becomes exact and being made exclusively by the RED ($\alpha(\tilde{\nu}_{rl}) = 0$), which implies that $\varepsilon_g(t) = 0, \forall t \geq \overline{T}$. In this case, an ISL is formed and applying [8, Lemma 1] to system (38)–(40), with $f(t)$ satisfying (8), one can conclude that the error state z will converge exponentially to zero and σ becomes identically zero after some finite time. ∎

REFERENCES

[1] S. Baev, Y. Shtessel, and I. Shkolnikov, "Nonminimum-phase output tracking in causal systems using higher-order sliding modes," *Int. J. Robust Nonlin. Control*, vol. 18, pp. 454–467, 2008.

[2] G. Bartolini, A. Levant, A. Pisano, and E. Usai, "Higher-order sliding modes for output-feedback control of nonlinear uncertain systems," in *Variable Structure Systems: Towards the 21st Century*, X. Yu and J.-X. Xu, Eds. New York: Springer-Verlag, 2002, ch. 4.

[3] S. Emelyanov and S. Korovin, *Control of Complex and Uncertain Systems*. New York: Springer Verlag, 2000.

[4] A. F. Filippov, "Differential equations with discontinuous right-hand side," *Amer. Math. Soc. Transl.*, vol. 42, no. 2, pp. 199–231, 1964.

[5] L. Fridman and A. Levant, "Higher order sliding modes," in *Sliding Mode Control in Engeneering*, W. Perruquetti and J. Barbot, Eds. New York: Marcel Dekker, Inc., 2002, pp. 53–101.

[6] L. Fridman, Y. Shtessel, C. Edwards, and X. G. Yan, "Higher-order sliding-mode observer for state estimation and input reconstruction in nonlinear systems," *Int. J. Robust Nonlin. Control*, vol. 18, pp. 399–412, 2008.

[7] L. Hsu, "Smooth sliding control of uncertain systems based on a prediction error," *Int. J. Robust Nonlin. Control*, vol. 7, pp. 353–372, 1997.

[8] L. Hsu, F. Lizarralde, and A. Araújo, "New results on output feedbak VS-MRAC: Design and stability analysis," *IEEE Trans. Auomat. Control*, vol. 42, no. 3, pp. 386–393, Mar. 1997.

[9] P. Ioannou and K. Sun, *Robust Adaptive Control*. Englewood Cliffs, NJ: Prentice Hall, 1996.

[10] Z. P. Jiang and M. Y. Mareels, "A small-gain control method for nonlinear cascade systems with dynamic uncertainties," *IEEE Trans. Automat. Control*, vol. 42, no. 3, pp. 292–308, Mar. 1997.

[11] Z. P. Jiang, A. R. Teel, and L. Praly, "Small-gain theorem for ISS systems and applications," *Math. Control Signals Syst.*, vol. 7, pp. 95–120, 1994.

[12] A. Levant, "Higher-order sliding modes, differentiation and output-feedback control," *Int. J. Control*, vol. 76, no. 9, pp. 924–941, 2003.

[13] A. Levant, "Exact differentiation of signals with unbounded higher derivatives," in *Proc. IEEE Conf. Dec. Control*, San Diego, CA, 2006, pp. 5585–5590.

[14] E. V. L. Nunes, L. Hsu, and F. Lizarralde, "Globally stable output-feedback sliding mode control with asymptotic exact tracking," in *Proc. Amer. Control Conf.*, Boston, MA, 2004, pp. 638–643.

[15] E. V. L. Nunes, L. Hsu, and F. Lizarralde, "Output-feedback sliding mode control for global asymptotic tracking of uncertain systems using locally exact differentiators," in *Proc. Amer. Control Conf.*, Minneapolis, MN, 2006, pp. 5407–5412.

[16] E. V. L. Nunes, L. Hsu, and F. Lizarralde, "Global output feedback tracking controller based on hybrid estimation for a class of uncertain nonlinear systems," in *Proc. 10th Int. Workshop Variable Structure Syst.*, Antalya, Turkey, 2008, pp. 141–146.

[17] T. R. Oliveira, A. J. Peixoto, E. V. L. Nunes, and L. Hsu, "Control of uncertain nonlinear systems with arbitrary relative degree and unknown control direction using sliding modes," *Int. J. Adaptive Control Signal Process.*, vol. 21, no. 8/9, pp. 692–707, 2007.

[18] S. S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence and Robustness*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[19] V. Utkin, J. Guldner, and J. Shi, *Sliding Mode Control in Electromechanical Systems*. London, U.K.: Taylor & Francis, 1999.

Risk Adjusted Set Membership Identification of Wiener Systems

Mario Sznaier, *Member, IEEE*, Wenjing Ma, Octavia I. Camps, *Member, IEEE*, and Hwasup Lim

Abstract—This technical note addresses the problem of set membership identification of Wiener systems. Its main result shows that even though the problem is generically NP-hard, it can be reduced to a tractable convex optimization through the use of risk-adjusted methods. In addition, this approach allows for efficiently computing worst-case bounds on the identification error. Finally, we provide an analysis of the intrinsic limitations of interpolatory algorithms. 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.

Index Terms—Risk-adjusted relaxations, Wiener systems identification, worst-case nonlinear identification.

I. INTRODUCTION

Identification of Wiener systems from time domain data has received considerable attention in the past decade. Stochastic approaches based on the use of white Gaussian inputs include [1]–[4]. More general inputs have been considered in [5]–[7], at the price of additional assumptions on the nonlinearity (either invertibility or a special structure). Set membership techniques [8], [9] provide an attractive alternative, since they require few additional assumptions about the nonlinearity or noise sets. Moreover, these approaches furnish hard bounds on the values of the unknown parameters of the plant in a form that can be directly used by robust control synthesis techniques. However, as recently shown in [10], set membership identification of Wiener systems is generically NP hard.

To avoid this difficulty, in this technical note we propose a risk-adjusted convex relaxation, where, 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 whose complexity scales linearly with the data. In addition, this approach allows for efficiently computing worst-case bounds on the identification error. Finally, we also provide an analysis of the convergence properties and intrinsic limitations of interpolatory algorithms. These results are illustrated 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 [11] where it has been argued that this motion can be explained by considering linear dynamics in a low

M. Sznaier and O. I. Camps are with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA (e-mail: msznaier@ece.neu.edu; camps@ece.neu.edu).

W. Ma and H. Lim are with the Department of Electrical Engineering, The Pennsylvania State University, University Park, PA 16802 USA (e-mail: wxm163@psu.edu; hxl211@psu.edu).

dimensional manifold, accounting for the physics of the motion, followed by a static non-linearity that accounts for appearance changes in the target.

II. PRELIMINARIES

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

$\overline{\sigma}(\mathbf{A})$ — Maximum singular value of matrix $\mathbf{A}$.

$\mathbf{A} \geq 0$ — $\mathbf{A}$ is positive semidefinite.

$(\mathcal{X}, m)$ — Metric space of elements in $\mathcal{X}$ equipped with the metric $m(x_1, x_2)$.

$d(\mathcal{A})$ — Diameter of $\mathcal{A} \subseteq \mathcal{X} : d(\mathcal{A}) \doteq \sup\limits_{x, a \in \mathcal{A}} m(x, a)$.

$\overline{\mathcal{BX}}(\gamma)$ — Closed γ-ball in a normed space $\{\mathcal{X}, \|.\|\} : \overline{\mathcal{BX}}(\gamma) = \{x \in \mathcal{X} : \|x\|_{\mathcal{X}} \leq \gamma\}$.

ℓ^{∞} — Extended Banach space of vector valued real sequences equipped with the norm: $\|x\|_{\infty} \doteq \sup_i \|x_i\|_{\infty}$.

$\mathcal{H}_{\infty, \rho}$ — Space of transfer functions analytic in $|z| \leq \rho$, equipped with the norm $\|G\|_{\infty, \rho} \doteq ess \sup_{|z| < \rho} \overline{\sigma}(G(z))$. The case $\rho = 1$ will be simply denoted $\mathcal{H}_{\infty}$.

$\mathcal{BH}_{\infty}^{N}$ — Set of $(N-1)^{\text{th}}$ order FIR transfer matrices that can be completed to belong to $\mathcal{BH}_{\infty}$, i.e. $\mathcal{BH}_{\infty}^{N} \doteq \{H(z) = \mathbf{h}_0 + \mathbf{h}_1 z + \ldots + \mathbf{h}_{N-1} z^{N-1} : H(z) + z^N G(z) \in \mathcal{BH}_{\infty}, \text{ for some } G(z) \in \mathcal{H}_{\infty}\}$. 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 associated with any finite sequence $\{x_k, k = 0, 1, \cdots, n-1\}$, or any column vector $x = [x_0, x_1, \cdots, x_{n-1}]^T$:

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

Given a Linear Time Invariant (LTI) system H, we will denote by h its impulse response sequence (Markov parameters) and by $\mathbf{T}_h$ the corresponding Toeplitz matrix. Finally, when dealing with finite sequences of length N, we will use h_N and $\mathbf{T}_h^N$ to denote the truncated sequence and the corresponding $N \times N$ upper left sub-matrix of $\mathbf{T}_h$.

III. PROBLEM STATEMENT

Consider the Wiener system shown in Fig. 1 consisting of the interconnection of a LTI system $H(z)$ and a memoryless nonlinearity $\mathbf{\Psi}(.)$. The corresponding equations are given by

$$\mathbf{y}_k = \mathbf{\Psi}(\boldsymbol{\omega}_k) + \boldsymbol{\eta}_k$$
$$\boldsymbol{\omega}_k = (h * u)_k \qquad (1)$$

where $*$ denotes convolution and the signals $u \in R^{n_u}$ and $y \in R^{n_y}$ represent the experimental data: a known input and its corresponding output, corrupted by unknown but norm-bounded measurement noise η. Our goal is to, given experimental data consisting of N_m measurements of the input/output sequences $\{u_k, y_k\}_{k=0}^{N_m-1}$ and some *a priori* information about the plant, establish whether they are consistent, and
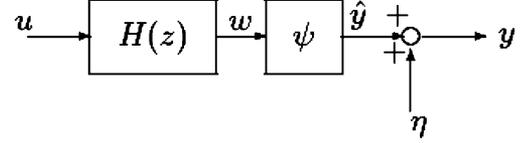
Fig. 1. Wiener system structure.

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 $\mathcal{S} \doteq \{S(z) : S(z) = H(z) + P(z)\}$, where $H(z) \in \mathcal{S}_{n_p} \doteq \mathcal{BH}_{\infty, \rho}(K)$, and $P(z) \in \mathcal{S}_p \doteq \{P(z) = p^T G_p(z), p \in \Pi \subset \mathcal{R}^{N_p}\}$, where the N_p components $G_{p_i}(z)$ of vector $G_p(z)$ are known, linearly independent functions.

A2.- The measurement noise satisfies: $\eta \in \mathcal{N} \doteq \{\eta : \|\eta_k\|_{\infty} \leq \epsilon\}$.

A3.- $\Psi \in \mathcal{F}$, a family of equi-bounded, uniformly equi-continuous functions.

Remark 1: Assumptions [A1.-] and [A2.-] are standard [12]. Assumption [A3.-], required in order to guarantee convergence of *any* interpolatory algorithm as the information is completed, is automatically satisfied in cases where the nonlinearity can be expressed as a bounded combination of a finite set of known, continuous basis functions:[1]

$$\mathcal{F} \doteq \left\{ \mathbf{\Psi}(.) : \mathbf{\Psi}(.) = \mathbf{B}\Phi, \Phi = \begin{bmatrix} \phi_1(.), \ldots, \phi_{n_f}(.) \end{bmatrix}^T, \right.$$
$$\left. \phi_i(.)\text{known}, \mathbf{B} \in \mathcal{B} \subset R^{n_y \times n_f} \right\}.$$

In principle the continuity assumption seems too strong, ruling out, among others, relay nonlinearities. However, note that these nonlinearities render all plants whose response to the given input has the same zero-crossings indistinguishable. Further, since the domain of the nonlinearity is restricted to a compact subset of R^{n_w}, it can be approximated arbitrarily close by a continuously differentiable function. Indeed, it can be argued that such smooth models provide a better representation of physically realizable nonlinearities.

Under assumptions [A1–A3], the problem under consideration can be precisely stated as:

Problem 1: Given the *a priori* information $\mathcal{S}, \mathcal{N}, \mathcal{F}$ and the experimental data $\{y_k, u_k\}_{k=0}^{N_m-1}$:

1) determine whether the information is consistent, i.e., the consistency set $\mathcal{T}(y, N_m, \mathcal{N}) \doteq \{H \in \mathcal{S} : y_k = \mathbf{\Psi}[(h*u)_k] + \eta_k, k = 0, \ldots, N_m - 1 \text{ for some } \mathbf{\Psi} \in \mathcal{F} \text{ and some sequence } \eta_k \in \mathcal{N}\}$ is nonempty.

2) If $\mathcal{T}(y, N_m, \mathcal{N}) \neq \emptyset$, find a nominal model $\{H, \mathbf{\Psi}(.)\}$ that interpolates the data.

IV. A POLYNOMIAL TIME RELAXATION

Unfortunately, as recently shown in [10], Problem 1 above is NP hard both in the number of inputs to the nonlinearity Ψ and the number of experimental data pairs (u, y), even in cases where the nonlinearity $\Psi(.)$ is known. To circumvent this difficulty, in this section we propose a polynomial time relaxation. The main idea is to exploit recently introduced results on sampling systems in $\overline{\mathcal{BH}}_{\infty}$ to sample the set of candidate linear plants $\mathcal{S}$, thus reducing the problem to a (noisy) interpolation problem in $\mathcal{F}$. To solve the latter problem, we will assume that

---

[1]Since H has a finite ℓ^2 induced norm, for a given input, the signal w, and hence the domain of ϕ_i belong to a compact subset of R^{n_w}. Uniform equicontinuity and equiboundedness follow from continuity and finiteness of the family $\{\phi_i\}$.

$\mathcal{F}$ is spanned by a known basis $\{\Phi_i(.)\}$. This assumption, standard in the field (see for instance [3], [5], [7]), is justified by the fact that, since $H \in \overline{\mathcal{BH}}_{\infty}^{N}(K)$ and u is bounded, the range and domain of Ψ are compact, and thus each of its components can be approximated there arbitrarily close for instance by a polynomial. These ideas are formalized in the following algorithm:

Algorithm 1:

1.- Generate N_s samples $\{\hat{h}_{np}^{i}\}_{i=1}^{N_s}$ of the set $\overline{\mathcal{BH}}_{\infty,\rho}^{N}(K)$ by using the algorithm proposed in [13] to sample $\overline{\mathcal{BH}}_{\infty}^{N}$, followed by the transformation $h_{np}^{i} = K\,\mathrm{diag}[1, \rho^{-1}, \rho^{-2}, \cdots, \rho^{-(N_m-1)}] \cdot \hat{h}_{np}^{i}$. Generate N_s uniformly distributed samples $\{p^{i}\}_{i=1}^{N_s}$ from the set Π. Set $i = 1$. Let $h^{i} = h_{np}^{i} + p^{i^{T}} \mathbf{G_p}$. Solve the following optimization problem:

$$\mu(h^{i}) = \min_{\mathbf{B}} \left\| y - \mathbf{B}\Psi\left(\mathbf{T}_{h^{i}}^{N_m} u\right) \right\|_{\infty} \qquad (2)$$

where $\Psi = [\Phi_1(.)\ \Phi_2(.)\ \ldots\ \Phi_m(.)]^{T}$, $\Phi_i(.)$ known. If $\mu(h^{i}) \le \epsilon$ 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 $\{H, \mathbf{B}\}$ 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 $\mathcal{T}(y, N_m, \mathcal{N}) = \emptyset$ can be made arbitrarily small.

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

$$N_s \ge \frac{\ln(1/\delta)}{\ln(1/(1-\nu))} \qquad (3)$$

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

Proof: Note that $\mathcal{T}(y, N_m, \mathcal{N}) \ne \emptyset$ if there exists at least one $H \in \mathcal{S}$ such that $\mu(h) \le \epsilon$. Direct application of the results in [14] shows that if the number of samples is at least N_s then

$$\mathbf{Prob}\left\{\mathbf{Prob}\left\{\exists H \in \overline{\mathcal{BH}}_{\infty,\rho}^{N} : \mu(H) \le \epsilon \text{ and}\right.\right.$$
$$\left.\left.\left\{\mu(H^{i})\right\}_{i=1}^{N_s} > \epsilon\right\} \le \nu\right\} \ge (1-\delta) \quad (4)$$

which yields the desired result. ∎

Remark 2: Note that as $\nu \to 0$, the number of samples (and hence the computational time) $N_s \to \infty$. This is consistent with the fact that the problem is NP hard and therefore no polynomial time exact solutions should be expected.

V. Convergence Analysis

In this section we briefly analyze the convergence properties of the proposed algorithm as the information is completed, that is $\epsilon \to 0$ and $N_m \to \infty$. To this effect, we will prove a more general result concerning the intrinsic limitations of *any* interpolatory algorithm.[2] Contrary to the case of linear plants, in the case of Wiener systems convergence to the true plant (in the sense that $\mathcal{T}(y, N_m, \mathcal{N}) \to \{h^o\}$) is no

[2]In the context of this technical note, interpolatory algorithms are those such that the linear portion of the *true* plant $h^o \in \mathcal{T}(y, N_m, \mathcal{N})$.

longer guaranteed (except in special cases, e.g. invertible nonlinearities). Rather, any interpolatory algorithm $\mathcal{A}^{I}$ will converge to the set

$$\mathcal{T}^{*}(y) \triangleq \{H \in \mathcal{S} : y_k = \Psi(h * u)_k, k = 0, 1, \ldots,$$
$$\text{for some } \Psi \in \mathcal{F}\}$$

e.g., the consistency set in case of complete and uncorrupted experimental information. The diameter of this set $e^{*}(y) \doteq d\{\mathcal{T}(y)\}$ defines an *intrinsic local* worst-case error for Wiener systems, in the sense that this is the best that can be achieved by any interpolatory algorithm.

Theorem 1: Assume that $\mathcal{N} \doteq \mathcal{B}\ell^{\infty}(\epsilon)$. Then as the information is completed

$$\lim_{N_m \to \infty, \epsilon \to 0} \mathcal{T}(y, N_m, \epsilon) = \mathcal{T}^{*}(y).$$

Proof: The proof proceeds by showing that $\{\mathcal{T}(y, N_m, \epsilon_m)\}$, the sequence of consistency sets indexed by (N_m, ϵ_m), converges to $\mathcal{T}^{*}(y)$. From the continuity assumption on $\Psi(.)$ and the fact that $\mathcal{S}$ is closed, it follows that the sets $\mathcal{T}(y, N_m, \epsilon_m)$ and $\mathcal{T}^{*}(y)$ are closed. Consider now sequences $\{N_m \uparrow \infty, \epsilon_m \downarrow 0\}$ as $m \to \infty$. The corresponding sequence $\mathcal{T}(y, N_m, \epsilon_m)$ satisfies

$$\mathcal{T}(y, N_{m+1}, \epsilon_{m+1}) \subseteq \mathcal{T}(y, N_m, \epsilon_{m+1}) \subseteq \mathcal{T}(y, N_m, \epsilon_m). \quad (5)$$

Therefore its limit $\lim_{m \to \infty} \mathcal{T}(y, N_m, \epsilon_m)$ exists and equals $\cap_{m>0} \mathcal{T}(y, N_m, \epsilon_m)$ ([15], page 19).

If the identification problem is well posed, i.e. $\mathcal{T}^{*}(y) \ne \emptyset$, then

$$\cap_{m>0} \mathcal{T}(y, N_m, \epsilon_m) \supseteq \cap_{m>0} \mathcal{T}(y, N_m, 0) = \mathcal{T}^{*}(y)$$

where we have used the fact that $0 \in \mathcal{N}$. Consider now an arbitrary plant $\tilde{H} \in \cap_{m>0} \mathcal{T}(y, N_m, \epsilon_m)$. Since $\tilde{H} \in \mathcal{T}(y, N_m, \epsilon_m) \forall m$, it follows that for each m there exists at least one nonlinearity $\Psi_m \in \mathcal{F}$ and an admissible noise sequence $\eta_m \in \mathcal{N}$ such that $y_k = \Psi[(h * u)_k] + \eta_k, k = 0, 1, \ldots N_m - 1$. Moreover, given any $\epsilon > 0$, there exists some $k_1(\epsilon)$ such that for all $m > k_1(\epsilon)$

$$\left\| y_j - \Psi_m\left[(\tilde{h} * u)_j\right] \right\| \le \frac{\epsilon}{2}, \quad j = 0, 1, \ldots, N_m. \quad (6)$$

By assumption [A.3-] the family of nonlinearities $\mathcal{F}$ is equibounded and equicontinuous, hence it follows from Arzela Ascoli's Theorem that it containts a convergent subsequence $\{\Psi_k\} \to \tilde{\Psi}$. Thus, there exists some $k_2(\epsilon)$ such that for all $m > k_2(\epsilon)$, $\|\Psi_m(.) - \tilde{\Psi}(.)\| \le (\epsilon/2)$. Thus, for all $m > max\{k_1, k_2\}$ we have that

$$\left\| y_j - \tilde{\Psi}\left[(\tilde{h} * u)_j\right] \right\| \le \epsilon, \quad j = 0, 1, \ldots, N_m - 1. \quad (7)$$

Since $N_m \uparrow \infty$ and ϵ is arbitrary, this implies that $y_j = \tilde{\Psi}[(\tilde{h}*u)_j], \forall j$ and hence $\tilde{H} \in \mathcal{T}^{*}(y)$. It follows that $\cap_{m>0} \mathcal{T}(y, N_m, \epsilon_m) \subseteq \mathcal{T}^{*}(y)$, which together with (6) establishes the equality. ∎

Remark 3: Note that assumption [A.3] plays a key role in the proof above. Without the sequential compactness on $\mathcal{F}$ induced by this hypothesis, it is not hard to build counterexamples where convergence to $\mathcal{T}^{*}(y)$ fails. For instance, consider the following family of non-linearities:

$$\mathcal{F} = \left\{ \Psi(.) : \Psi_n(x) = \frac{1}{\sqrt{n}} tan^{-1}(nx), n \ge 1 \right\} \quad (8)$$

and experimental data $u_k = 1, \forall k$, $y \equiv 0$. Clearly $\mathcal{T}^{*}(y) = \{0\}$. However, all plants of the form $H = K$ are in $\mathcal{T}(y, N, \epsilon)$ for all N and all $\epsilon > 0$. Thus $\mathcal{T}^{*}(y) \subset \cap_{m>0} \mathcal{T}(y, N_m, \epsilon_m)$.

Applying the result above to Algorithm 1 leads to the following result:

Corollary 1: Denote by H_o the actual plant and let $e_o = \sup_y d\{T^*(y)\}$. Then, given any $e > e_o$, there exists a pair (N^*, ϵ^*) such that if $N_m > N^*$ output measurements are collected with a noise level $\epsilon < \epsilon^*$ and the number of samples N_s is selected according to (3), Algorithm 1 will generate, with probability greater than $1 - \nu$, a plant H_{id} such that $\|H_{id} - H_o\| < e$.

VI. COMPUTING BOUNDS OF THE IDENTIFICATION ERROR

In this section we briefly address the problem of computing bounds on the identification error. In the case of linear plants, this error can be bounded by the local and global diameters of information. In turn, depending on the characterization of the set $\mathcal{S}$, these quantities or upper bounds can be computed by solving a convex optimization problem. On the other hand, this approach is no longer feasible here, due to the existence of the nonlinearity. Indeed, in the case of Wiener systems, three different quantities are relevant:

1) $e_h(y, N, \mathcal{N}) \doteq \sup_{H_1, H_2 \in \mathcal{T}(y, N, \mathcal{N})} \|H_1 - H_2\|$, that is the worst case identification error for the linear portion of the plant.

2) $e_B(y, N, \mathcal{N}) \doteq \sup_{H_1, H_2 \in \mathcal{T}(y, N, \mathcal{N})} \|B_1 - B_2\|$, where $\Psi_i \doteq B_i \Phi(.)$ denotes the identified nonlinearities corresponding to the linear portion H_i.

3) $e_y(y, N, \mathcal{N}) \doteq \sup_{H_1, H_2 \in \mathcal{T}(y, N, \mathcal{N})} \|\Psi_1[(H_1 * u)_T] - \Psi_2[(H_2 * u)_T]\|$, that is the prediction error at some future time $T > N$.

In principle, computing these errors (or suitable upper bounds) leads to challenging infinite dimensional non-convex optimization problems. However, the same risk-adjusted approach used in Section IV can be used to obtain computationally tractable relaxations with risk-adjusted optimality certificates, as follows. Begin by noting that since $H_{np} \in \overline{\mathcal{BH}}_{\infty,\rho}(K)$, $\bar{\sigma}(h_i) \le K\rho^{-i}$. It follows that:

$$\|H_{np}\|_\infty \le \|\sum_{i=0}^{N-1} h_{np,i}\|_\infty + K \frac{1}{\rho^{N-1}(\rho-1)}.$$

Hence by selecting N large enough, $\|H_{1,np} - H_{2,np}\|_\infty$ can be approximated by $\bar{\sigma}(T^N_{h_{1,np}} - T^N_{h_{2,np}})$. This observation allows for computing the identification errors e_h, e_f, e_y as follows:

Algorithm 2

0.- Given $\epsilon > 0$, set $N = \lceil log(1-\rho) - log(K)/log(\rho) \rceil$. Set $e_h = 0$, $e_f = 0$, $e_y = 0$.

1.- Generate N_s samples $\{\hat{h}, B_i\}_{i=1}^{N_s}$ of the set $\mathcal{T}(y, N_m, \mathcal{N}) \times \mathcal{B}$ using Algorithm 1

2.- For each sample $\hat{h}^i$, find a matrix $\tilde{h}^i_{np} \doteq [(\tilde{h}^i_1)^T, \ldots, (\tilde{h}^i_{N-N_m})^T]^T$ such that $h^i \doteq [(\hat{h}^i)^T \ (\tilde{h}^i)^T] \in \overline{\mathcal{BH}}^N_{\infty,\rho}(K)$ by proceeding as in step 2 of Algorithm 2 in [13].

3.- Set $\hat{e}_h = \hat{e}_B = \hat{e}_y = 0$

4.- For $i = 1, \ldots, N_s - 1$ and $j = i+1, \ldots, N_s$ do:

4.1 set $\hat{e}_h = max\{\hat{e}_h, \|h^i_{np} + p_i \mathbf{G}_p - h^j - p\mathbf{G}_p\|_\infty\}$

4.2 Let:

$$e_B^{i,j} = \max_{\substack{B_i, B_j \in \mathcal{B} \\ \eta^i, \eta^j \in \mathcal{N}}} \|B_i - B_j\| \text{ subject to :}$$

$$y_k = B_i \Psi\left[(h^i \star u)_k\right] + \eta^i_k, \quad k = 0, 1, \ldots, N_m - 1$$

$$y_k = B_j \Psi\left[(h^j \star u)_k\right] + \eta^j_k \tag{9}$$

Set $\hat{e}_B = \max\{\hat{e}_B, e_B^{i,j}\}$

4.3 Let

$$e_y^{i,j} = \max_{\substack{B_i, B_j \in \mathcal{B} \\ \eta^i, \eta^j \in \mathcal{N}}} \left\| B_i \Psi\left[(h^i * u)_T\right] - B_j \Psi\left[(h^j * u)_T\right]\right\|$$

subject to :

$$y_k = B_i \Psi\left[(h^i \star u)_k\right] + \eta^i_k, \quad k = 0, 1, \ldots, N_m - 1$$

$$y_k = B_j \Psi\left[(h^j \star u)_k\right] + \eta^j_k \tag{10}$$

Set $\hat{e}_y = \max\{\hat{e}_y, e_y^{i,j}\}$

As before, if N_s is chosen according to (3) then:

$$\mathbf{Prob}\left\{\mathbf{Prob}\left\{e(h, B) \ge \hat{e} : \left\{e(h^i, B^i) \le \hat{e}\right\}_{i=1}^{N_s}\right\}\right.$$
$$\left. \le \nu\right\} \ge (1-\delta), \quad (11)$$

where $e(.,.)$ and $\hat{e}(.,.)$ denote the true worst case errors and their estimates, respectively.

VII. 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 [11], [16], [17] and references therein). A difficulty with existing approaches stems from the high dimensionality of the data: even using small size images requires processing hundreds of pixels from each frame. In [11] it has been conjectured that the problem can be decoupled into a *linear* tracking problem in a low dimensional manifold, accounting for the *dynamics* of the motion, and a nonlinear, static mapping that accounts for the changes in appearance of the target. However, no formal proof of this conjecture is available. In this section we will show that this is indeed the case, by recasting the problem of human motion modelling and tracking into a Wiener systems identification form.

The starting point is to *postulate* that human motion can be modeled as the impulse response of a Wiener system whose output is the observed image sequence. 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 work in [17] on estimating the dimension required for human motion modelling).

2.- The static nonlinearity $\Psi(\omega)$ is given by $\Psi(\omega) = \mathbf{B}[\phi(|\omega - t_1|), \phi(|\omega - t_2|), 1, \omega^T]^T$ with $\phi(x) = \exp(-0.8x^2)$ and

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

This hypothesis is motivated by the bases proposed in [11], [17] to map human silhouettes to lower dimensional spaces.

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, partially shown in Fig. 2(a), consists of the first 20 frames of a human walking on a treadmill, each having 1728

(a)

(b)
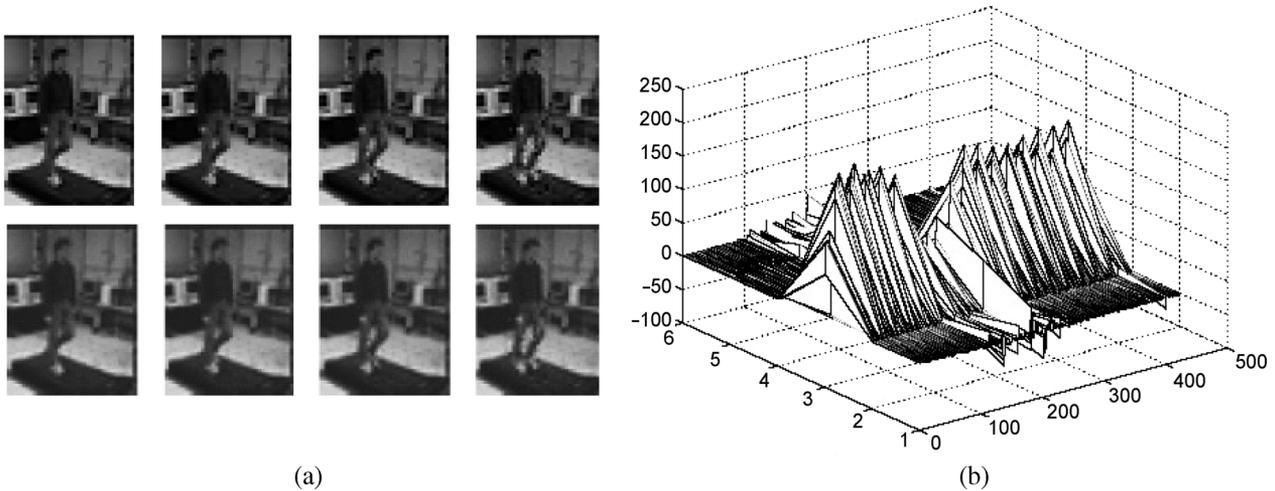
Fig. 2. (a) Top: Frames 19 through 22 of a walking person video sequence (Frames 21 and 22 were not used in the identification). Bottom: Frames 19 through 22 predicted by the identified system. (b) Surface of $\mathbf{B}$ for the nonlinear part of the model.

TABLE I
ACTUAL IDENTIFICATION ERROR VERSUS WORST CASE BOUND

Frame	Actual Error ($\|y_{actual} - y_{predicted}\|_\infty$)	Worst Case Bound
21	0.4915	1.74
22	0.9289	3.7383

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 $\{H, \mathbf{B}\}$ and hence a model explaining the experimental data. Model reducing the central interpolant found using the formulae in [18] leads to the following 5th order system for the linear portion of the model:

$$H(z) = \begin{bmatrix} \frac{-0.03z^5+0.19z^4-0.30z^3+0.22z^2-0.08z+0.02}{z^5-1.84z^4+1.13z^3+0.92z^2-1.57z+0.83} \\ \frac{-0.08z^5+0.09z^4-0.02z^3-0.08z^2+0.07z-0.03}{z^5-1.84z^4+1.13z^3+0.92z^2-1.57z+0.83} \\ \frac{0.14z^5-0.35z^4+0.27z^3+0.12z^2-0.30z+0.16}{z^5-1.84z^4+1.13z^3+0.92z^2-1.57z+0.83} \end{bmatrix}^T \quad (12)$$

The corresponding static output nonlinearity is given by $\mathbf{B}\Psi$, where a surface plot of the matrix $\mathbf{B}$ is shown in Fig. 2(b). Here the x, y axes correspond to the index of the matrix and the z axis to the matrix value. 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 Fig. 2(a). As illustrated there, the system is able to correctly predict the appearance of the target in frames 21 and 22, not used for training. Table I shows a comparison between the actual identification errors in the predicted image and the worst-case bounds computed using Algorithm 2.

VIII. CONCLUSION

In this technical note we propose an algorithm for set membership identification of Wiener systems using time-domain data. As shown in the technical note, although the problem is known to be generically NP hard, exploiting recently introduced results on sampling of transfer functions leads to a tractable 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

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

of human motion in a sequence of two-dimensional images. By modelling 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.

REFERENCES

[1] T. Wigren, "Convergence analysis of recursive identification algorithms based on the nonlinear Wiener model," *IEEE Trans. Automat. Control*, vol. 39, no. 11, pp. 2191–2206, Nov. 1994.
[2] D. Westwick and M. Verhaegen, "Identifying MIMO Wiener systems using subspace model identification methods," *Signal Processing*, vol. 52, pp. 235–258, 1996.
[3] B. Haverkamp, C. Chou, and M. Verhaegen, "Subspace identification of continuous-time Wiener models," in *Proc. IEEE Conf. Decision Control*, 1998, pp. 1846–1847.
[4] W. Greblicki, "Nonparametric approach to Wiener system identification," *IEEE Trans. Circuits Syst.*, vol. 44, no. 6, pp. 538–545, Jun. 1997.
[5] R. Raich, T. Zhou, and M. Viberg, "Subspace based approaches for Wiener system identification," *IEEE Trans. Automat. Control*, vol. 50, no. 10, pp. 1629–1634, Oct. 2005.
[6] E. Bai, "An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol. 34, pp. 333–338, 1998.
[7] E. W. Bai, "A blind approach to the Hammerstein-Wiener model identification," *Automatica*, vol. 38, pp. 967–979, 2002.
[8] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertain: An overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, 1991.
[9] W. Ma, H. Lim, M. Sznaier, and O. Camps, "Risk adjusted identification of Wiener systems," in *Proc. 45th IEEE Conf. Dec. Control*, San Diego, CA, 2006, pp. 2512–2517.

[10] M. Sznaier, "Computational complexity analysis of set membership identification of Hammerstein and Wiener systems," *Automatica*, vol. 45, no. 3, pp. 701–705, Mar. 2009.

[11] V. Morariu and O. Camps, "Modelling correspondences for multi camera tracking using nonlinear manifold learning and target dynamics," in *Proc. IEEE Comp. Vision Pattern Recogn. Conf. (CVPR)*, New York, 2006, pp. 545–552.

[12] J. Chen and G. Gu, *Control Oriented System Identification, an $\mathcal{H}_\infty$ Approach*. New York: Wiley, 2000.

[13] M. Sznaier, C. M. Lagoa, and M. Mazzaro, "An algorithm for sampling balls in $\mathcal{H}_\infty$ with applications to risk-adjusted performance analysis and model (in)validation," *IEEE Trans. Automat. Control*, vol. 50, no. 3, pp. 410–416, Mar. 2005.

[14] R. Tempo, E. W. Bai, and F. Dabbene, "Probabilistic robustness analysis: Explicit bounds for the minimum number of sampling points," in *Proc. IEEE Conf. Decision Control*, Kobe, Japan, 2006, pp. 3418–3423.

[15] J. P. Aubin and H. Frankowska, *Set Valued Analysis*. Boston, MA: Birkhauser, 1990.

[16] H. Sidenbladh, M. Black, and D. Fleet, "Stochastic tracking of 3D human figures using 2D image motion," in *Proc. Eur. Conf. Comput. Vision*, Dublin, Ireland, 2000, pp. 702–718.

[17] A. Elgammal and C. Lee, "Inferring 3D body pose from silhouettes using activity manifold learning," in *Proc. IEEE Comp. Vision Pattern Recogn. Conf. (CVPR)*, Washington, DC, 2004, pp. 681–688.

[18] P. A. Parrilo, M. Sznaier, and R. S. S. P. Na, "Mixed time/frequency domain based robust identification," *Automatica*, vol. 34, no. 11, pp. 1375–1389, 1998.

Uniform Global Asymptotic Stability of a Class of Adaptively Controlled Nonlinear Systems

Frédéric Mazenc, Marcio de Queiroz, and Michael Malisoff

Abstract—We give a new explicit, global, strict Lyapunov function construction for the error dynamics for adaptive tracking control problems, under an appropriate persistency of excitation condition. We then allow time-varying uncertainty in the unknown parameters. In this case, we construct input-to-state stable Lyapunov functions under suitable bounds on the uncertainty, provided the regressor also satisfies an affine growth condition. This lets us quantify the effects of uncertainties on both the tracking and the parameter estimation. We illustrate our results using Rössler systems.

Index Terms—Adaptive control, input-to-state stability, Lyapunov functions, uniform asymptotic stability.

I. INTRODUCTION

Consider a nonlinear system

$$\dot{x} = f(t, x, \theta, u) \qquad (1)$$

F. Mazenc is with Projet MERE INRIA-INRA, UMR Analyse des Systèmes et Biométrie INRA, Montpellier 34060, France (e-mail: frederic.mazenc@supagro.inra.fr).

M. de Queiroz is with the Department of Mechanical Engineering, Louisiana State University, Baton Rouge, LA 70803-6413 USA (e-mail: dequeiroz@me.lsu.edu).

M. Malisoff is with the Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803-4918 USA (e-mail: malisoff@lsu.edu).

Digital Object Identifier 10.1109/TAC.2009.2013053

where θ is a vector of uncertain constant parameters. The *adaptive tracking control* problem for (1) is: Given a sufficiently smooth reference trajectory $x_r(t)$, find a dynamic feedback

$$u = u(t, x, \hat{\theta}), \quad \dot{\hat{\theta}} = \tau(t, x, \hat{\theta}) \qquad (2)$$

where $\hat{\theta}$ is the estimate of θ, that ensures that $x_r(t) - x(t) \to 0$ as $t \to \infty$ while keeping all closed-loop signals bounded. *In general*, solving the adaptive tracking problem does not guarantee that $\theta - \hat{\theta}(t) \to 0$ as $t \to \infty$; i.e., parameter identification is not assured. In fact, one does not know in general whether $\hat{\theta}$ even converges to a *constant* vector [5].

Persistency of excitation (PE) has been linked to the asymptotic stability of adaptive systems [13]. PE establishes that a necessary (and sometimes sufficient) condition for parameter identification is that the reference trajectory be *sufficiently rich* so that the regressor satisfies a PE inequality [3] along the reference trajectory. For large classes of systems, PE implies that tracking error convergence can only happen when the adaptation law identifies the actual parameters [15]. The relationship between parameter identification, uniform asymptotic stability and PE was first shown for linear systems, and has been established for certain types of nonlinear systems as well. (Uniformity with respect to initial times has important implications for robustness. For example, this property ensures stability in the face of persistent disturbances [2] and provides rate of convergence information [12]. In general, PE is neither necessary nor sufficient for uniform asymptotic stability [13].) One notable example is the nonlinear dynamics of robot manipulators, where PE ensures asymptotic parameter error convergence under the Slotine-Li adaptive controller [15]. Recently, PE was shown to be necessary and sufficient for uniform global asymptotic stability (UGAS) of a class of nonlinear systems that includes the manipulator dynamics [6], [7].

When an adaptive controller does not yield GAS, this means that the corresponding closed-loop system does not admit a *strict* Lyapunov function (as defined precisely in the next section). However, even when the controller yields UGAS, the classical Lyapunov approach does not give an *explicit* strict Lyapunov function. Explicit strict Lyapunov functions are generally more useful than nonstrict ones when computing stability gains or quantifying the effects of uncertainty.

The present work provides a global, explicit, strict Lyapunov function construction for the error dynamics for adaptive tracking problems under a PE condition. It belongs to a family of results that transform nonstrict Lyapunov functions into explicit strict Lyapunov functions; see [9], [10] for constructions of this type for large classes of time-invariant systems. The paper [11] contains a very general result on constructing strict Lyapunov functions for nonlinear time-varying systems for which so-called auxiliary functions are known; i.e., the strict Lyapunov function construction in [11] is *nonexplicit*, unless the auxiliary functions are known.

By contrast, this note provides explicit expressions for auxiliary functions, which make our Lyapunov function completely explicit. The Lyapunov functions we obtain here are much simpler than the ones that would be obtained by applying [11]. Finally, the Lyapunov functions we provide here are lower bounded near 0 by positive definite quadratic functions, while the Lyapunov function construction of [11] would not have this property. We also use the idea of weighting functions, which have been used in other contexts [1], [4], [19]. The global strict Lyapunov-based framework can potentially generalize the UGAS proofs for adaptive systems. This note takes the first step towards this generalization.