On the behavior of autonomous Wiener systems

Ivan Markovsky and Philippe Dreesen

Department ELEC
Vrije Universiteit Brussel
1050 Brussels, Belgium
ivan.markovsky@vub.ac.be
philippe.dreesen@vub.ac.be

Abstract

Wiener systems are nonlinear dynamical systems, consisting of a linear dynamical system and a static nonlinear system in a series connection. Existing results for analysis and identification of Wiener systems assume zero initial conditions. In this paper, we consider the response of a Wiener system to initial conditions only, *i.e.*, we consider autonomous Wiener systems. Our main result is a proof that the behavior of an autonomous Wiener system with a polynomial nonlinearity is included in the behavior of a finite-dimensional linear system. The order of the embedding linear system is at most $\binom{n+d}{d}$ — the number of combinations with repetitions of d elements out of n elements — where n is the order of the linear subsystem and d is the degree of the nonlinearity. The relation between the eigenvalues of the embedding linear system and the linear subsystem is given by a rank-1 factorization of a symmetric d-way tensor. As an application of the result, we outline a procedure for exact (deterministic) identification of autonomous Wiener systems.

Keywords: block-oriented models, Wiener system, behavioral approach, system realization, nonlinear system identification.

1 Introduction

Interconnections of linear dynamic and nonlinear static systems is a popular class of nonlinear systems, refered to as block-oriented models [Giri and Bai(2010), Billings and Fakhouri(1982), Schoukens and Tiels(2017)]. Block-oriented models are simpler to identify from data and simpler to use for simulation and control due to the restriction of the nonlinear subsystems to be static. Among the variety of block-oriented models, the simplest special case is the Wiener system. A Wiener system consists of a linear system followed by a nonlinear static system. Despite of its limited modeling power in comparison to other block-oriented models, the Wiener system is the natural first step in the study of the class of block-oriented models and has practical applications.

A special case of an input-output system when the input dimension is zero is the autonomous system. To the best of our knowledge, currently there are no methods for autonomous Wiener system identification. The existing methods depend on a persistently exciting input and can not be used in the autonomous case (which corresponds to a response due to nonzero initial conditions and zero input). In contrast, linear time-invariant identification methods such as the prediction error and subspace methods can deal seamlessly with the autonomous case.

Our main result is that an autonomous Wiener system with a polynomial nonlinearity is embedded in a finite-dimensional linear system. In order to outline the result, consider the autonomous Wiener system $\mathscr{B}_{\mathrm{w}}$, shown in Figure 1. It consists of an order-n linear time-invariant subsystem $\mathscr{B}$ and a degree-d polynomial nonlinearity g. We prove that $\mathscr{B}_{\mathrm{w}}$ is included in a linear time-invariant system of order $n_{\mathrm{w}} \leq \binom{n+d}{d}$ — the number of combinations with repetitions of d elements out of n elements. Moreover, there is a relation between the eigenvalues of the embedding system and the eigenvalues of $\mathscr{B}$: an eigenvalue of the embedding system is a product of up to d eigenvalues of $\mathscr{B}$. This relation is characterized by a rank-1 factorization of a symmetric d-way tensor, constructed from the eigenvalues of the embedding system.

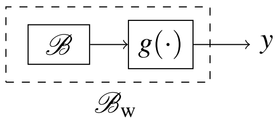

Figure 1: An autonomous Wiener system $\mathscr{B}_{\mathrm{w}}$ is a series connection of an autonomous linear time-invariant system $\mathscr{B}$ and a nonlinear static system g. In this paper, g is polynomial. We prove that in this case the behavior of $\mathscr{B}_{\mathrm{w}}$ is a subset of the behavior of a finite-dimensional linear time-invariant system.

2 Notation

The notation used in the paper is standard: $\mathbb{R}$ is the set of real values, $\mathbb{C}$ is the set of complex values, and $\mathbb{N}$ is the set of natural numbers. The set of scalar real-valued signals over $\mathbb{N}$ is denoted by $\mathbb{R}^{\mathbb{N}}$. An autonomous linear time-invariant system $\mathscr{B}$ admits a minimal state space representation

$$\mathscr{B} = \mathscr{B}(A,c) := \{\, z \in \mathbb{R}^{\mathbb{N}} \mid \text{there is } x \text{, such that } \sigma x = Ax,\ z = cx,\ x(1) \in \mathbb{R}^n \,\}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ and $c \in \mathbb{R}^{1\times n}$ are parameters of the system and σ is the shift operator $(\sigma x)(t) = x(t+1)$. The eigenvalues $\lambda_1,\ldots,\lambda_n$ of A are invariant of the representation and are, therefore, a property of the system $\mathscr{B}$.

In this paper, we consider a single output Wiener system and assume that the eigenvalues of its linear subsystem are distinct. In this case, the linear time-invariant subsystem admits a sum-of-damped-exponentials representation

$$\mathscr{B} = \mathscr{B}(\lambda) := \Big\{ z \in \mathbb{R}^{\mathbb{N}} \mid z = \sum_{i=1}^{n} \alpha_i \exp_{\lambda_i},\ \alpha \in \mathbb{C}^n \Big\}, \tag{2}$$

where $\exp_{\lambda_i}(t) := \lambda_i^t$ and λ is the vector of the system's eigenvalues $\lambda = \begin{bmatrix} \lambda_1 & \cdots & \lambda_n \end{bmatrix}^\top \in \mathbb{C}^n$.

Restricting ourselves to the single output case with distinct eigenvalues simplifies the notation. The results in the paper can be generalized mutatis mutandis to the case of multi-output systems. Dealing with repeated eigenvalues requires a generalization of the sum-of-damped-exponentials representation, which complicates the analysis but does not change our main results.

The static nonlinearity g is a dth order polynomial, represented by a given monomial basis v

$$y = g(z) := \theta^\top v(z), \qquad \text{where} \quad v(z) = \begin{bmatrix} z^0 & z^1 & \cdots & z^d \end{bmatrix}^\top \in \mathbb{R}^{d+1}. \tag{3}$$

Putting together (2) and (3), we obtain the autonomous Wiener system

$$\mathscr{B}_{\mathrm{w}}(\lambda, \theta) := \{ y \in \mathbb{R}^{\mathbb{N}} \mid (2,3) \text{ hold for } \alpha \in \mathbb{C}^n \},$$

parameterized by the vector of the coefficients $\theta = \begin{bmatrix} \theta_0 & \theta_1 & \cdots & \theta_d \end{bmatrix}^\top \in \mathbb{R}^{d+1}$ of the nonlinear part and the eigenvalues λ of the linear part.

3 Main result

Theorem 1. *Consider an autonomous Wiener system $\mathscr{B}_w(\lambda, \theta)$ with order-n linear subsystem and a degree-d nonlinear subsystem. Assume that the eigenvalues λ are distinct* , i.e., *$\lambda_i \neq \lambda_j$ for all i and 1 is not an eigenvalue of $\mathscr{B}(\lambda)$. Then, there is an autonomous linear time-invariant system $\mathscr{B}(\lambda_w)$ with eigenvalues $\lambda_w \in \mathbb{C}^{n_w}$, where*

$$n_w \leq \bar{n}_w := \binom{n+d}{d} = \frac{(n+1)(n+2)\cdots(n+d)}{d!}, \tag{4}$$

such that

$$\mathscr{B}_w(\lambda, \theta) \subseteq \mathscr{B}(\lambda_w). \tag{5}$$

The eigenvalues λ_w of the embedding system $\mathscr{B}(\lambda_w)$ are products of d elements of the set $\{\lambda_0, \lambda_1 \ldots, \lambda_n\}$, where $\lambda_0 := 1$, i.e., *there are indices $k_{i,1}, \ldots, k_{i,d} \in \{0, 1, \ldots, n\}$, such that*

$$\lambda_{w,i} = \prod_{j=1}^{d} \lambda_{k_{i,j}}, \quad \text{for } i = 1, \ldots, n_w. \tag{6}$$

Proof. By definition

$$\mathscr{B}(\lambda_{\mathrm{w}}) := \Big\{ y \in \mathbb{R}^{\mathbb{N}} \mid y = \sum_{i=1}^{n_{\mathrm{w}}} \beta_i \exp_{\lambda_{\mathrm{w},i}},\ \beta \in \mathbb{C}^{n_{\mathrm{w}}} \Big\}, \tag{7}$$

In order to prove the relation (5), we compare the output of (7) with the expression for the output of the autonomous Wiener system $\mathscr{B}_{\mathrm{w}}(\lambda, \theta)$.

Consider a general basis element

$$v_j\big(z(t)\big) = \big(z(t)\big)^j = \left(\sum_{i=1}^{n} \alpha_i \lambda_i^t\right)^j.$$

For $j = 0$ and 1, $v_0 = 1$ and $v_1 = y$ are of the form of sum-of-damped-exponentials with $n_0 = 1$ and $n_1 = n$ exponents, respectively. For $j > 1$, v_j is also of the form of a sum-of-damped-exponentials with exponents that are products of j elements of the set λ, *i.e.*,

$$v_j\big(z(t)\big) = \sum_{i=1}^{n_j} \gamma_i \mu_{i,j}^t, \quad \text{where } \mu_{i,j}^t = \prod_{\ell=1}^{j} \lambda_{k_{i,j,\ell}}$$

for some indices $k_{i,j,\ell} \in \{1, \ldots, n\}$. The number of terms n_j is equal to the number of combinations with repetitions of j elements out of the n elements of λ. Therefore,

$$n_j = \binom{n+j-1}{j} = \frac{(n+j-1)\cdots n}{j!}.$$

Consider now the output

$$y(t) = g\big(z(t)\big) = \theta^\top v\big(z(t)\big).$$

It is also of the form of a sum-of-damped-exponentials

$$y(t) = \sum_{i=1}^{n_\mathrm{w}} \zeta_i \lambda_{\mathrm{w},i}^t, \qquad \text{where } \{\lambda_{\mathrm{w},1}, \ldots, \lambda_{\mathrm{w},n_\mathrm{w}}\} = \bigcup_{i=0}^{j} \bigcup_{j=0}^{d} \mu_{i,j}. \tag{8}$$

The elements of λ_w are products of d elements of the set $\{1, \lambda_1, \ldots, \lambda_n\}$. The number of such products is

$$\bar{n}_\mathrm{w} = \sum_{j=0}^{d} n_j = \binom{n+d}{d} = \frac{(n+1)(n+2)\cdots(n+d)}{d!}.$$

The number of distinct elements n_w, which is the order of $\mathscr{B}(\lambda_\mathrm{w})$ is therefore upper bounded by $\bar{n}_\mathrm{w}$.

We've shown that both the output (7) of the autonomous Wiener system $\mathscr{B}_\mathrm{w}(\lambda, \theta)$ and the output (8) of the linear system $\mathscr{B}(\lambda_\mathrm{w})$ are of the form of sum-of-damped-exponentials with the same exponents. The coefficients β_i in (7) however are restricted only by the condition of being in complex conjugate pairs (since the signal is real), while the coefficients ζ_i in (8) range over an n-dimensional manifold of $\mathbb{C}^{n_\mathrm{w}}$. This proves (5). $\square$

Next, we give an alternative characterization of the relation (6) between the eigenvalues of λ and λ_w, using the notation "$\circ$" for the vector outer product.

Corollary 2 (Link between λ_w and λ)**.** *The symmetric, rank-1, d-way tensor*

$$T := \underbrace{\lambda \circ \lambda \circ \cdots \circ \lambda}_{d \text{ times}},$$

has as unique elements $\lambda_{w,1}, \ldots, \lambda_{w,n_w}$.

4 Application of the result in system identification

The problem considered in this section is: Given a monomial basis v, a finite trajectory

$$y_{\mathrm{d}} = \big(y_{\mathrm{d}}(1), \dots, y_{\mathrm{d}}(T)\big)$$

of an autonomous Wiener system $\mathscr{B}_{\mathrm{w}}(\lambda,\theta)$, and the order n of its linear part, find parameters $\widehat{\lambda}, \widehat{\theta}$, such that

$$\mathscr{B}_{\mathrm{w}}(\lambda,\theta) = \mathscr{B}_{\mathrm{w}}(\widehat{\lambda},\widehat{\theta}).$$

Theorem 1 suggests the following solution method:

1. identify the embedding system $\mathscr{B}(\lambda_{\mathrm{w}})$ from the given output data,
2. compute the linear subsystem $\mathscr{B}(\lambda)$ from $\mathscr{B}(\lambda_{\mathrm{w}})$, and
3. compute the nonlinear subsystem g from $\mathscr{B}(\lambda_{\mathrm{w}})$ and $\mathscr{B}(\lambda)$.

Assuming that the given trajectory y_{d} is persistently exciting of order n_{w}, the embedding system $\mathscr{B}(\lambda_{\mathrm{w}})$ is identifiable from y_{d} [Willems *et al.*(2005)]. The remaining problems, resolved in steps 2 and 3, are to find from the identified system $\mathscr{B}(\lambda_{\mathrm{w}})$, the linear and nonlinear subsystems of the autonomous Wiener system. Note that due to exchange of gain, between $\mathscr{B}$ and g, the linear and nonlinear subsystems are not identifiable from the data alone. As shown next, however, the eigenvalues of $\mathscr{B}$ can be determined uniquely and g can be determined up to a scaling factor.

Identification of $\mathscr{B}(\lambda_{\mathrm{w}})$ from the given output data

The identification of an autonomous linear time-invariant system $\mathscr{B}_{\mathrm{w}}(\lambda,\theta)$ from the finite trajectory $y_{\mathrm{d}} \in \mathscr{B}_{\mathrm{w}}(\lambda,\theta)$ is a classical problem, see for example, [Kung(1978), Kumaresan and Tufts(1982)] and [Markovsky(2019), Section 5.1.3]. One possible solution [Kumaresan and Tufts(1982)] is to form the Hankel matrix

$$\mathscr{H}_{n_{\mathrm{w}}+1}(y_{\mathrm{d}}) := \begin{bmatrix} y_{\mathrm{d}}(1) & y_{\mathrm{d}}(2) & \cdots & y_{\mathrm{d}}(T-n_{\mathrm{w}}) \\ y_{\mathrm{d}}(2) & y_{\mathrm{d}}(3) & \cdots & y_{\mathrm{d}}(T-n_{\mathrm{w}}+1) \\ y_{\mathrm{d}}(3) & y_{\mathrm{d}}(4) & \cdots & y_{\mathrm{d}}(T-n_{\mathrm{w}}+2) \\ \vdots & \vdots & & \vdots \\ y_{\mathrm{d}}(n_{\mathrm{w}}+1) & y_{\mathrm{d}}(n_{\mathrm{w}}+2) & \cdots & y_{\mathrm{d}}(T) \end{bmatrix}$$

and compute its left kernel (which can be shown to be one dimensional)

$$p\mathscr{H}_{n_{\mathrm{w}}+1}(y_{\mathrm{d}}) = 0.$$

The roots of the polynomial

$$p(s) = p_0 + p_1 s + \cdots + p_{n_\mathrm{w}} s^{n_\mathrm{w}}$$

are the eigenvalues of the embedding system. Another solution (called Kung's method [Kung(1978)]) is based on realization theory: 1) compute the rank revealing factorization

$$\mathscr{H}_{n_\mathrm{w}+1}(y_\mathrm{d}) = \mathscr{O}\mathscr{C}, \quad \text{with } \mathscr{O} \in \mathbb{R}^{L\times n_\mathrm{w}} \text{ and } \mathscr{C}^{n_\mathrm{w}\times(T-L)}$$

of the Hankel matrix $\mathscr{H}_L(y_\mathrm{d})$, where L is a design parameter, satisfying the constraints

$$n_\mathrm{w} + 1 \leq L \leq T - n_\mathrm{w}$$

and 2) solve the system of linear equations $\overline{\mathscr{O}}\widehat{A} = \underline{\mathscr{O}}$, for $\widehat{A}$, where $\overline{\mathscr{O}}$ is the matrix $\mathscr{O}$ with the first row removed and $\underline{\mathscr{O}}$ is the matrix $\mathscr{O}$ with the last row removed. The eigenvalues of $\widehat{A}$ are the eigenvalues of the embedding system.

The minimal number of samples needed for the identification of the system $\mathscr{B}(\lambda_\mathrm{w})$ is $T_{\min} = 2n_\mathrm{w} + 1$. The identification data, however, can be collected from n_w experiments with $n_\mathrm{w} + 1$ samples instead of a single experiment with $T_{\min}$ samples. Let $y_\mathrm{d}^1, \dots, y_\mathrm{d}^{n_\mathrm{w}}$ be the data of the multiple experiments of length $n_\mathrm{w} + 1$. Then, the identification procedure is modified by replacing the Hankel matrix $\mathscr{H}_{n_\mathrm{w}+1}(y_\mathrm{d})$ by the matrix

$$\begin{bmatrix} y_\mathrm{d}^1 & \cdots & y_\mathrm{d}^{n_\mathrm{w}} \end{bmatrix}$$

of the stacked next to each other responses. More generally, using date from multiple experiments of length $T_1, \dots, T_{n_\mathrm{w}} > n_\mathrm{w}$, the identification method is based on the computation of the left kernel or rank revealing factorization of the mosaic Hankel matrix [Heinig(1995), Markovsky and Pintelon(2015)]

$$\mathscr{H}_{n_\mathrm{w}+1}(y_\mathrm{d}^1, \dots, y_\mathrm{d}^{n_\mathrm{w}}) := \begin{bmatrix} \mathscr{H}_{n_\mathrm{w}+1}(y_\mathrm{d}^1) & \cdots & \mathscr{H}_{n_\mathrm{w}+1}(y_\mathrm{d}^{n_\mathrm{w}}) \end{bmatrix}.$$

Computation of the linear subsystem $\mathscr{B}(\lambda)$ from $\mathscr{B}(\lambda_\mathrm{w})$

After finding $\mathscr{B}(\lambda_\mathrm{w})$, the next step is the computation of the linear subsystem. We are interested in the transition from λ_w to λ, *i.e.*, extracting the linear subsystem $\mathscr{B}(\lambda)$ from $\mathscr{B}(\lambda_\mathrm{w})$. Using Corollary 2, we can find λ by computing a rank-1 factorization of a symmetric tensor T constructed from λ_w. Checking whether T has rank equal to one can be done by checking the rank of the d unfoldings of the tensor: T is rank-1 if and only if all unfoldings of T are rank-1 [De Lathauwer *et al.*(2000), De Lathauwer(1999)].

Another characterization of (6) that leads to a more efficient method is given in terms of the "frequencies" $\omega_i := \angle\lambda_i$ and $\omega_{\mathrm{w},i} := \angle\lambda_{\mathrm{w},i}$ of $\mathscr{B}(\lambda)$ and $\mathscr{B}(\lambda_\mathrm{w})$, respectively. From (6), we have the following linear relation among the $\omega_{\mathrm{w},i}$'s and the ω_i's

$$\omega_{\mathrm{w},i} = \sum_{j=1}^{d} \omega_{k_{i,j}} \pmod{2\pi}.$$

Therefore, there is an $n_\mathrm{w} \times n$ matrix K, such that

$$\omega_\mathrm{w} = K\omega \pmod{2\pi}. \tag{9}$$

Relation (9) shows that the problem of extracting $\mathscr{B}(\lambda)$ from $\mathscr{B}(\lambda_\mathrm{w})$ can be solved by computing the frequencies of $\mathscr{B}(\lambda_\mathrm{w})$ and solving a system of linear equations. The ordering of the ω_w's however is unknown, so that all permutations of the ω_w's should be tested for existence of an exact solution. (The order of the to-be-found frequencies ω is not important.) This method requires the same number of subproblems to-be-solved as in the procedure using Corollary 2. The subproblem (9) however is a linear system, which is simpler and faster to solve than the rank-1 factorization of a symmetric tensor.

Computation of the nonlinear subsystem g from $\mathscr{B}(\lambda_\mathrm{w})$ and $\mathscr{B}(\lambda)$

Finally computing the nonlinear function g requires a simultaneous rank-1 factorization of d tensors. General theory and methods, called structured data fusion, for solving simultaneous tensor factorization problems is developed in [Sorber *et al.*(2015), Vervliet *et al.*(2016)]. In general, the structured data fusion problem has no analytical solution and requires iterative solution methods. Applied to the autonomous Wiener system identification problem, however, when g contains a first and/or second order terms the structured data fusion problem has a trivial solution. When g has a first order term, the coefficients θ can be obtained directly from the coefficients γ in (8) without extra computations. When g has a second order term, the coefficients θ can be obtained from the coefficients γ in (8) by a Cholesky factorization of a symmetric matrix constructed from the γ's.

5 Numerical example

The autonomous Wiener system $\mathscr{B}_\mathrm{w}(\lambda, \theta)$ used in the simulation example consists of a second order linear subsystem with eigenvalues $\lambda_{1,2} = -0.5 \pm 0.7\mathbf{i}$ and a (dead zone) nonlinear subsystem defined by the third order polynomial

$$g(z) = \theta_0 + \theta_1 z + \theta_2 z^2 + \theta_3 z^3,$$

with coefficients

$$\theta = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^\top.$$

According to Theorem 1, $\mathscr{B}_\mathrm{w}(\lambda, \theta)$ is included in a linear time-invariant system of order

$$n_\mathrm{w} \leq \binom{n+d}{d} = \binom{5}{3} = 10. \tag{10}$$

In order to verify this property empirically, we generate a $T = 25$ samples long trajectory $y_{\rm d}$ of the system $\mathscr{B}_{\rm w}(\lambda,\theta)$ due to a random initial condition and check the rank of the square Hankel matrix $\mathscr{H}_{13}(y_{\rm d})$. The fact that $\operatorname{rank}\big(\mathscr{H}_{13}(y_{\rm d})\big) = 10$ confirms (10).

Next, we verify (6), namely the statement that the eigenvalues $\lambda_{\rm w}$ of the embedding system are products of up to d eigenvalues λ of the linear subsystem. First, using an exact identification method, *e.g.*, Kung's method described in Section 4, we obtain the linear time-invariant system $\mathscr{B}(\lambda_{\rm w})$ that contains $\mathscr{B}_{\rm w}(\lambda,\theta)$. Then, we form the set

$$\{\, \lambda_{k_0}\lambda_{k_1}\cdots\lambda_{k_d} \mid \lambda_0 := 1 \text{ and } k_0,k_1,\dots,k_d \in \{\,0,1\dots,n\,\}\,\} \tag{11}$$

of all products of up to d eigenvalues of $\mathscr{B}$. Finally, we compare the identified eigenvalues $\lambda_{\rm w}$ and the theoretically predicted ones (11). For the simulation example, described above, we confirm that they coincide, see Figure 2.

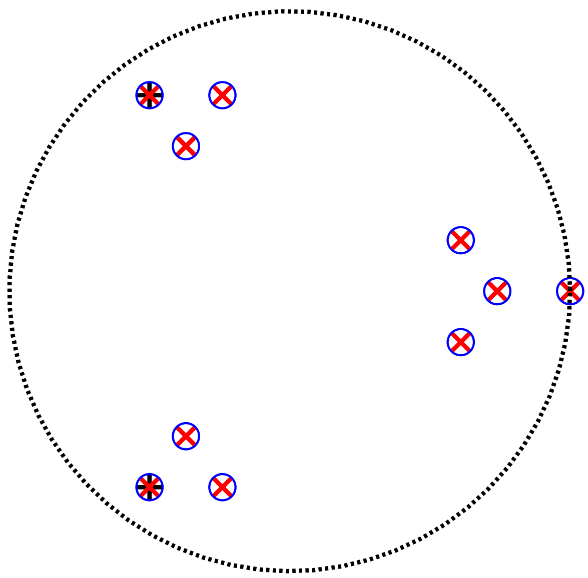

Figure 2: The eigenvalues $\lambda_{\rm w}$ (plotted as red $\times$'s) of the embedding system coincide with the products (plotted as blue $\circ$'s) of up to d eigenvalues of the linear subsystem $\mathscr{B}(\lambda)$. $+$ — eigenvalues λ, dotted line — unit circle.

For higher values of n and d it is possible to obtain after identification from data only a subset of the eigenvalues of $\mathscr{B}_{\rm w}$. This is due to ill-conditioning of the identification problem and the finite precision arithmetic used in the numerical computations. The problem can be partially resolved using data of multiple experiments generated by properly selected initial conditions (design of the experiments). However, the ill-conditioning of the system identification problem remains an important practical issue that will be addressed elsewhere.

6 Conclusion

We showed that the behavior of an autonomous Wiener system with polynomial nonlinearity is included in the behavior of a finite-dimensional linear time-invariant system. The order of the embedding linear system depends combinatorially on the order of the linear subsystem and the degree of the static nonlinearity. The relation between the eigenvalues of the embedding system and the linear subsystem is given by a rank-1 factorization of a symmetric tensor: the unique elements of the tensor are the eigenvalues of the embedding system and the factors contain the eigenvalues of the linear subsystem. The result suggests an autonomous Wiener system identification procedure that is based on linear time-invariant system identification followed by a rank-1 tensor factorization. Challenges that need to be addressed in order to make this procedure practically useful are ill-conditioning of the linear identification step and combinatorial number of rank-1 factorization problems that have to be solved for the computation of the eigenvalues of the linear subsystem from the eigenvalues of the identified system.

7 Acknowledgments

The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007–2013) / ERC Grant agreement number 258581 "Structured low-rank approximation: Theory, algorithms, and applications" and Fund for Scientific Research Vlaanderen (FWO) projects G028015N "Decoupling multivariate polynomials in nonlinear system identification" and G090117N "Block-oriented nonlinear identification using Volterra series"; and Fonds de la Recherche Scientifique (FNRS) – FWO Vlaanderen under Excellence of Science (EOS) Project no 30468160 "Structured low-rank matrix / tensor approximation: numerical optimization-based algorithms and applications".

References

[Billings(2013)] Billings, S. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.

[Billings and Fakhouri(1982)] Billings, S. and Fakhouri, S. (1982). Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, **18**(1), 15–26.

[De Lathauwer(1999)] De Lathauwer, L. (1999). *Signal processing based on multilinear algebra*. Ph.D. thesis, K.U.Leuven, ESAT-SISTA.

[De Lathauwer *et al.*(2000)] De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, **21**(4), 1253–1278.

[Giri and Bai(2010)] Giri, F. and Bai, E.-W. (2010). *Block-oriented nonlinear system identification*. Springer.

[Heinig(1995)] Heinig, G. (1995). Generalized inverses of Hankel and Toeplitz mosaic matrices. *Linear Algebra Appl.*, **216**(0), 43–59.

[Khalil(1996)] Khalil (1996). *Nonlinear Systems*. Prentice Hall.

[Kumaresan and Tufts(1982)] Kumaresan, R. and Tufts, D. (1982). Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. *IEEE Trans. Acoust., Speech, Signal Process.*, **30**(6), 833–840.

[Kung(1978)] Kung, S. (1978). A new identification method and model reduction algorithm via singular value decomposition. In *Proc. 12th Asilomar Conf. Circuits, Systems, Computers*, pages 705–714, Pacific Grove.

[Markovsky(2017)] Markovsky, I. (2017). Application of low-rank approximation for nonlinear system identification. In *25th IEEE Mediterranean Conference on Control and Automation*, pages 12–16, Valletta, Malta.

[Markovsky(2019)] Markovsky, I. (2019). *Low-Rank Approximation: Algorithms, Implementation, Applications*. Springer, second edition.

[Markovsky and Pintelon(2015)] Markovsky, I. and Pintelon, R. (2015). Identification of linear time-invariant systems from multiple experiments. *IEEE Trans. Signal Process.*, **63**(13), 3549–3554.

[Noël and Kerschen(2013)] Noël, J.-P. and Kerschen, G. (2013). Frequency-domain subspace identification for nonlinear mechanical systems. *Mechanical Systems and Signal Processing*, **40**(2), 701–717.

[Schoukens and Tiels(2017)] Schoukens, M. and Tiels, K. (2017). Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, **85**, 272–292.

[Sorber *et al.*(2015)] Sorber, L., Van Barel, M., and De Lathauwer, L. (2015). Structured data fusion. *IEEE Journal of Selected Topics in Signal Processing*, **9**(4), 586–600.

[Vervliet *et al.*(2016)] Vervliet, N., Debals, O., Sorber, L., Van Barel, M., and De Lathauwer, L. (2016). Tensorlab v3.0. `http://www.tensorlab.net/`.

[Willems *et al.*(2005)] Willems, J. C., Rapisarda, P., Markovsky, I., and De Moor, B. (2005). A note on persistency of excitation. *Control Lett.*, **54**(4), 325–329.