Wienerization of systems in nonlinear control canonical normal form

Hans A. Engmark*, Håvard B. Bjørkøy*, Adil Rasheed, and Damiano Varagnolo

Abstract— **We extend the concept of model approximation via wienerization to systems in nonlinear control canonical normal form. We elaborate on the conditions for, and implications of, analytically separating nonlinear input affine dynamical systems in state space form in a static part plus a dynamic one. In doing so, we discuss under which conditions Wiener models may approximate the resulting models well. More precisely, we report that a specific bijective transformation of the original nonlinear model will separate the system into a multidimensional state space structure for which it is possible to compare nonlinear Wiener control against linear control for underactuated nonlinear systems. We finally assess how the former type of control has better closed-loop performance than the latter by means of quantitative examples.**

I. Introduction

We consider a situation where the inputs of an asymptotically stable system are known to change only very slowly in time with respect to the characteristic time constant of the system. The structure and the parameters of such a model are unknown, and if the plant is operated in closed loop as close as possible to an equilibrium point, or if the transitions among operating regimes are performed slowly, the data collected will be in a static-like modality.

In other words, the measured trajectory is governed more by the steady state response of the system (i.e., static gain) than by its transient response. Data-driven modelling in such a situation, for which the data is non-persistently exciting (PE), is to the best of our knowledge problematic.

Inspired by this practical problem, we propose and analyze system representations that allow to formally separate and investigate the static and dynamic properties of a system, and we investigate how one may build model approximations that leverage such a separation. By doing so, we aim at contributing towards answering the question of what information may be extracted from non-PE data.

We consider what can be said about a model, given that one just has information about its equilibria (potentially obtained by means of opportune identification efforts, expert knowledge, physics based simulators, or other suitable sources of information for the specific system). The focus of the paper is however not on obtaining this information, but rather presenting a specific model structure and discussing how it may be approximated, and the control capabilities such approximations offer.

Our first main purpose is thus to clarify how the set of equilibria of the system may be seen (and used) as structural information about the system. Our second purpose stems from the consequent consideration: how may such structural information about the equilibria be used for feedback control purposes? Again, intuitively for now (but more formally below), knowing the equilibria of the model should enable drafting at least part of a model-based control strategy (potentially nonlinearly, potentially only partially solving the problem of controlling the plant).

Literature review: The concepts above strongly relate to Hammerstein Wiener modelling, due to their capability of embedding nonlinear static information. Hammerstein and Wiener models are simple nonlinear models that have linear parts but nonlinear equilibria, thus emphasizing nonlinear static effects. For these reasons, and the advantages mentioned above, such models have proven useful in many areas, such as chemical processes, biological processes, and signal processing [1], [11]–[13]. The controllers based on Wiener and Hammerstein models have much in common with linear controllers, but generally tend to exceed linear model based controllers for nonlinear systems.

Data-driven identification of Wiener models is both a mature and active field of research [6], [17], and it may be considered a sub-class of identification of general block-oriented systems [5]. The block structure may in many cases allow for inclusion of prior knowledge into the model [9] (e.g., a titration curve in a pH process [7]). This is also the case here, when relevant. One general and central problem when identifying Wiener models is static gain estimation, described in the recent work [16].

We also note that Wiener models are not only considered to be strictly empirical or having a physical interpretation. Other interpretations of Wiener models can be found through Koopman theory [14], or Volterra series approximations [3]. The approach here resembles and connects to the latter, though the starting point is a system in state space form.

Finally, system identification regularized with steady state gain (i.e., static information) has been successfully tested in, e.g., [4], though we note that this paper has the different focus of characterizing what the static information means for the system itself.

Statement of contributions: We introduce a nonlinear transformation for systems in nonlinear control canonical normal form that separates the system into a dynamical part with constant gain followed by a memory-less nonlinear part. We characterize when this transformation exists and

* the authors have contributed equally to this work. All the authors are with the Department of Engineering Cybernetics at the Norwegian University of Science and Technology, Trondheim, Norway. Emails: { hans.a.engmark — havard.bjorkoy — adil.rasheed — damiano.varagnolo } @ntnu.no. Damiano Varagnolo is also with the Department of Information Engineering, University of Padova, Italy. This project received funds from the Norwegian Research Council through the Knowledge Building project number 326802, “Electrical Conditions in Submerged Arc Furnaces - Identification and Improvement”.

is invertible, and generalize the results from [2] to systems that also can be underactuated, in this way broadening the class of systems that are known to be separable to the class of systems expressable in nonlinear control canonical normal form (i.e., models containing integrators).

Based on this, we define the process of wienerization for this class of systems and connect this operation to linearization, in this way finding results that are of immediate use for the design of control strategies. We then derive conditions for when the dynamical part of a wienerized model is equivalent to the linearization of the original system. Finally we show that control based on wienerization may be compared against control based on linearization, and demonstrate this with numerical examples that confirm the expected performance gain, especially for slowly varying input references.

In this way we show how Wiener models are particularly suitable for situations where modelling should start from non-PE data, and how Wiener models may exceed the generalization capabilities of linear ones by extracting meaningful nonlinear static information (note that leveraging non-PE information is particularly useful in the considered scenario when static behaviour is dominant in the plant). Non-PE data is available as information or estimates on, e.g., equilibria or gain of a system computed through a physics based simulator as in a hybrid modeling setting [15].

Structure of the manuscript: Section II describes how dynamic and static effects may be separated, exemplifies the limitations of previous works, and defines a transformation that can be used for separating underactuated systems. Section III describes how a linearization and a wienerization of the original system are similar and connected, how this results in related control schemes, and how the schemes are different. Section IV includes numerical experiments that illustrate the findings from the two previous sections. Finally we make some conclusions and outline future research in Section V.

II. A MODEL STRUCTURE THAT SEPARATES STATIC AND DYNAMIC NONLINEAR COMPONENTS

We start by recalling a particular model structure, proposed and analyzed first in [2], that serves the purpose of separating static and dynamic nonlinear effects as two distinct and separately identifiable parts. This paper focuses on generalizing these concepts to a broader class of systems.

Consider the following nonlinear input affine system,

$$\dot{y} = f(y) + G(y)u \,, \tag{1}$$

where $u \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$, and where both $f(y) : \mathbb{R}^n \mapsto \mathbb{R}^n$ and $G(y) : \mathbb{R}^n \mapsto \mathbb{R}^{n\times n}$ are smooth functions of class C^1. As shown in [2], to separate this system into a dynamic model that has purely dynamic nonlinearities in series with a static nonlinearity, it is required that the steady state map $h(u) : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a smooth function in C^1 that is invertible in the region of interest. This enables defining the condition for an input-output equilibrium as

$$\bar{u} := h^{-1}(\bar{y}) := -G(\bar{y})^{-1} f(\bar{y}),$$

where the *bar* notation indicate steady states (constant in time). Then the model structure (1) can be equivalently rewritten into the state space representation

$$\dot{x} = k(x)(x - Tu) \tag{2a}$$

$$y = h\left(T^{-1}x\right) \tag{2b}$$

where $x \in \mathbb{R}^n$, and $T \in \mathbb{R}^{n\times n}$ is an arbitrary invertible matrix. The representation (2) is the aforementioned separation of the static and dynamic part of the original system, as (2a) essentially contains no steady state information (it has constant gain); it is all included in (2b), the memory-less part.

The function $k(x) : \mathbb{R}^n \mapsto \mathbb{R}^{n\times n}$ in (2a) is shown in [2] to be

$$k(x) := -T\nabla h\left(T^{-1}x\right)^{-1}(G\circ h)\left(T^{-1}x\right)T^{-1}\,.$$

Representing a system as in the form (2) may thus be interpreted as rewriting the model to highlight a specific type of separation of the static and the dynamic system nonlinearities. Indeed at the equilibria we have $\bar{x} = T\bar{u}$, implying that $\bar{y} = h(\bar{u})$ may be interpreted as an equilibrium map. Moreover, as shown in [2], representation (2) can immediately be used for nonlinear control design purposes.

Consider the linearization versus the *wienerization* of (1), the latter obtained by linearizing only (2a). Starting from both these models one may design an associated linear feedback controller, and apply it to the system. However, the controller designed according to the Wiener model is designed for control in x. Thus, if the equilibrium map h is known and invertible, its inverse has to be used in the feedback loop to enable the Wiener model based controller, making it a nonlinear controller w.r.t. y.

Example: Consider the following Duffing equation, physically representing a mass-spring system with nonlinear spring stiffness (i.e., a hardening spring),

$$\dot{y}_1 = y_2 \tag{3a}$$

$$\dot{y}_2 = \frac{1}{m}\big(-cy_2 - ky_1\left(1 + a^2 y_1^2\right) + u\big) \tag{3b}$$

where y_1 is position of the spring relative to equilibrium, y_2 is the speed, m is the mass, c is a damping or friction constant, k and a are constants related to the spring stiffness, and u is some input force.

As system (3) contains an integrator, its equilibria must be so that the integrated states in equilibrium must be zero. Indeed the equilibria of system (3) correspond to $y_2 = 0$ and $u = \frac{k}{m}y_1(1+a^2y_1^2)$, and thus only y_1 can have a static non-zero component. Rewriting (3) in the form (2) is not possible from the theory summarized above, since $G(y)$ cannot be full rank for any y (being not even square). □

We also note that requiring that $\dim u = \dim y$, as needed to produce (2), is too restrictive to be useful for modelling many real-world systems that are typically underactuated, i.e., $\dim u < \dim y$.

A first contribution of this paper is to show how to define an alternative nonlinear differential equation with a structure that is inspired by (2) and that shows similar separations of static and dynamic components, that can also model situations for which $\dim u < \dim y$. We soon describe an alternative, invertible map based on the nontrivial equilibria in more general terms than above.

A. On the existence of invertible equilibrium maps

As exemplified above, systems that admit representations as in (1) do not generally admit representations of the form (2). Vice versa, systems that may be defined as in (2) may not admit a representation of the form (1). Any system that has an output equation like (2b) that is not invertible, can not be put on the form (1), i.e. can not be represented through a single, explicit ODE. This is easily seen by any attempt to do so, as from (2b) it follows that

$$\dot{y} = \nabla h\left(T^{-1}x\right)T^{-1}\dot{x}$$

for which we can conclude that it is not possible to substitute out x for y.

We argue now that asking when the two specific representations (1) and (2) are equivalent comes down to asking when an invertible equilibrium map $h(u) = y$ exists.

Consider now the equilibria $(\bar{u}, \bar{y})$ of the system (1) and the implicit function theorem, describing when there exists a mapping $h(\cdot)$ in the neighborhood of an equilibrium. In this case, the equilibria are given by

$$0 = r(\bar{u}, \bar{y}) = f(\bar{y}) + G(\bar{y})\bar{u}\,,$$

an equation that is solvable with respect to $\bar{u}$ if the matrix

$$\frac{\partial r}{\partial \bar{u}}(\bar{u}, \bar{y}) = G(\bar{y})$$

is full rank, meaning that a function $\bar{u} = p(\bar{y})$ exists near the equilibrium, giving the equilibria values $\bar{u}$ from $\bar{y}$. Solvability with respect to $\bar{y}$ equivalently requires that

$$\frac{\partial r}{\partial \bar{y}}(\bar{u}, \bar{y}) = \frac{\partial f}{\partial y}(\bar{y}) + \bar{u}^{\intercal}\nabla_y G(\bar{y})$$

is full rank, implying existence of a function $y = q(u)$ near the equilibrium.

When both $p(y), q(u)$ exist, then the function h exists and is invertible. Clearly, the existence of both $p(y)$ and $q(u)$ from (1) is generally not possible, as for example a non-square $G(y)$ will lead to an undefined $p(y)$. However, in such cases we may define appropriate transformations to still separate dynamics and statics of systems as (1), though the resulting form is somewhat different from (2).

B. Equilibria of underactuated systems

To extend the ideas of separation for static and dynamic nonlinearities to systems where $\dim u = m < n$ (and thus the cases where $G(y)$ is not square and the equilibrium map is thus non-invertible) we consider the following as a prototype of such a system, recalled from [18]:

Definition 1 A system is said to be in nonlinear control canonical normal form (NCCNF) when, for $y \in \mathcal{Y} \subseteq \mathbb{R}^n$,

$$\dot{y} = \begin{bmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ f(y) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ g(y) \end{bmatrix} u\,. \tag{4}$$

This definition of the NCCNF coincides with the definition found in literature on input-to-state feedback linearization of SISO nonlinear systems.

Accordingly, any input affine nonlinear system $\dot{z} = f_z(z) + G_z(z)u$, $z \in \mathcal{Z}$ can be represented by (4) as long as there exists a diffeomorphism $\mathrm{T} : \mathcal{Z} \mapsto \mathbb{R}^n$ s.t. $\mathcal{Y} = \mathrm{T}(\mathcal{Z})$ contains the origin, and where transforming the states according to $y = \mathrm{T}(z)$ results in the NCCNF representation [8]. In the following, it is assumed that there exists such a diffeomorphism, and that systems are given on the form (4).

We now show how such nonlinear control canonical normal forms enable performing the sought extension to the case $m < n$, through defining a nonlinear coordinate transformation based on the nontrivial equilibria of the original system.

Consider the following additional assumption, posed to guarantee the existence of an invertible steady state map for the state y_1 for systems represented as in (4):

Assumption 2 Considering system (4), $0 = f(y) + g(y)u$ is solvable with respect to y_1 and u.

Elaborating on the consequences of Assumption 2, it implies that the output equilibrium in (4) is $\bar{y} = [\bar{y}_1, 0, \ldots, 0]^{\intercal}$, i.e., only y_1 is non-zero at equilibrium. Since all the other states are zero at the equilibrium, the steady state from u to y_1 has to be a function of y_1 only, meaning that we may define

$$h_1^{-1}(y_1) := u = -\big(g(\bar{y})\big)^{-1} f(\bar{y})\,. \tag{5}$$

Moreover, Assumption 2 implies that (5) is bijective, allowing us to define the mapping

$$x_1 := h_1^{-1}(y_1) \tag{6a}$$
$$y_1 := h_1(x_1)\,. \tag{6b}$$

When the mapping is not bijective we may not find a closed form solution to the inversion, and so methods for locally approximating the inverse will likely have to be used. This is outside the scope of this work.

In addition to (6), for the purpose of extending the sought separation property to the case $m < n$, we define a transformation for each of the remaining states s.t.

$$h^{-1}(y) := \left[h_1^{-1}(y_1), h_2^{-1}(y), \ldots, h_i^{-1}(y)\right]^{\intercal}$$
$$h(x) := \left[h_1(x_1), h_2(x), \ldots, h_i(x)\right]^{\intercal}\,.$$

Then, since y_1 is the only state with non-zero equilibria, we may freely choose the remaining maps $h_i^{-1}(y)$, $h_i(x)$ as long as they satisfy

C1) $h_i^{-1}(\bar{y}) = 0$ and $h_i(\bar{x}) = 0$, and
C2) $h^{-1}(y)$ is invertible (and the inverse is $h(x)$) .

Given this, we define the transformation as:

Definition 3

$$x_i := h_i^{-1}(y) = \begin{cases} h_1^{-1}(y_1) & i = 1 \\ \sum_{j=1}^{i-1} \frac{\partial h_{i-1}^{-1}}{\partial y_j} y_{j+1} & i = 2, \ldots, n \end{cases} \tag{7a}$$

$$y_i := h_i(x) = \begin{cases} h_1(x_1) & i = 1 \\ \sum_{j=1}^{i-1} \frac{\partial h_{i-1}}{\partial x_j} x_{j+1} & i = 2, \ldots, n\,. \end{cases} \tag{7b}$$

Note that this transformation is defined such that $x_{i+1} = \dot{x}_i$, seen from e.g.

$$x_2 := h_2^{-1}(y) = \frac{\partial h_1^{-1}}{\partial y_1} y_2 = \frac{\partial h_1^{-1}}{\partial y_1} \dot{y}_1 = \frac{d}{dt} h_1(y) = \dot{x}_1\,.$$

This ensures that the relation between states are preserved through the transformation, as then both the linearization and wienerization of (4) result in models that are on NCCNF.

To show this, we start by noting that Definition 3 trivially satisfies C1) above. Moreover, C2) may be proven using the implicit function theorem, i.e., by ensuring that $\nabla h^{-1}(y)$ is non-singular for all $y \in \mathcal{Y}$. The latter requires the eigenvalues of the Jacobian matrix to never be zero. Computing the Jacobian using definition (7b) leads to

$$\nabla h^{-1}(y) = \begin{bmatrix} \frac{\partial h_1^{-1}}{\partial y_1} & 0 & 0 & \cdots & 0 \\ \frac{\partial h_2^{-1}}{\partial y_1} & \frac{\partial h_2^{-1}}{\partial y_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \frac{\partial h_{n-1}^{-1}}{\partial y_1} & \frac{\partial h_{n-1}^{-1}}{\partial y_2} & \cdots & \frac{\partial h_{n-1}^{-1}}{\partial y_{n-1}} & 0 \\ \frac{\partial h_n^{-1}}{\partial y_1} & \frac{\partial h_n^{-1}}{\partial y_2} & \cdots & \frac{\partial h_n^{-1}}{\partial y_{n-1}} & \frac{\partial h_n^{-1}}{\partial y_n} \end{bmatrix},$$

where the matrix structure emerges from the fact that each subsequent h_i^{-1} will be a function of one more state than the last. This is a lower triangular matrix, and so its eigenvalues are given by the elements along the diagonal. From (7b) we may also deduce that

$$\frac{\partial h_i^{-1}}{\partial y_i} = \frac{\partial h_1^{-1}}{\partial y_1}$$

regardless of the index i, thus the Jacobian is non-singular as long as $\frac{\partial h_1^{-1}}{\partial y_1} \neq 0$ for all $y \in \mathcal{Y}$, which holds by Assumption 2. Consequently, the inverse of $h^{-1}(y)$ is well defined and it is trivial to see that it is $h(x)$.

To summarize, under Assumption 2 it is possible to define a transformation based on the equilibrium map as in (7). Next we show how this transformation enables extending the structure in (2) so that it is possible to separate static and dynamic nonlinearities for systems with $\dim u < \dim y$.

C. Separating static and dynamic nonlinearities

Before presenting the separated system, we define an invertible matrix T for some equilibrium point $\bar{x}$

$$T := \nabla h(\bar{x}) \quad \in \mathbb{R}^{n \times n} \tag{8}$$

Remark 4 We note that T in the following can be any invertible matrix, analogous to (2). However, for reasons discussed in Section III, we will limit our discussion to the case where T is defined as in (8).

Lemma 5 Consider system (4). Let Assumption 2 hold, $h(\cdot)$ be as in Definition 3, and the corresponding T be as in (8). Then, transforming y in (4) by means of $y = h(T^{-1}x)$ results in the system structure

$$\dot{x} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ \widetilde{f}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \widetilde{g}(x) \end{bmatrix} u \tag{9a}$$

$$y = h(T^{-1}x) \tag{9b}$$

where

$$\widetilde{f}(x) = \frac{h_1'(T^{-1}\bar{x}_1)}{h_1'(T^{-1}x_1)} \left(f(h(T^{-1}x)) - \sum_{j=1}^{n-1} \frac{\partial h_n}{\partial x_j} x_{j+1} \right) \tag{9c}$$

$$\widetilde{g}(x) = \frac{h_1'(T^{-1}\bar{x}_1)}{h_1'(T^{-1}x_1)} g\big(h(T^{-1}x)\big) \tag{9d}$$

Moreover, this representation separates dynamic and static nonlinearities, i.e. the static gain from u to x is constant.

Proof: To show that (9a) indeed holds, observe that

$$\dot{y} = \frac{d}{dt}\big(h(T^{-1}x)\big) = \nabla h(T^{-1}x) T^{-1} \dot{x}$$

Note that since $\bar{x} = [\bar{x}_1, 0, \ldots, 0]^\intercal$ and x_j is a factor in $\frac{\partial h_i}{\partial x_j}$ for $i \neq j$ (seen by expanding and differentiating (7b)), we have that T is a diagonal matrix with $h_1'(\bar{x}_1)$ in every diagonal element. Furthermore, since $\nabla h(x)$ is lower triangular, we may use forward substitution to solve the above equation for $\dot{x}_i$. When $i < n$ we get

$$\begin{aligned} \dot{x}_i &= \Big(\nabla h(T^{-1}x)T^{-1}\Big)_{ii}^{-1} \left(\dot{y}_i - \sum_{j=1}^{i-1} \frac{\partial h_i}{\partial x_j} \dot{x}_j \right) \\ &= \frac{h_1'(\bar{x}_1)}{h_1'(x_1)} \left(h_{i+1}(x) - \sum_{j=1}^{i-1} \frac{\partial h_i}{\partial x_j} x_{j+1} \right) \\ &= \frac{h_1'(\bar{x}_1)}{h_1'(x_1)} \left(\frac{h_1'(x_1)}{h_1'(\bar{x}_1)} x_{i+1} + \sum_{j=1}^{i-1} \frac{\partial h_i}{\partial x_j} x_{j+1} - \sum_{j=1}^{i-1} \frac{\partial h_i}{\partial x_j} x_{j+1} \right) \\ &= x_{i+1} \end{aligned} \tag{10}$$

where we have omitted T^{-1} from the function arguments in the last three lines for ease of notation. When $i = n$, we may rewrite (10) to get

$$\dot{x}_n = \frac{h_1'(\bar{x}_1)}{h_1'(x_1)} \left(f(h(x)) + g(h(x))u - \sum_{j=1}^{n-1} \frac{\partial h_n}{\partial x_j} x_{j+1} \right)$$

again omitting T^{-1} in the function arguments.

To see that (9a) indeed has equilibrium $\bar{x}_1 = T_{11}\bar{u}$ (where the subscript denotes the row-column index) we consider $\dot{x}_n = 0$. Using that $x = Th^{-1}(y)$, the definition of h_1^{-1} and T, and considering the equilibrium $(\bar{u}, \bar{x}, \bar{y})$, we get

$$\widetilde{f}(\bar{x}) = f(\bar{y}_1)$$
$$\widetilde{g}(\bar{x}) = g(\bar{y}_1)$$

so that

$$\bar{u} = g(\bar{y}_1)^{-1} f(\bar{y}_1) = h_1^{-1}(\bar{y}) = h_1^{-1}\Big(h_1\left(T^{-1}\bar{x}\right)\Big) = \frac{1}{T_{11}}\bar{x}_1$$

■

Under Lemma 5, the equation (9a) only explains dynamical behaviour of the original system, while the static information is completely contained in the transformation (9b).

III. USING STRUCTURE (9) FOR CONTROL DESIGN PURPOSES

We now demonstrate that structure (9) enables improving closed loop performances by means of moving from linear controllers to nonlinear Wiener controllers.

A. *Linear Output Feedback Control*

A classical strategy to controlling system (4) is by using a linear output feedback controller

$$u = -K_l y + N_l y_r \tag{11}$$

where K_l is the controller gain, y_r is the reference we wish to track, and N_l is a matrix that is chosen such that the steady state tracking error is ideally zero with respect to some opportune linearization of the nonlinear system.

Such a linearization may be either identified from data or provided through linearizing (4) around a specific equilibrium point $(\bar{x}, \bar{u})$. We denote such a linear model by

$$\dot{y}_l = Ay_l + Bu \ . \tag{12}$$

Assuming w.l.o.g. that the equilibrium is $(\bar{u}, \bar{y}) = (0, 0)$, and assuming that (12) has been obtained through linearizing a model whose structure is (4), it follows that

$$A := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \dfrac{\partial f}{\partial y_1} & \dfrac{\partial f}{\partial y_2} & \dfrac{\partial f}{\partial y_3} & \dots & \dfrac{\partial f}{\partial y_n} \end{bmatrix}_{y=\bar{y}} \quad B := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ g(\bar{y}) \end{bmatrix} . \tag{13}$$

This is a controllable linear model, since it is in a linear control canonical form. This in turns allows us to define the controller gain K_l so that the closed loop system matrix

$$A_l = A - BK_l$$

is not only Hurwitz but also with eigenvalues chosen according to some criterion on the desired closed loop performance.

Following this approach to design a closed loop controller for the nonlinear system (4) is rather immediate, since it mimics classical output feedback strategies, and it results in a controller whose closed loop performances are ideally as specified when designing K_l and N_l for a reference that is kept close to the linearization point.

However, if reference tracking is sought, then one should expect, in general, that the trajectory of the output will deviate from the desired one, as the static gain of the original system is generally nonlinear.

B. *Wiener Control*

Another strategy to control (9) is by means of the linear output feedback controller

$$u = -K_w Th^{-1}(y) + N_w Th^{-1}(y_r) \tag{14}$$

where again K_w is the controller gain and N_w is a matrix that is chosen such that the tracking error is nominally zero in steady state conditions. Note that although this controller is linear in the states, the controller itself is in practice a *nonlinear* one, since it is based on applying the (in general) nonlinear static "block" h^{-1}. Still, the procedure to design the controller is similar to designing a linear output feedback controller (indeed K_w may be designed according to some opportune linearization of the original model).

We now note how the model structure (9) facilitates linearizing (9a) around the equilibrium point $(\bar{u}, \bar{x}) = (\bar{u}, Th^{-1}(\bar{y}))$. Indeed, assuming w.l.o.g. $(\bar{u}, \bar{x}) = (0, 0)$ as before, and choosing T as in (8) immediately yields the Wiener model

$$\dot{x}_w = Ax_w + Bu \tag{15a}$$
$$y_w = h\left(T^{-1}x_w\right) \tag{15b}$$

where A and B are the same as defined in (13), owing mainly to the fact that $\bar{x}_i = 0$, $\forall\, i > 1$, resulting in $\frac{\partial \widetilde{f}}{\partial x_i} = \frac{\partial f}{\partial y_i}$ and $\widetilde{g}(\bar{x}) = g(\bar{y})$.

The overall combination of separating the system into static and dynamic nonlinear components as in (9), together with the linearization as above, may be referred to as a *wienerization* [2].

In the following we let the subscript w indicate a model that is in fact a Wiener approximation of the original nonlinear system. Thus, the feedback control approach proposed in this section may equivalently be called *Wiener output feedback control*, as discussed in [10]. Once again, due to the controllable form of the linear model, we may define K_w so that the closed loop system matrix

$$A_w = A - BK_w$$

is Hurwitz, with eigenvalues chosen according to some criterion on the desired tracking performance. We regard the Wiener output feedback control and the linear output feedback control as comparable when the performance criterion is the same for both models.

C. *Comparing the structures of controllers* (11) *and* (14)

The two control strategies (11) and (14) are structurally similar, with the unique difference being whether we design the controller for x (since $Th^{-1}(y) = x$) or for y.

At the same time, the Wiener output feedback controller is based on a model that does not approximate the static nonlinearity of the system. In a sense, and laddering on the intuitions developed in the first part of the paper, the model structure (15) includes a globally accurate representation of the equilibrium map, something that a linearized version of the original system does not have. This raises the following ansatz: *provided that the reference signal is slow enough, controlling the nonlinear plant via the Wiener control strategy should lead to better closed loop performance than with the linear one*. This ansatz is based on the intuition that the Wiener output feedback control provides a smaller simulation error, and the system will behave closer to what is desired.

The formal investigation of the validity of such an ansatz may be made by opportune error bounds on the control of the original system (4) under Wiener feedback, using the idea of "slowly varying systems" [8], similar to what is proved in [2]. This is though beyond the scope of this paper, and is left as an open research question.

D. Stability of the Wiener simulation error

The discussions above formalize how the Wiener model approximates the dynamics of the original system, while retaining the static parts (that will therefore in the remainder of the paper be assumed to be exact). To analyze the simulation error induced by the fact that the dynamics are approximated, consider the Wiener model error

$$e_w(t) = y(t) - y_w(t)\,. \tag{16}$$

As shown in Section II-C, $y(t)$ and $y_w(t)$ have identical equilibrium u. When the original system is exponentially stable (in closed or open loop), so will the wienerized model be. Thus, y and y_w are converging and bounded, and converging to the same point. This implies that e_w also is convergent and bounded, i.e., exponentially stable to zero.

IV. Examples

In the following, we present two input affine nonlinear systems on which we demonstrate

1) the possibility of separating dynamic and static nonlinearities by transforming the systems from (4) to (9)
2) the possibility of designing an improved control system using the Wiener feedback control discussed above.

As for point 2) we utilize the two controllers considered in (11) and (14), with

$$\begin{aligned} K_l &= K_w = K \\ N_l &= N_w = -B^{\dagger}\left(A - BK\right) \end{aligned}$$

where $\dagger$ indicates the Moore-Penrose pseudo inverse. Moreover, K is chosen as the solution to the LQR problem

$$\begin{aligned} K &= B^{\intercal} P \\ A^{\intercal} P + PA - PBB^{\intercal}P &= -I_n\,. \end{aligned}$$

The performance is assessed using (16), when wienerization based control is used, and

$$e_l(t) = y(t) - y_l(t),$$

when linear control is used.

A. Example 1

As a first example, consider the NCCNF system

$$\dot{y} = \begin{bmatrix} y_2 \\ f(y) \end{bmatrix} + \begin{bmatrix} 0 \\ g(y) \end{bmatrix} u \tag{17a}$$

where

$$f(y) = \frac{y_1 y_2^2 - y_2\left(y_1^2+1\right) + 5\operatorname{asinh}\left(y_1\right)\left({y_1}^2+1\right)^2}{y_1^2+1} \tag{17b}$$

$$g(y) = y_1^2 + 1. \tag{17c}$$

1) Transforming the system: From (5) we find that

$$h_1^{-1}(y_1) := \operatorname{arcsinh}(y_1).$$

From Definition 3 we have

$$h^{-1}(y) = \begin{bmatrix} \operatorname{arcsinh}(y_1) \\ \dfrac{y_2}{\sqrt{y_1^2+1}} \end{bmatrix} \tag{18}$$

$$h(x) = \begin{bmatrix} \sinh(x_1) \\ \cosh(x_1)x_2 \end{bmatrix}, \tag{19}$$

which gives

$$\nabla h(x) = \begin{bmatrix} \cosh(x_1) & 0 \\ x_2\sinh(x_1) & \cosh(x_1) \end{bmatrix} \tag{20}$$

and $T = I_{2\times 2}$. Using this together with (19), (20), the definition of T, and Lemma 5, we may rewrite (17a) into the separated system representation

$$\dot{x} = \begin{bmatrix} x_2 \\ 5x_1\cosh(x_1) - x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \cosh(x_1) \end{bmatrix} u \tag{21a}$$

$$y = h\left(T^{-1}x\right) \tag{21b}$$

2) Control Performance: Linearization and wienerization of (17a) and (21) around $(\bar{u}, \bar{x}, \bar{y}) = (0, 0, 0)$ results in the two models

$$\dot{y}_l = Ay_l + Bu \quad \text{and} \quad \left\{ \begin{array}{ll} \dot{x}_w & = Ax_w + Bu \\ y_w & = h(T^{-1}y) \end{array} \right.$$

where

$$A = \begin{bmatrix} 0 & 1 \\ 5 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

To assess the model performance we simulate the original system and the two models using a reference $y_r = [5\sin(2\pi\omega t), 0]^{\intercal}$. We recall that the goal of this experiment is not for the systems to track the references, but to assess if the system under the wienerization and linearization based controllers behave similarly to the Wiener and linear models with those controllers, respectively. Accordingly, Figures 1 and 2 present the norm of the errors $e_l(t)$ and $e_w(t)$ with this reference when $\omega = 0.2$ and $\omega = 0.02$, respectively.

We observe that the wienerization based controller is effective in improving the model-simulation error over using a linear controller for this particular system. Moreover, we see that the relative improvement increases as the frequency of the reference decreases, i.e., the reference is slower, which is as expected.

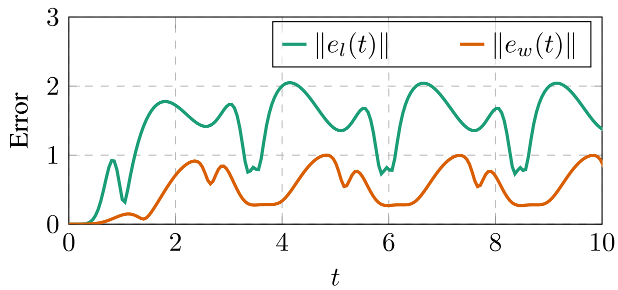

Fig. 1. Error associated to (17a) in feedback control with a sinusoidal reference.

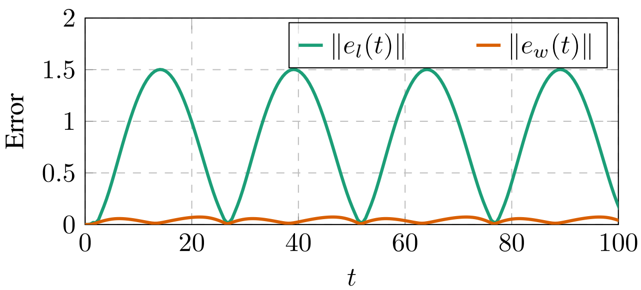

Fig. 2. Error associated to (17a) in feedback control with a slowly varying sinusoidal reference.

B. Example 2

In the following we revisit the mass-spring system from (3). Defining $y := [y_1, y_2]^\intercal$, we may write it as

$$\dot{y} = \begin{bmatrix} y_2 \\ \frac{1}{m}\left(-cy_2 - ky_1(1 + a^2y_1^2)\right) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \tag{22}$$

which one can recognise as a NCCNF.

1) Transforming the system: Seeing that $u = \frac{1}{m}\left(cy_2 + ky_1(1 + a^2y_1^2)\right)$ in equilibrium, we define $h_1^{-1}(y_1) := \frac{k}{m}y_1(1 + a^2y_1^2)$, which is invertible.

Using this with Definition 3, (8) and Lemma 5, we rewrite (22) into the separated system representation

$$\dot{x} = \begin{bmatrix} x_2 \\ \dfrac{m}{kh_1'(x_1)}\left(f\left(h(T^{-1}x)\right) - \dfrac{\partial h_2}{\partial x_1}x_2\right) \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{m}{kh_1'(x_1)} \end{bmatrix} u \tag{23a}$$

$$y = h\left(T^{-1}x\right). \tag{23b}$$

Due to the presence of the cubic polynomial in $h^{-1}(y)$, the resulting $h_1(x)$, $h(x)$, $\nabla h(x)$ and system transformation are too long and complex to write out in full. However, analytical solutions of these are found using Cardano's formula and (7).

2) Control Performance: Linearization and wienerization of (22) and (23) around $(\bar{u}, \bar{x}, \bar{y}) = (0, 0, 0)$ result in the two models

$$\dot{y}_l = Ay_l + Bu \quad \text{and} \quad \begin{cases} \dot{x}_w &= Ax_w + Bu \\ y_w &= h(T^{-1}y) \end{cases}$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -\dfrac{k}{m} & -\dfrac{c}{m} \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For this example we assess the performance by simulating the system and models using a reference $y_r = [\sin(2\pi\omega t), 0]^\intercal$. Figures 3 and 4 present the norm of the errors $e_l(t)$ and $e_w(t)$ corresponding to this reference when $\omega = 0.2$ and $\omega = 0.02$, respectively.

For this system observe that the wienerization based controller performs worse than the linear controller for the case when the reference has a higher rate of change, and that the performance is improved once the rate change of the reference is lower. This is an unfortunate side-effect of the wienerization procedure; the dynamic and static nonlinear components of the spring are somewhat similar in nature, resulting in the system being more linear-like than its individual parts when they are represented in a combined form (apparent by the involved expression for the dynamical part after transformation). That linearizing effect is lost during the separation in the wienerization, with the result being a model and controller that behaves worse in some scenarios. This is subject to future analysis. Nonetheless, the wienerization based controller is seen to be effective for a reference that is sufficiently slowly varying, as expected.

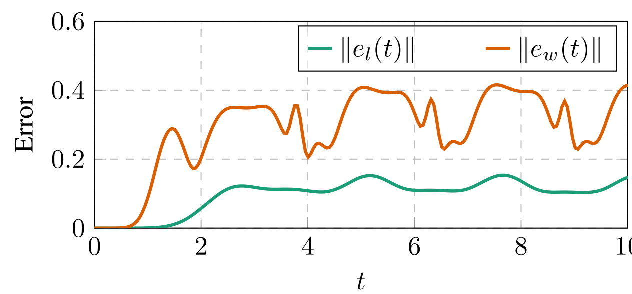

Fig. 3. Errors associated to the mass-spring system in feedback control with a sinusoidal reference.

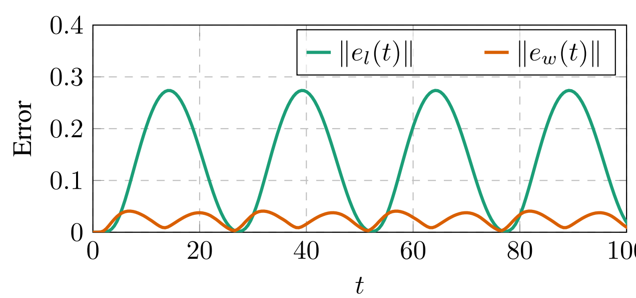

Fig. 4. Error associated to the mass-spring system in feedback control with a slowly varying sinusoidal reference.

V. Conclusion

This work presented an extension of existing concepts about how to separate static and dynamic behaviours in nonlinear systems, and obtained generalizations that allow performing such separations for a class of underactuated systems on nonlinear control canonical normal form (NCCNF). Underactuated systems generally do not have invertible equilibrium maps, and therefore do not immediately yield separated, wienerizable structures. We have characterized and clarified these issues, and shown a way to still separate dynamics and statics that in turn enables wienerization, and discussed what this means for state space representations. This leads to a new system theoretical approach to Wiener modeling that highlights related and more sophisticated model structures too.

Extending the wienerization further to even broader classes of systems and proving stability of the controller and modeling error is a high priority for future works. Moreover, since the model approximations proposed above extend the ideas of linearization and control schemes based on linearized models, we foresee Hammerstein models to be interesting candidates for similar approaches, putting this type of model approximation on more solid grounds.

REFERENCES

[1] S. I. Biagiola, O. E. Agamennoni, and J. L. Figueroa. H-infinity control of a Wiener-type system. *International Journal of Control*, 77(6):572–583, 2004.

[2] H. B. Bjørkøy, H. A. Engmark, A. Rasheed, and D. Varagnolo. Wienerization based control of nonlinear systems. In *2022 European Control Conference (ECC)*, pages 641–648. IEEE, 2022.

[3] S. Boyd and L. Chua. Fading memory and the problem of approximating nonlinear operators with volterra series. *IEEE Transactions on circuits and systems*, 32(11):1150–1161, 1985.

[4] Y. Fujimoto and T. Sugie. Kernel-based impulse response estimation with a priori knowledge on the DC gain. *IEEE Control Systems Letters*, 2(4):713–718, 2018.

[5] F. Giri and E.-W. Bai. *Block-oriented nonlinear system identification*, volume 1. Springer, 2010.

[6] A. Hagenblad. *Aspects of the identification of Wiener models*. Linköpings Universitet, PhD thesis, 1999.

[7] A. D. Kalafatis, L. Wang, and W. R. Cluett. Linearizing feedforward–feedback control of pH processes based on the Wiener model. *Journal of Process Control*, 15(1):103–112, 2005.

[8] H. K. Khalil. *Nonlinear systems, 3rd edition*. Prentice hall, Upper Saddle River, NJ, 2002.

[9] G. Mzyk. *Combined parametric-nonparametric identification of block-oriented systems*. Springer, 2014.

[10] D. Nesic. Output feedback stabilization of a class of wiener systems. *IEEE Transactions on Automatic Control*, 45(9):1727–1731, 2000.

[11] S. J. Norquay, A. Palazoglu, and J. Romagnoli. Model predictive control based on wiener models. *Chemical Engineering Science*, 53(1):75–84, 1998.

[12] S. J. Norquay, A. Palazoğlu, and J. A. Romagnoli. Nonlinear model predictive control of pH neutralisation using Wiener models. *IFAC Proceedings Volumes*, 29(1):5829–5834, 1996.

[13] G. Pajunen. Adaptive control of wiener type nonlinear systems. *Automatica*, 28(4):781–785, 1992.

[14] J. C. Schulze, D. T. Doncevic, and A. Mitsos. Identification of MIMO Wiener-type Koopman Models for Data-Driven Model Reduction using Deep Learning. *arXiv preprint 2201.12669*, 2022.

[15] B. Sohlberg and E. W. Jacobsen. Grey box modelling–branches and experiences. *IFAC Proceedings Volumes*, 41(2):11415–11420, 2008.

[16] J. Wang, M. Wei, and X. Xing. Static gain estimation for nonlinear dynamic systems from steady-state values hidden in historical data. *ISA transactions*, 120:78–88, 2022.

[17] A. Wills, T. B. Schön, L. Ljung, and B. Ninness. Identification of hammerstein–wiener models. *Automatica*, 49(1):70–81, 2013.

[18] M. Zeitz. Canonical forms for nonlinear systems. In *Nonlinear Control Systems Design 1989*, pages 33–38. Elsevier, 1990.