data, shown in Figures 2 and 4, consists of the first 20 frames of a human walking on a treadmill, each having 1728 pixels[3], taken from the CMU MOBO database. Applying Algorithm 1 with $N_s = 470$ samples, which yields a probability of 99% of establishing consistency with confidence 99% led to a feasible pair $\{\mathbf{h}, \mathbf{B}\}$ and hence a model explaining the experimental data. The central interpolant[4] corresponding to the linear portion of the model is given by $H(z) = KT_1(z/\rho)T_2^{-1}(z/\rho)$, where $H(.)$ has the following state-space realization (see [17] for details):

$$\begin{aligned} \mathbf{A}_H =& \{\mathbf{A} - [\mathbf{C}_-^T\mathbf{C}_- + (\mathbf{A}^T - \mathbf{I})]^{-1}\mathbf{C}_-^T\mathbf{C}_-(\mathbf{A} - \mathbf{I})\}^{-1} \\ \mathbf{B}_H =& [\mathbf{C}_-^T\mathbf{C}_-(\mathbf{A}^T - \mathbf{A} - \mathbf{I}) - (\mathbf{A}^T - \mathbf{I})\mathbf{A}]^{-1}\mathbf{C}_-^T \\ \mathbf{C}_H =& K\mathbf{C}_+ - K\mathbf{C}_+ \\ & \{\mathbf{A} - [\mathbf{C}_-^T\mathbf{C}_- + (\mathbf{A}^T - \mathbf{I})]^{-1}\mathbf{C}_-^T\mathbf{C}_-(\mathbf{A} - \mathbf{I})\}^{-1} \\ \mathbf{D}_H =& K\mathbf{C}_+\{[\mathbf{C}_-^T\mathbf{C}_- + (\mathbf{A}^T - \mathbf{I})]\mathbf{A} \\ & - \mathbf{C}_-^T\mathbf{C}_-(\mathbf{A} - \mathbf{I})\}^{-1}\mathbf{C}_-^T \end{aligned} \tag{6}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & \mathbf{I}_{(N-1)\times(N-1)} \\ 0 & 0 \end{bmatrix}, \mathbf{C}_- = [\overbrace{1, 0, \cdots, 0}^{N}], \mathbf{C}_+ = \frac{\mathbf{h}^T}{K}.$$

Replacing the value of $\mathbf{h}$ in the formulas above leads (after a model reduction step) to the following 5^{th} order system for the linear portion of the model:

$$H(z) = \begin{bmatrix} \frac{-0.03z^5+0.19z^4-0.30z^3+0.22z^2-0.08z+0.02}{z^5-1.84z^4+1.13z^3+0.92z^2-1.57z+0.83} \\ \\ \frac{-0.08z^5+0.09z^4-0.02z^3-0.08z^2+0.07z-0.03}{z^5-1.84z^4+1.13z^3+0.92z^2-1.57z+0.83} \\ \\ \frac{0.14z^5-0.35z^4+0.27z^3+0.12z^2-0.30z+0.16}{z^5-1.84z^4+1.13z^3+0.92z^2-1.57z+0.83} \end{bmatrix} \tag{7}$$

The corresponding static output nonlinearity is given by $\mathbf{B}\mathbf{\Psi}$, where a surface plot of the matrix $\mathbf{B}$ is shown in Figure 6. Here the x, y axes correspond to the index of the matrix and the z axis to the matrix value. Note that in most cases, the rows of $\mathbf{B}$ are sharply peaked around one or two values, indicating that these pixels can be explained using fewer elements of the bases $\Psi_{i,j}$.

The impulse response of the identified system is shown in Figures 3 and 5. As illustrated there, the system is able to correctly predict the appearance of the target in the next two frames (21 and 22). For comparison purposes, the actual output, not used for training, is shown in the last 2 frames of Figure 4[5].

VI. CONCLUSION

In this paper we propose an algorithm for deterministic set membership identification of Wiener systems using time–domain data. As shown in the paper, in principle this formulation leads to a non-convex, computationally hard to solve optimization problem. However, by exploiting recently introduced results on sampling of transfer functions, the problem can be relaxed to a convex optimization, at the price of an arbitrarily small probability of mis-identifying the plant. These results were illustrated with a problem that has been the object of considerably attention in the computer vision community: modelling the evolution of human motion in a sequence of two-dimensional images. By modelling

[3]The measurements vector $\mathbf{y}$, of dimension 1728, was constructed by row-wise stacking the grey-scale values of the pixels in each frame.

[4]All solutions in the consistency set can be parameterized in terms of a free parameter $Q(z) \in \overline{\mathcal{B}}\mathcal{H}_{\infty,\rho}$. The central solution corresponds to $Q(z) = 0$.

[5]The worst case prediction error in this interval, found by comparing the actual and predicted sequences for the 1728 pixels, is 21.73% of the ℓ_∞ norm of the actual output. The mean square error is 0.1393 in terms of the ℓ_∞ norm of the error signal.

this evolution as the impulse response of a Wiener system, we were able to establish that the problem can be indeed decoupled into two simpler subproblems: (i) tracking the trajectory of an LTI system in a low dimensional subspace and (ii) finding a nonlinear static mapping that accounts for appearance changes. By decoupling the intrinsic dynamics of the target from changes in its appearance, this decomposition is the first step towards designing faster, more robust trackers.

Efforts are currently underway to generalize the techniques proposed here to cases where both the linear dynamics and the nonlinearity are slowly time varying.

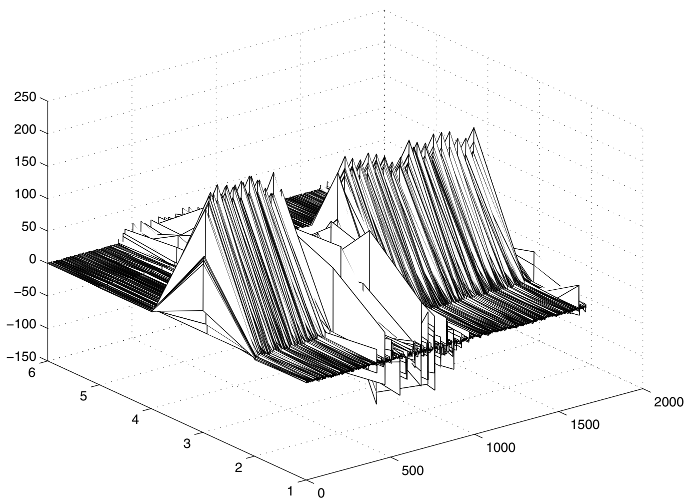

Fig. 6. Surface of $\mathbf{B}$ for the nonlinear part of the model

VII. REFERENCES

[1] E.W. Bai, ”A blind approach to the Hammerstein-Wiener model identification”, *Automatica*, vol. 38, 2002, pp. 967-979.

[2] E.W. Bai ”An Optimal Two-Stage Identification Algorithm for Hammerstein-Weiner Nonlinear Systems”, *Automatica*, vol. 34, 1998, pp. 333-338.

[3] E.W. Bai, ”Frequency Domain Identification of Wiener Models”, *Automatica*, vol. 39, 2003, pp 1521-1530.

[4] A.C. Brinker, ”A Comparison of Results from Parameter Estimation of Impulse Responses of the Transient Visual Systems”, *Biological Cybern*, vol. 61, 1989, pp 139-151.

[5] P. Celka and P. Colditz, ”Nonlinear Nonstationary Wiener model of infant EEG seizures”, IEEE *Transactions on Biomedical Engineering*, vol. 49, 2002, pp 556-564.

[6] J. Chen and G. Gu. *Control–oriented System Identification : an $\mathcal{H}_\infty$ Approach*, John Wiley & Sons, New York, 2000.

[7] S.C. Cripps, *RF Power Amplifiers for Wireless Communications*, Artech House, Norwood, MA, 1999.

[8] A. Elgammal and C.S. Lee, ”Inferring 3D body pose from silhouettes using activity manifold learning”, IEEE *Conference on Computer Vision and Pattern Recognition*, Washington, DC, 2004, pp 681-688.

[9] D. Gavrila and L. Davis, ”3-d model-based tracking of humans in action: a multi-view approach”, IEEE *Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 1996, pp 73-80.

[10] G. Giunta, G. Jacoviti and A. Neri, ”Bandpass Nonlinear System Identification by Higher Cross Correlation”, IEEE *Transactions on Signal Processing*, vol. 39, 1991, 2092-2095.

[11] W. Greblicki, ”Nonparametric Approach to Wiener System Identification”, IEEE *Transactions on Circuits and Systems*, vol. 44, 1997, pp 538-545.

[12] W. Greblicki, ”Nonparametric Identification of Wiener system”, IEEE *Transactions on Information Theory*, vol. 38, 1992, pp 1487-1493.

[13] B.R.J. Haverkamp, C.T. Chou, and U. Verhaegen, ”Subspace Identification of Continuous-time Wiener Models”, IEEE *Conf. of Decision and Control*, vol. 2, 1998, pp 1846-1847.

[14] D. Hogg, ”Model-based vision: a program to see a walking person”, *Image and Vision Computing*, vol. 1, 1983, pp 5-20.

[15] I.A. Kakadiaris and D. Metaxas, ”Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection”, IEEE *Conference of Computer Vision and Pattern Recognition*, Los Alamitos, CA, 1996, pp 18-20.

[16] V. Morariu and O. Camps, “Modelling Correspondences for Multi Camera Tracking using NonLinear Manifold Learning and Target Dynamics,” Proc. 2006 IEEE Comp. Vision and Pattern Recogn. Conf. (CVPR), N. Y., June, pp. 545–552.

[17] P. A. Parrilo, M. Sznaier, R. S. Sánchez Peña and T. Inanc, “Mixed time/frequency domain based robust identification,” *Automatica,* Vol. 34, No. 11, pp. 1375–1389, 1998.

[18] J. O’Rourke and N.I. Badlet, ”Model-based image analysis of human motion using constraint propagation”, IEEE *Transactions on Pattern Analysis and Machine Intelligence*, vol. 6 , 1980, pp 522-536.

[19] R. Raich, G.T. Zhou and M Viberg, ”Subspace Based Approaches for Wiener System Identification”, IEEE *Transactions on Automatic Control*, vol. 50, 2005, pp 1629-1634.

[20] J.M. Rehg and T. Kanade, ”Model-based tracking of self-occluding articulated objects”, *International Conference on Computer Vision*, Boston, MA, 1995, pp 612-617.

[21] K. Rohr, ”Towards model-based recognition of human movements in image sequence”, *Computer Vision and*

Fig. 2. First 11 frames of a walking person video sequence

Fig. 3. First 11 frames of the output of the identified Wiener System

Fig. 4. Frames 12 to 22 of the actual walking sequence (Frames 21 and 22 were not used in the identification)

Fig. 5. Frames 12 to 22 predicted by the identified system

Graphic Image Processing Conference, vol. 59, 1994, pp 94-115.

[22] S. Roweis and L. Saul, ”Nonlinear dimensionality reduction by locally linear embedding”, *Science*, v. 290, 2000, pp. 2323-2326.

[23] H. Sidenbladh, M.J. Black and D.J. Fleet, ”Stochastic tracking of 3d human figures using 2d image motion”, *European Conference on Computer Vision*, Dublin, Ireland, 2000, pp 702-718.

[24] M. Sznaier, C. M. Lagoa and M.C. Mazzaro, “ An Algorithm for Sampling Balls in $\mathcal{H}_\infty$ with Applications to Risk–Adjusted Performance Analysis and Model (In)Validation,” *IEEE Trans. Autom. Contr.,* March 2005, pp. 410–416.

[25] R. Tempo, E. W. Bai and F. Dabbene (1996). “Probabilistic Robustness Analysis: Explicit Bounds for the Minimum Number of Sampling Points”, *Proceedings of the IEEE Conference on Decision and Control*, pp. 3418–3423, Kobe, Japan.

[26] D. Westwick and M. Verhaegen, ”Identifying MIMO Wiener Systems using Subspace Model Identification Methods”, *Signal Proces*, vol. 52, 1996, pp 235-258.

[27] T. Wigren, ”Convergence Analysis of Recursive Identification Algorithms based on the Nonlinear Wiener Model”, IEEE *Transactions on Automatic Control*, vol. 39, 1994, pp 2191-2206.